# A systematic review of non-invasive manual therapy in the management of rheumatoid arthritis patients

By

Kyle Donald Sclanders

Dissertation submitted in partial compliance with the requirements for the Master's Degree in Technology: Chiropractic

Durban University of Technology

I, Kyle Donald Sclanders, do declare that this dissertation is representative of my own work in both conception and execution.

23/11/2021

Kyle Donald Sclanders — Date

BSc (Stell), M.Tech: Chiropractic

Final Submission

23/11/2021

Dr C.M. Korporaal — Date

M.Tech: Chiropractic, CCFC, CCSP, ICSSD

# DEDICATION

I dedicate this dissertation to my parents (Owen and Jean) and to my inspiration for this study, Taryn Purchase.

In adding to the support, I have from family and friends, a poem by Berton Braley helped focus me through the ebbs and flows of writing this dissertation:

"If you want a thing bad enough
To go out and fight for it,
Work day and night for it,
Give up your time and your peace and your sleep for it

If only desire of it
Makes you quite mad enough
Never to tire of it,
Makes you hold all other things tawdry
and cheap for it

If life seems all empty and useless without it
And all that you scheme and you dream is about it,

If gladly you'll sweat for it,
Fret for it, Plan for it,
Lose all your terror of God or man for it,

If you'll simply go after that thing that you want.
With all your capacity,
Strength and sagacity,
Faith, hope and confidence, stern pertinacity,

If neither cold poverty, famished and gaunt,
Nor sickness nor pain
Of body or brain
Can turn you away from the thing that you want,

If dogged and grim you besiege and beset it,
You'll get it!"

~ Berton Braley

## ACKNOWLEDGEMENTS

I would like to thank my parents, Owen, and Jean, for being the unshakable support that has always put their children's best interest above all else. Without your dedication, selflessness, generosity and sacrifice I would have never had the opportunities I have had and been able to pursue my passions in life. You are the perfect example of what a nurturing and supportive family environment can have on a child.

To my friends, I am finally done… you can stop stressing! Thank you for the support and encouragement to finish my studies and start working. Your pressure was real but so necessary. Thank you.

To my inspiration for this study, Taryn Purchase, you have been a close friend and support. Thank you for allowing me to see the impact Rheumatoid Arthritis has had on your life so that I could understand your experiences and the difficulties you deal with on a daily basis, which drove my interest in this field and pushed me to look for avenues to assist you and others who suffer like you. I have learnt so much both in writing this dissertation and in our everyday interaction. I hope I have contributed in some way towards therapeutic efforts in the management of RA.

To my Supervisor, Dr Charmaine Korporaal, I greatly appreciate all your guidance and generous contribution to this dissertation. The complexity of the methodology is not something I tackle alone and without your input I don't think I would have reached the end point. In addition to your supervisory role, you have provided great mentorship over the last six years and offered me many opportunities to better myself both clinically and personally. I hope that as I entre this new chapter as a practitioner I can carry myself in the way you do and provide the patient care and management that you have put so much time and effort into teaching me. Thank you for the part you played in shaping me as a chiropractor.

To all my reviewers both locally and internationally, thank you for your willingness to contribute to my dissertation, and for taking time away from your busy schedules and patients to assist me. It has been an honour to collaborate with you and I am deeply appreciative of the benefit of your insight.

To Mrs Gillian Parker, thank you for your time taken to edit and proof-read my dissertation. Your incredible grasp on the English language is greatly appreciated and admired.

To Mrs Avenal Finlayson, thank you for your willing assistance with my systematic search for articles and citations.

# ABSTRACT

**Background:** Rheumatoid arthritis (RA) is a common, chronic, and progressive autoimmune condition manifesting in a polyarticular pattern, with an agnogenic pathophysiology. The progressive effect on physical function (e.g., pain, quality of life, funcaitonal impairment and disability) and the side effects of pharmaceutical interventions, have led to patients seeking adjunctive therapy. Complementary and alternative medicine (CAM) therapy, is thought to reduce pain, increase range of motion, and improve quality of life. Yet, the evidence base for CAM therapies is vague. This systematic review of studies relating to non-invasive manual therapy (vis. exercise, manipulation, mobilisation, and massage), determined the level of evidence available for their use, and provided a summation of the available evidence to support evidence-based clinical practice. Finally, the study provides recommendations for further investigations.

**Method:** This mixed methods systematic review complying with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol and Cochrane Collaboration guidelines, required an electronic literature search (viz: Google Scholar and DUT Summons) using key terms including rheumatoid arthritis and variable combination one or more of the following: treatment, management, chiropractic, osteopathy, physiotherapy, physical therapy, massage, mobilization, manipulation, joint oscillation, and exercise. This along with hand search strategies and reviewer input produced 65 articles that were incorporated into the review (40 RCTs. 15 nRCTs, and 10 Cases). All selected articles were available in English in electronic format and represented RCTs, nRCTs and CR/CS (not limited to peer reviewed studies) and reflected interventions applied to adult, human subjects. The articles were reviewed by 13 reviewers, using as appropriate the PEDro, Newcastle-Ottawa or Liddle scales, to evaluate the methodological rigour (internal validity) of the studies, while qualitatively contextualizing these outcomes (external validity) to determine the impact of bias on the reported clinical outcomes. Thereafter an analysis of evidence per modality was completed utilising published grading systems.

**Results:** The above evaluations and aggregated evidence available was then determined for the modalities. It was found that there is limited evidence available for manipulation, mobilisation, massage, and exercise in the treatment of the RA patient. The notable exceptions were the use of aerobic exercise with acetaminophen; exercise as part of an intervention programme in the treatment of the hands; and mixed exercise programme to improve bone mineral density and attaining muscular strength gains.

**Conclusion:** Thus, excluding the exceptions, the reviewed interventions should be used cautiously in practice and patients should not be provided with expectations that are not confidently supported by available literature. The evidence suggests that for the most part, intervention by means of manipulation, mobilisation, soft tissue manipulation and massage, along with exercise, requires further research efforts to provide sufficient high-quality evidence for the practitioner to implement in clinical practice with confidence.



**Key terms:** Manual therapy, Arthritis, rheumatoid, Systematic review

# TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## DEFINITIONS:

For purpose of consistency, where definitions are relevant to papers reviewed in this study. They have been quoted in this list accordingly.

| | |
|---|---|
| **Absenteeism** | You are not present at the workplace and are unproductive |
| **ACR response criteria** | A criteria by the American college of rheumatology used to define response rate to therapy. |
| **Adaptive immune system** | the deviations of the immune system responsible for acquired immunity. |
| **Adhesive capsulitis** | Also known as frozen shoulder. A painful restrictive shoulder condition |
| **Agnogenic** | Unknown origin or idiopathic. |
| **Amyloidosis** | In this context serum amyloid A. A condition relating to the abnormal build-up of the protein amyloid A, resulting in pathology. |
| **Auto-immune diseases** | "These diseases are characterized by the body's own immune system attacking self-antigens. This results from changes in the processes of central and peripheral immune tolerance" (Valesini *et al*. 2015). |
| **Boutonniere deformities** | Also known as buttonhole deformities. "Presenting as an abnormal flexed position of the proximal interphalangeal joint with a hyperextension of the interphalangeal joint" (Cima et al., 2013). |
| **Carpal tunnel** | A common condition resulting from the compression of the median nerve as it passes through the wrist in the carpal tunnel. |
| **Chiropractors** | "These are practitioners that treat neuro-muscular and musculoskeletal complaints through adjusting muscles, tendons and joints and using manipulation and massage techniques" (Cooper, 2013). |
| **Citrullination** | The mechanism by which there is a post-translation modification of the protein-bound amino acid arginine into citrulline. |
| **Cock-up toe deformities** | A form of hammertoe, whereby the proximal fifth phalanx of the foot deviates from it natural position and articulates at a 90-degree angle to the metatarsal. |
| **De Quervain's** tenosynovitis | A painful condition resulting from the inflammation of the tendons of the extensor pollicis brevis and the abductor pollicis longus. |
| **Felty's syndrome** | A syndrome comprising of rheumatoid arthritis, splenomegaly, and neutropenia |
| **Innate immune system** | the deviations of the immune system responsible for first response immune reaction. |
| **Kaltenborn mobilization** | "This is a three-tiered technique that incorporates small amplitude movements with traction to loosen and counteract any compressive forces on the joint (grade I), to gentle mobilization of joints with oscillations without overriding any tightening of the connective tissues surrounding the joint (grade II), or oscillations followed by additional stretching (grade III)" (Levitsky et al 2019). |

| | |
|---|---|
| **Non-invasive therapy** | For the purposes of this thesis, the term “non-invasive therapy” has been used as an umbrella term under which the terms “non-medicinal interventions” and “manual therapies” are included. |
| **Manipulation** | “Commonly referred to as “an adjustment” by chiropractors, is another form of manual therapy which involves a joint positioned at its restrictive barrier and then the practitioner applies a quick, short impulse (HVLA) thrust” (Chung 2015). |
| **Manual oscillations** | “A technique defined as rhythmical passive mobilisations of the joint with continuously equal amplitude” (Dhondt et al 1999) |
| **Mobilizations** | “This involves positioning a joint at its restrictive barrier (i.e., end of passive range of motion) and then the practitioner uses a series of gentle, repetitive movements towards and through the restriction” (Chung 2015). |
| **Osteopaths** | “These are practitioners that use a system of diagnosis and treatment, usually by manipulation, which mainly focuses on musculo-skeletal problems, though some branches aim to treat a wider spectrum of disorders” (Cooper, 2013). |
| **Pack years** | A value utilised to define the how many cigarettes a patient has smoked during over their lifetime. |
| **Presenteeism** | You are at the workplace, but you are not productive. |
| **Rheumatoid cachexia** | “A loss of body cell mass which predominates in skeletal muscle… usually characterised by stable bodyweight as the decrease in muscle mass is masked by a concomitant increase in fat mass” (Cooney et al. 2011). |
| **Swan-neck deformities** | “Hyperextension of the proximal interphalangeal joint with flexion of distal interphalangeal joint” (Cime et al. 2013). |
| **Treat-to-target principles** | the practice of therapeutic management of a specific condition or disease by regular assessment of the disease status a regular modification according to these assessments. |
| **Ulnar deviations** | “Deviation of the metacarpophalangeal joints” of the wrist toward the ulnar (medially) (Cima et al. 2013). |
| **Z-deformities** | “Flexion of the metacarpophalangeal joint and hyperextension of the interphalangeal joint” of the thumb (Cima et al. 2013). |

## ABBREVIATIONS:

| | | |
|---|---|---|
| A | ACPA | Auto-Citrullinated protein antibodies / anticyclic citrullinated peptide antibodies |
| | ACR | American College of Rheumatology |
| | ACR20/50/70 | America College of Rheumatology response criteria |
| | ADL | Activities of daily life |
| | AGREE | Appraisal of Guidelines for Research and Evaluation |
| | αvβ3 | Alpha-v-beta-3-integrin |
| B | B-cell | B lymphocyte |
| | BMI | Body Mass Index |
| C | CAM | Complementary and alternative medicine |
| | CCP | Cyclic-citrullinated peptide |
| | CD | Cluster of differentiation |
| | CD196 / CCR6 | Chemokine receptor 6 protein which in humans is encoded by the CCR6 gene |
| | CD4+ | T helper cell |
| | CD8+ | T cytotoxic cell |
| | CI | Confidence interval |
| | CNS | Central nervous system |
| | CONSORT | Consolidated Standards of Reporting Trials |
| | CQ | Chloroquine |
| | CRD | Centre for reviews and dissemination |
| | CRP | C-reactive proteins |
| | CS/CR | Case study / series / report |
| | CSA | Clinically suspected arthritis |
| | CVD | Cardiovascular disease |
| D | DALY | Disability-adjusted life year |
| | DAMPs | Damage-associated molecular patterns |
| | DAS (-28) | Disease Activity Score (in 28 joints) |
| | DC | Dendritic cell |
| | DIP | Distal interphalangeal joints |
| | DMARD | Disease modifying anti-rheumatic drugs |
| | DNA | Deoxyribonucleic acid |
| | DRC | Department Research Committee |
| | DUT | Durban University of Technology |
| E | ESR | Erythrocyte sedimentation rate |
| | EULAR | European League Against Rheumatism |
| F | FBC | Full blood count |
| | Fc | Fc portion of immunoglobulin |

| | | |
|---|---|---|
| | FRC | Faculty Research Committee |
| G | GM-CSF | Granulocyte-monocyte-colony-stimulating factor |
| | GR | Glucocorticoid receptor |
| | GRADE | Grading of recommendations, assessment, development, and evaluations |
| H | HCQ | Hydroxychloroquine |
| | HDL | High-density lipoprotein (cholesterol) |
| | Hi-oxLDL | High-level oxidation low-density lipoprotein |
| | HLA | Human leucocyte antigen |
| | HLA-DRB1 | Class 2 classic human leucocyte antigen gene |
| | HVLA thrust | High velocity low amplitude thrust |
| I | IFN | Interferon |
| | IFN-α | Interferon - alpha |
| | IFN-β | Interferon - beta |
| | IFN-γ | Interferon - gamma |
| | IFN-type 1 | Interferon type 1 |
| | IFN- ε | Interferon type 1 - epsilon |
| | IFN-κ | Interferon type 1 - kappa |
| | IFN-ω | Interferon type 1 - omega |
| | IFN-type 2 | Interferon type 2 |
| | Ig | Immunoglobulin |
| | IL | Interleukins: -1, -1beta(β), -1RA, -2, -3 -4, -5 -6, -8, -10, -12, -13, -15, -17, -22 |
| | IL2RA | Interleukin-2 receptor alpha chain a protein (also called CD25) that in humans is encoded by the IL2RA gene |
| | iPR | Intracellular progesterone receptors |
| | IREC | Institutional Research Ethics Committee |
| J | JAK | Janus kinase |
| L | LDL | Low-density lipoprotein (cholesterol) |
| | Low-oxLDL | Low-level oxidation low-density lipoprotein |
| M | MAPK | Mitogen-activated protein kinase |
| | MCID | Minimal clinically important differences |
| | MCP | Metacarpal joints |
| | mDC | Myeloid dendritic cell |
| | MHC | Major histocompatibility locus<br>- MHC class II / type 2<br>- MHC class I / type 1 |
| | MMPs | Matrix metalloproteinase |
| | mPR | Progesterone membrane-associated receptors -alpha (α) / beta (β) or gamma (γ) |
| | MS | Multiple sclerosis |
| | MTP | Metatarsal phalangeal joint |
| | MTX | Methotrexate |

| | | |
|---|---|---|
| | MyD88 | Cytosolic adaptor myeloid differentiation factor 88 |
| | M1 and M2 | Macrophage types |
| | MoA | Memorandum of Agreement |
| N | NET | Neutrophil extracellular trap |
| | NK | Natural killer cells |
| | NLRP3 | Nucleotide-binding oligomerization domain, leucine rich repeat and pyrin domain-containing protein 3 |
| | NSAID | Non-steroidal anti-inflammatory drugs |
| | nRCT | Non-Randomised controlled study |
| | NOD | Nucleotide-binding oligomerization domain receptors |
| | NOS | Newcastle-Ottawa Scale |
| O | OR | Odds Ratio |
| | OS | Observational Studies |
| P | PA | Physical activity |
| | PADI | Protein-arginine deiminase |
| | PADI4 | Protein-arginine deiminase type-4 |
| | PAK | Progesterone-activated killer cells |
| | PAMP | Pathogen-associated molecule |
| | pDC | Plasmacytoid dendritic cell |
| | PEDro | Physiotherapy Evidence Database |
| | PGRNC1 | Progesterone receptor membrane component 1 |
| | PIP | Proximal interphalangeal joints |
| | Pr | Prolactin |
| | Prg | Progesterone |
| | PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| | PROSPERO | Prospective Register of Systematic Review |
| | PRR | Pattern recognition receptor |
| | PTP | Protein tyrosine phosphatase |
| | PTPN22 | Protein tyrosine phosphatase, non-receptor type 22, which is a protein that in humans is encoded by the PTPN22 gene |
| Q | QoL | Quality of life |
| R | RA | Rheumatoid arthritis |
| | RCT | Randomised controlled study |
| | RF | Rheumatoid Factor |
| | RNA | Ribonucleic acid |
| | ROS | Reactive oxygen species / oxygen radical |
| | ROM | Range of motion |
| S | SNP | Single nucleotide polymorphisms |
| | SR | Scavenger receptors |
| | SSZ | Sulfasalazine |

| | | |
|---|---|---|
| | STAT1 | Signal transducer and activator of transcription 1 |
| | STAT4 | Signal transducer and activator of transcription 4 |
| T | T-cell | T lymphocyte |
| | TGF- β | Transforming growth factor beta |
| | Th or CD4+ | T helper cell |
| | TLR | Toll-like receptors: -2, -3, -4, -7, -8, -9 |
| | TNF-α | Tumour necrosis factor – alpha |
| | TNFR | Tumor necrosis factor receptor |
| | TRAF1 | Tumor necrosis factor receptor associated factor 1, a protein encoded by this gene |
| | Treg | T regulatory cell |
| V | VECs | Vascular endothelial cells |
| | VIP | Vasoactive intestinal peptide |
| W | WHO | World Health Organization |

# CHAPTER 1: Introduction

## 1.1 Introduction to the topic of this research: reviewing evidence for Complementary and Alternative Medicine as an alternative approach to the treatment of Rheumatoid Arthritis.

Rheumatoid arthritis (RA) is a common, chronic and progressive autoimmune condition (Sharma *et al.* 2018) evolving from immune dysregulation (Ridgley, Anderson and Pratt 2018) that manifests primarily in the synovial joint in a polyarticular pattern, effecting the tendons, synovial tissue, bursae, and the articular bone and cartilage. The condition is characterised by one or more of the following pathognomonic signs and symptoms: morning stiffness, symmetrical arthritis of three or more joint areas (typically involving the hand joints), rheumatoid nodules, serum rheumatoid factor (RF) (positive) and specific radiological findings (e.g., juxta-articular osteopenia, reduced joint spacing and periarticular soft tissue swelling) (Brahee, Pierre-Jerome and Kettner 2003; So *et al.* 2003; Rindfleisch and Muller 2005; Longmore *et al.* 2014).

Inflammation is the key symptom of RA pathophysiology. Controlling and treating inflammation is therefore a major therapeutic target pharmaceutically (Smolen, Aletaha and McInnes 2016). The main goals of pharmaceutical treatment are to prevent or control joint damage and loss of function whilst reducing pain with the goal of remission (Christie *et al.* 2007). The pharmaceutical interventions include a combination of one or more of the following: non-steroidal anti-inflammatory drugs (NSAIDs), disease-modifying anti-rheumatic drugs (DMARDs) and steroids (Emery 2006).

NSAIDs primarily focus on pain control and stiffness in order to improve the physical function of the patient. Unfortunately, they do not aid in the modification of the disease, nor do they protect against structural damage. Steroids (such as glucocorticoids) are primarily used for the rapid control of symptoms, but also have a role in disease modification (Smolen, Aletaha and McInnes 2016). As for DMARDs, they target inflammation and attempt to control the structural damage synonymous with the condition. There are three main groups of DMARDs:

- conventional synthetic DMARDs (methotrexate falls into this group and is considered the gold standard for treatment of RA),
- biological DMARDs (such as, TNF inhibitors, Anti-B-cell, Anti-T-cell co-stimulation and Anti-IL 6R); and
- targeted synthetic DMARDs (such as Janus kinase inhibitors) (Smolen *et al.* 2007; Smolen, Aletaha and McInnes 2016).

These drug groups and many other pharmaceuticals used in treatment of RA are part of an ever-growing catalogue of new drugs being developed in the pursuit of a safer profile, as well as fewer side effects, a reduced toxicity and a greater chance of remission (Smolen, Aletaha and McInnes 2016).

The many limitations of pharmaceutical management of RA fall into three areas, which are listed and discussed individually below:

**1. Treatment goals not met:**

Treatment goals include the therapeutic (pharmaceutical) efforts to achieve stringent remission and at the very least suppress the disease to a low activity. Regrettably, many RA patients never achieve remission, while those patients who achieve remission need to continue therapy to maintain it. This implies the need for new therapeutic agents (Smolen, Aletaha and McInnes 2016). If the treatment goals of remission or low disease activity are not substantially met by the sixth month of therapy, the protocol needs to be reviewed. It is important to balance the choice of treatment protocol against various factors relating to the patient as well as the risks associated with the agents used (Smolen, Aletaha and McInnes 2016).

**2. Co-morbidities:**

The patient's co-morbidities could influence the treatment options and must therefore be taken into consideration. For example, Rituximab is the drug of choice when the patient has multiple sclerosis (MS) as a co-morbidity because of its efficacy in treating RA and MS. By contrast, the use of TNF inhibitors to treat RA may elicit flare-ups of MS (Smolen, Aletaha and McInnes 2016).

**3. Reduced efficacy of medications and lack of compliance:**

Over time patients may become less responsive to the treatment regimen implemented. The reason for this is still not fully understood, though it has been proposed that lack of adherence to the regimen and even immunogenicity may play a role (Pascual-Salcedo *et al.* 2011; Smolen, Aletaha and McInnes 2016). In either case a loss of efficacy may develop. This suggests that treating a patient with RA effectively requires regular evaluation of disease activity versus drug dosage. These evaluations should drive decisions relating to adapting the course of therapy (Smolen, Aletaha and McInnes 2016). It should be noted that lack of compliance to the treatment regimen involves many factors. Side effects and the toxicity of the agents used are reported as being strong influencers in this regard (Breedveld and Kalden 2004).

Notwithstanding the above limitations of the pharmaceutical management of RA, RA patients are required to consume significant amounts of medication with an eye to controlling the progression of their condition. All such medications carry with them significant side effects (both positive and negative) and toxicity (Breedveld and Kalden 2004). In addition, despite the use of pharmaceutical therapy, RA progresses in a linear fashion over a patient's life, contributing to increased morbidity and premature mortality. The result is a significant decrease in the length and quality of life of RA patients' (QoL), which forces them to adapt their daily activities accordingly (Breedveld and Kalden 2004; Malm *et al.* 2017). Furthermore, research reveals that pharmaceutical interventions place a significant burden on patients, with the annual medical and disability-related costs not

only escalating but correlating closely with the degree of disability resulting from RA (Breedveld and Kalden 2004; Malm *et al.* 2017).

While pharmaceutical management as discussed above is an option for RA patients, the treatment strategy of Complementary and Alternative Medicine (CAM) offers other possibilities. CAM is a system of diverse medical practices and products excluded from what is regarded as orthodox medicine insofar as they are not offered in the standard medical school curriculum and therefore are not taught as interventions to use in the standard care of RA patients (Fernandez-Llanio Comella, Fernandez Matilla and Castellano Cuesta 2016; Zhao *et al.* 2017). It is necessary to take into account the fact that 33% of all RA patients in the United States used a form of CAM in 2013 while an estimated range of 28 – 90% of RA patients have done so at least once in their lives (Tokem *et al.* 2014). This indicates that a high number of patients with rheumatological conditions tend to seek or try out practitioner-based CAM care (Phang *et al.* 2018). CAM is thought to assist with reducing pain, thus reducing the need to increase the dosage of prescription medication with its various negative side-effects (for example, a reduction of painkillers means a reduction of such side-effects). In addition, it is seen as helping to increase the patients' range of motion (ROM), the benefits being an improvement in the patient's QoL as a result of more exercise, enhanced strength, reduced muscle-wasting and greater independence (Moquin *et al.* 2009; Phang *et al.* 2018; Della-Peruta 2019).

CAM intervention/s may also help with treatment compliance (Zhao *et al.* 2017) as the reduced pain and increased ROM enable the patient to hold gainful employment, which, along with the reduced use of medication, will lessen the financial burden on the patient, the patient's family and the health care system (Della-Peruta 2019). Concurrently, the improvement in the patient's QoL, physical condition and independence will decrease the social burden the patient places on others (Chen and Michalsen 2017). The appointments with the CAM practitioner also provide additional clinical evaluation and monitoring points (Della-Peruta 2019) which aid in assessing and adjusting the treatment, including the pharmaceutical regimen (Della-Peruta 2019), especially if the practitioners engage in co-management (Della-Peruta 2019). This requires referral between CAM practitioners and other relevant medical practitioners (such as orthopaedic surgeons, occupational therapists, physiotherapists, biokineticists, rheumatologists and psychiatrists) for co-management purposes in order to enable and maintain the best QoL for the patient (Maravic *et al.* 2000; Chandrashekara 2011; Briggs *et al.* 2013).

Implementing the co-management model of patient care requires that the various practitioners involved collaborate in an evidence-informed, patient-centred practice model for the patient to achieve the best possible clinical outcomes through receiving the most appropriate care at the best time within the pathogenesis of the RA progression (Smith *et al.* 2013; Brosseau *et al.* 2014).

A small body of research within the context of RA and CAM interventions for RA patients has produced a negative response to the above approach as the research is considered compromised by several authors due to:

- Contrasting scientific data on the efficacy, effectiveness and safety of CAM (Furlan *et al.* 2012).
- Limited numbers of randomised controlled trials or trials considered to be high evidence trials (Phang *et al.* 2018).
- A large variance in the outcome measures utilised to chart clinical progress in those studies that have been completed (Dhondt 1999; de Jong *et al.* 2004; Häkkinen 2004; Eversden *et al.* 2007; Durcan, Wilson and Cunnane 2014; Chung and Mior 2015), which results from the unclear RA pathogenesis (McInnes and Schett 2011; Yang *et al.* 2017).

This leaves the public, the physician and multidisciplinary team and their patients faced with a conundrum (Chandrashekara 2011) arising out of the following factors:

- the lack of capacity firstly to provide adequate information to patients and, secondly, to gain full informed consent from them (Allied Health Professions of South Africa 2015; Boucher, Brousseau and Chahine 2016),
- the inability to produce practice guidelines and follow them (Gambrill 2003),
- the reduction of support from the clinical setting, and
- the erosion of inter-disciplinary confidence (Makaram 1995).

## 1.2 Motivation for this research project

Chandrashekara (2011) sums up the current situation in his statement that health policy-makers and the medical community can only advocate correctly for and against CAM modalities and enable optimal co-management subject to the research quantifying safety profiles and efficacy for these modalities. Taken in conjunction with the growing number of RA patients seeking out interventions for themselves (Phang *et al.* 2018; Schlaeger *et al.* 2018), the issues mentioned above highlight the need for continued summation and review of clinical research evaluating CAM interventions for the RA patient group.

This research project therefore proposes to address the above concerns by targeting the existing lack of correlated research findings on the efficacy or otherwise of the manual therapy aspect of CAM in relation to the various RA treatment models identified in the preceding pages. It aims at a critical evaluation of the current literature on non-invasive manual therapies (mobilisation, manipulation, massage and exercise) in the treatment of RA patients in order to provide an evidence base for both allopathic medical and CAM practitioners (e.g., chiropractors) as well as for patients and future research in this field (Higgins, Green and Scholten 2008; Dagenais and

Haldeman 2011). A mixed method systematic review (Harden *et al.* 2018) was decided on as the research model most suitable for the purposes and scope of the study.

### 1.2.1 The research problem:

Does evidence exist for or against the provision of non-invasive manual therapies (massage, manual therapies, and exercise prescription) for the treatment of patients suffering with RA?

### 1.2.2 **Aim and methodology**:

As indicated in 1.2 above, the aim of this systematic review of the efficacy of non-invasive manual therapy in the management of RA is to determine the level of evidence available for the results of the treatment of rheumatoid arthritis by means of non-invasive manual therapy. This evaluation will be achieved by employing an evidence-based, non-biased critical analysis of studies including randomised control trials (RCTs), non-randomised clinical trials (nRCTs), and case reports / studies / series (CS/CR) (Roopchand 2014).

### 1.2.3 Objectives:

1. To determine the level of methodological rigour of studies involving non-invasive manual therapies for patients suffering with RA.
2. To determine the level of evidence in support of non-invasive manual therapies (manipulation, mobilisation, massage, and exercise) for patients suffering from RA.
3. To make recommendations for further investigations regarding non-invasive manual therapies for patients suffering with RA.

### 1.2.4 Rationale and benefits:

Despite the accumulation of research (Michalsen 2013), collectively the various fields of CAM still require evidentiary validation such as provided by RCTs for specific musculoskeletal conditions (Macfarlane *et al.* 2012). This is essential in view of how many fields of CAM there are and how many of these themselves offer several possible modes of therapy. Most of these modes of therapy have not yet undergone any trials in order to be evaluated (Macfarlane *et al.* 2012).

Thus, owing to the small body of existing research together with the ongoing efforts to grow the evidence body (Michalsen 2013; Phang *et al.* 2018), uncertainty is created by contradictory scientific data (Furlan *et al.* 2012). This therefore highlights both the shortfall in knowledge in this field and the need for further research because, as mentioned above, the status quo leaves the public, the physician, and patients without reliable answers (Chandrashekara 2011) for the following reasons, among others:

- If therapies are untested there is no information available to enable patients to give informed consent to such manual therapies as may be utilised to assist them (Allied Health Professions of South Africa 2015; Boucher, Brousseau and Chahine 2016)
- The lack of evidence-informed practice (Joseph 2004) decreases the cultural authority and contributes to a perception that medical professionals may be employing practices the merits of which are unsubstantiated by research that supports their use either in general or for a particular condition (Gambrill 2003; Chandrashekara 2011);
- Erosion of inter-professional practice interactions, with a consequent negative impact on patient care (Makaram 1995; McQuoid-Mason and Dada 2011);
- The lack of positive contribution to evidence-informed practice guidelines (Tranfield, Denyer and Smart 2003; McQuoid-Mason and Dada 2011).
- If the concerns conventional medical science expresses with regard to the growing patient seeking behaviour (Breedveld and Kalden 2004; Bhalerao *et al.* 2013; Phang *et al.* 2018) are added to problems created by the existing lack of accredited research and substantive reviews, this further advocates the need not only for continued extensive clinical research evaluating CAM as a collective but also for the scientific community to prioritize research topics in this area in order to provide profiles of efficacy and safety (Michalsen 2013).

To sum up the motivations for this project, it is hoped that by addressing the current gap in relevant knowledge at the level of a systematic review, the research undertaken will contribute to service the concerns raised in the study by Chandrashekara (2011), which argues that health policy makers and the medical scientific community can only advocate correctly for and against modalities according to the profile of safety and efficacy as revealed by research and review. This would then provide practitioners with a basis from which to present the most appropriate and reliable evidence to patients. Additionally, this project can highlight the areas where high quality research efforts are needed to support the current evidence or to establish evidence where none currently exists (Dagenais and Haldeman 2011).

### 1.2.5 Limitations:

#### 1.2.5.1 Choice of modalities

Given that CAM is an area of complementary medical interventions that includes multiple different modalities that fall under numerous professional designations, this study is limited to the use of massage, manual therapies (viz., manipulation and mobilisation) and exercise prescription in the treatment of RA. Among the reasons for the inclusion of these modalities is their use by a range of practitioners (Carlesso *et al.* 2014; Grant, Steffen and Palmer 2020). For example, chiropractic ranks as the third most used health service in the United States after medicine and dentistry (Davis *et al.* 2012; Alcantara and Leach 2015; Grand-View-Research; 2018). In addition, massage and exercise can be implemented in the practitioner-based environments and, following

patient education, also be extended into the home environment as self-directed care (Gillman *et al.* 2018). These modalities are easily accessible to the patient, besides which they are considered relatively more affordable by comparison with other interventions and potentially have fewer side effects (Tick *et al.* 2018).

### 1.2.5.2 Publication bias

From a methodological vantage point, a potential for publication bias in a systematic review is always a threat (Knobloch, Yoon and Vogt 2011). This limitation implies that some articles (and thus evidence) might not have been published and therefore be excluded from this systematic review. This may affect a systematic review in that those articles not published tend to be those with a negative outcome for the therapy under investigation (Knobloch, Yoon and Vogt 2011), thus requiring the systematic review of the published literature to be tempered by the acknowledgement that positive outcomes may need to be accepted with caution. This limitation has restricted ability to be controlled as it is publication / author based (Knobloch, Yoon and Vogt 2011). In the present study, however, all articles found in the systematic search, regardless of study design and outcome, have been included provided that the publication met the inclusion criteria and did not fall outside the scope of the study by reason of the exclusion criteria.

### 1.2.5.3 Language bias

Language bias is a considered limitation of this study, given that the language of communication between the reviewers here and the principal researcher is English. This therefore does not allow for the inclusion in this systematic review of articles and studies in languages other than English. Consequently, an unknown number of non-English articles suited to the scope of this study (Higgins, Green and Scholten 2008) have been lost to it, which introduces an inherent bias into the study. In order to reduce this bias in future systematic reviews, it is suggested that non-English articles be included either by means of translation or a multi-linguistic study in order to ensure a greater potential for inclusion for all articles within the scope of the study (Liberati *et al.* 2009; Moher *et al.* 2009a). This suggestion, though, is cautioned when due consideration of translation bias is not included into the methodological rigour of the future studies.

### 1.2.5.4 Absence of meta-analysis

A final limitation within this study to be noted is that it is not intended to be a meta-analysis (Dagenais and Haldeman 2011). In terms of purpose and structure this study was designed to collect, critically review, and present the available evidence specific to its scope, namely, non-invasive manual therapies, in particular the use of manipulation, mobilisation, massage, and exercise in the treatment of RA patients. The study therefore focuses on the methodological rigour of the included articles. This decision was based on the variance of study designs in the published

articles as well as the range of outcomes measures utilised (Dhondt 1999; de Jong *et al.* 2004; Häkkinen 2004; Eversden *et al.* 2007; Durcan, Wilson and Cunnane 2014; Chung and Mior 2015). In this respect it differs from meta-analysis designs which utilise statistical analysis to extract and collaborate data from uniform outcomes across similar studies' designs (usually RCTs), from which conclusions are then drawn.

### 1.2.6 Conclusion

The above chapter provides the rationale behind the need for the investigation of non-medicinal interventions for the treatment of RA, which in this study, as previously stated, is limited to massage, manual therapies and exercise prescription. The following chapters will lay out the study and present the results and evaluation thereof as follows:

- **Chapter 2:** Literature Review – A review of publications outlining the condition of RA, the primary RA treatment modalities and the mechanisms of action/application in RA patients, with the aim of providing further insights into the need for this systematic review as well as a basis for the decisions articulated in Chapter 3.
- **Chapter 3:** Research Methodology – An exposition of the research design, the procedure, the research sample, the involved reviewers, the measurement tools and a description of the analysis.
- **Chapter 4:** Results – A compilation of the results obtained after data collection, in which conclusions are drawn on the evidence provided by each of the reviewed articles.
- **Chapter 5:** Discussion of Results – A review of the interventions (massage, manual therapies and exercise prescription) for RA that includes discussion of the criteria used for ranking evidence in each domain and the outcomes of the ranking of evidence for each modality in the context of RA.
- **Chapter 6:** Conclusion and Recommendations – A summation of the conclusions arising from this study together with pertinent recommendations relating both to the treatment of RA in respect of the modalities herein evaluated and to further research in the field.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Introduction:

This chapter is an overview of RA. Beginning with a definition of the disease, it looks at the impact of the disease on patients and society, discusses the risk and protection factors associated with RA development, reviews the limited understanding of the pathophysiology of the disease, and then covers methods of diagnosis and treatment. In particular, the chapter addresses the use of manual therapy in the treatment of RA and the surrounding controversies, following this with a description of a systematic review and its importance with regard to outlining the evidence available both in favour of and against the use of manipulation, mobilisation, exercise and massage in the treatment of RA.

## 2.2 Definition of Rheumatoid arthritis:

As mentioned in the introduction to Chapter 1, RA is a common, chronic and progressive autoimmune condition confirmed by the presence of RF and Auto-Citrullinated protein antibodies (ACPA) in 70 to 80% of all RA patients (Sharma *et al.* 2018; Okada *et al.* 2019) and evolves from immune dysregulation (Ridgley, Anderson and Pratt 2018) that manifests primarily in synovial joints in a polyarticular pattern (effecting the tendons, synovial tissue, bursae, and the articular bone and cartilage). The development of synovial cell hypertrophy, inflammatory infiltration, and vascular changes results in the formation of the characteristic pannus, the ensuing destruction of the affected bone and cartilage and the consequent periarticular osteopenia (Lewis and Battaglia 2019). However, one could consider RA as a syndrome, owing to the extra-articular manifestations (Smolen, Aletaha and McInnes 2016), which include rheumatoid nodules, pulmonary involvement and / or vasculitis among other systemic comorbidities such as Sjögren's syndrome, pericardial effusion, fibrosing alveolitis, Raynaud’s phenomenon, peripheral neuropathy, splenomegaly (Longmore *et al.* 2014). RA represents a major challenge to global health and unfortunately remains agnogenic (Yang *et al.* 2017; Safiri *et al.* 2019). It is particularly complicated to treat and leads to a guarded prognosis by medical practitioners (McInnes and Schett 2011). Despite ongoing effort, there is no pathognomonic test for the diagnosis (Sharma *et al.* 2018). This complex problem has an impact on patients, their social dynamics and work circumstances as well as on society in general (Phang *et al.* 2018; Myasoedova *et al.* 2020).

## 2.3 The impact of RA on patients, society, and health care:

### 2.3.1 Quality of life:

As indicated in Chapter 1, RA severely impacts the QoL of patients (Intriago *et al.* 2019). The negative impact of RA on QoL presents a major public health concern (Ethgen *et al.* 2002; Myasoedova *et al.* 2020). This is illustrated by the study of Intriago *et al.* (2019) which points out that:

- female RA patients tend to experience higher levels of pain, disease activity and disability rates.
- patients with a higher level of education have better QoL, thought to be owing to better coping mechanisms.
- the support structure of marriage has exhibited a positive influence over QoL.
- the financial and social support of an occupation enhances patients' QoL.
- decreasing physical activity with advancing age has a negative impact on patients' QoL.
- the duration of the disease has a bidirectional effect on patients' QoL. Some studies have reported an association between longer duration of RA and decreased QoL, while other studies have found a reverse correlation, stating that with time the patients develop a higher degree of acceptance for the imparted effects of RA, thus reducing anxiety, depression, stress, and pain and resulting in improved sense of well-being.

In addition, co-morbidities, of which dyslipidaemia, hypertension, diabetes, depression, and metabolic disorders are among the most common, are known to exacerbate the RA disability and place limitations on patients' independence, for example, their work capacity and the activities of daily living (ADL), while RA also increases the burden on the patient's support system (Naqvi, Hassali and Aftab 2019). According to Walker *et al.* (2013), 40% of all RA patients will experience a disease progression that disables them to some extent within three years from onset. Within the first 20 years of disease, 80% are designated as at least moderately disabled. RA also negatively affects patients' activities, among them work, housework, service activities, running errands and climbing stairs. If patients do not receive treatment within two to three years of diagnosis, 20% to 30% of them become disabled to the point of being unable to work and then cannot work (Tokem *et al.* 2014). Thus, a comprehensive understanding of the systemic effect of RA on QoL looks for and requires innovations in interventions that focus on improving the patient's QoL (Moquin *et al.* 2009). According to the World Health Organisation (WHO), the complexity of QoL is reflected in the impact on the physical, psychological, and social domains that create the life experienced by the individual (Malm *et al.* 2017). The chronically afflicted RA patient therefore benefits from a multidisciplinary team that focuses on a holistic approach to the treatment of RA as well as providing support and improving the QoL (Ethgen *et al.* 2002; Moquin *et al.* 2009; Naqvi, Hassali and Aftab 2019).

### 2.3.2 Epidemiology of RA

Safiri *et al.* (2019) report the lack of comprehensive global epidemiological studies, highlighting the reliance on modelling algorithms that estimate that RA has a prevalence of 20 million, a yearly incidence rate of 1.2 million (or 0.5 to 1% as defined by Smolen, Aletaha and McInnes (2016)), and a disability-adjusted life year (DALY – the measure of burden a disease imposes, expressed in number of years lost due to the condition) of 3.4 million. In addition, the overall global age-standardised prevalence and incidence rates of RA have been showing a trend of increase since 1990 (Myasoedova *et al.* 2020). This age-standardised prevalence and the incidence rates show a high female preponderance as well as an increase with age up to the age of 70, after which there is no statistical difference between male and female cases (Safiri *et al.* 2019).

Table 2.1: Epidemiological statistics

| **Country** | **Available data on trends, incidence, and prevalence:** | |
|---|---|---|
| Algeria, | Prevalence 0.13% | (Naqvi, Hassali and Aftab 2019) |
| Canada | Prevalence of 0.9%<br>Period 2001 to 2014 showed no change in RA incidence. | (Naqvi, Hassali and Aftab 2019)<br>(Myasoedova *et al.* 2020) |
| Congo | Prevalence 0.9% | (Naqvi, Hassali and Aftab 2019) |
| Egypt | Prevalence 0.2% | (Naqvi, Hassali and Aftab 2019) |
| Finland | Period 2000 to 2014 showed a decrease in seronegative RA incidence and a stable seropositive RA incidence<br>Total RA prevalence 1.90% | (Myasoedova *et al.* 2020)<br><br>(Almutairi *et al.* 2021) |
| France | Prevalence 0.31% | (Naqvi, Hassali and Aftab 2019) |
| India | Prevalence 0.75% | (Naqvi, Hassali and Aftab 2019) |
| Italy | Prevalence 0.41% | (Naqvi, Hassali and Aftab 2019) |
| Japan | Prevalence 0.6% - 1% | (Naqvi, Hassali and Aftab 2019) |
| Lithuania | Prevalence 0.55% | (Naqvi, Hassali and Aftab 2019) |
| Nigeria | Prevalence less than 0.5% | (Naqvi, Hassali and Aftab 2019) |
| Serbia | Prevalence 0.35% | (Naqvi, Hassali and Aftab 2019) |
| South Africa | Prevalence 0.9% | (Naqvi, Hassali and Aftab 2019) |
| South Korea | Prevalence 0.26% | (Naqvi, Hassali and Aftab 2019) |
| UK | Period 1990 to 2015 showed a decline in annual incidence of 1.6% with the estimates of 38.1 cases per 100 000<br>Prevalence 0.81% | (Myasoedova *et al.* 2020)<br><br>(Symmons *et al.* 2002) |
| USA | Prevalence 0.6 – 1% (over 1.3 million adults) | (Myasoedova *et al.* 2020) |

The data presented in Table 2.1, compiled from Symmons *et al.* (2002), Naqvi, Hassali and Aftab (2019), Myasoedova *et al.* (2020) and Almutairi *et al.* (2021), shows temporal and geographic variability (such as the apparent reduction noted from north to south within the northern hemisphere (as seen in the North American and Eastern Mediterranean regions)). Additionally, authors have provided insight into the variability in prevalence between urban and rural areas. These along with the previously stated temporal and geographic variability may possibly be influenced by genetics, ethnic and environmental factors (Smolen, Aletaha and McInnes 2016; Myasoedova *et al.* 2020).

When the data is stratified to incorporate serological phenotype, the incidence and prevalence of RA collectively remains unchanged for the data derived in the USA. However, the trend reported in the RA incidence and prevalence data for Olmsted County in Minnesota between 2005 and 2014 compared with previous decades shows that the incidence of seropositive RA had

decreased while seronegative RA had increased (Myasoedova *et al.* 2020). Some potential factors influencing these observed incidence trends include:

- the criteria used for diagnosis (the 1987 American College of Rheumatology (ACR) criteria verses the 2010 ARC criteria),
- the inherent low sensitivity for seronegative RA by the 2010 ARC criteria,
- changes in exposure to environmental factors (smoking and obesity are two factors noted to have change in exposure over the last few decades (Hill 2018; Organization 2019)) that influence the risk of RA development, some of which are serological phenotype specific (Myasoedova *et al.* 2020).

Although these stratified data can provide insight into the pathophysiology, the course of the disease, the treatment of choice, the healthcare use, and the policy planning for the serological phenotypes (Myasoedova *et al.* 2020), limited information is available.

### 2.3.3 Impact of RA on healthcare and society

RA places a substantial burden on both the patient and society (Ernst and Posadzki 2011; Smolen, Aletaha and McInnes 2016). For the patient, this involves conditions of malignant pain, progressive debilitation and disablement, systemic complications, and the likelihood of premature death. In addition, there are the costs, both direct and indirect, to the patient and society (Ernst and Posadzki 2011; McInnes and Schett 2011). Fiscally, RA accounts for up to 1% - 2.5% of the gross national product of Western nations as part of the health care budget (Ernst and Posadzki 2011). These costs are synonymous with but not limited to those identified in the literature below:

- the consumption of significant amounts of medication to control or limit the progression of the condition (Breedveld and Kalden 2004),
- medication-induced side effects such as NSAID-induced adverse events, changes in systolic blood pressure, increased cardiovascular risk and morbidity (Ernst and Posadzki 2011),
- the adaptation of the patient's daily activities (Christie *et al.* 2007),
- a reduction in the patient's QoL (Christie *et al.* 2007; Moquin *et al.* 2009),
- the psychological impact of the RA which may include depression, anxiety, and stress (Moquin *et al.* 2009; Naqvi, Hassali and Aftab 2019),
- absenteeism from work (Naqvi, Hassali and Aftab 2019),
- presenteeism (Naqvi, Hassali and Aftab 2019) and
- medical boarding with increased reliance on state support for medical and subsistence costs.

Given the impact RA has on the patients, their social framework, and health sector, understanding the risk factor that predisposes the patients to developing or protecting against the development of RA can offer areas of focus for preventative medical care. For the is reason the next section of this chapter deals with these risk and protective factors associated with RA.

## 2.4 Risk and protective factors for developing RA

RA, a chronic inflammatory condition, is a complex interaction of multiple genetic traits and environmental risk factors (McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019).

### 2.4.1 Genetic factors

Research into the genetic link to developing RA indicates that over 100 genetic risk loci have been identified, but evaluation of the causal link to the pathophysiology of RA is still required (McInnes and Schett 2011; Smolen, Aletaha and McInnes 2016). Missense mutations cause a change in the DNA sequence, which results in a functional change in the amino acid and its protein structure sequence. This is thought to be pathogenic, based on the fact that genes have been identified as important in the role of T-cell immunity. This process carries significance in the pathogenesis of RA, including the HLA, PTPN22, STAT4, TRAF1 and IL2RA loci (McInnes and Schett 2011). Of these, the most acknowledged genetic link to autoimmune susceptibility is the class 2 classic human leucocyte antigen gene allele (HLA-DRB1, especially the QKRAA amino acid motif) within the major histocompatibility locus (MHC) of chromosome 6, which is linked to the positive RF and ACPA (McInnes and Schett 2011; Smolen, Aletaha and McInnes 2016). This change in the locus confers a 50% genetic risk (McInnes and Schett 2011). Additionally, a non-MHC locus (PADI4) has been identified to interact with environmental factors, conferring significant risk for the likelihood of developing RA across several generations.

This risk has also been linked with the non-MHC loci associated with the autoimmune susceptibility (PTPN22 gene encoding for protein tyrosine phosphatase (PTP)) expressed in haematopoietic cells. PTP regulates immunoreceptor signalling cascades of T-cell activation and development of the T-cell in the thymus (Smolen, Aletaha and McInnes 2016). This negatively regulates antigen receptor signalling in both T-cells and B-cells, as a result of a missense single nucleotide polymorphism's (SNP) mutation in this gene (the PTPN22 R620W + 1858C/T). In this context, the R620W risk allele mutation is a gain-of-function mutation, as a result of which T-cell receptors and B-cell receptors present with a reduced signalling that affects the clonal selection of the receptor/s of the lymphocytes. The reduction in regulatory function of the T-cell regulators has been linked to toll-like receptor (TLR) 7 signalling of the plasmacytoid dendritic cells (pDC) and hypercitrullination in blood mononuclear lymphocytes (Smolen, Aletaha and McInnes 2016). Thus, the PTPN22 R620W + 1858 C/T polymorphism plays a role in the autoantibody component of the autoimmune disease and makes the patient susceptible to RA (McInnes and Schett 2011).

In the context of the autoimmune disease, the Fc receptors of the autoantibodies bind to the autoantigens, forming immune complexes and initiating the immunoreceptor signalling cascades. The antigen uptake, presentation and cytokine production are of particular importance and contribute to RA pathophysiology (Derksen, Huizinga and van der Woude 2017).

Risk variants associated with upregulation of cytokine production, such as seen in the STAT4 and CCR6 genes, impart their effect by the cytokine pathways within lymphocytes. An example is the activation of the tumour necrosis factor – alpha (TNF-α) pathway in CD4 T-cells (McInnes and Schett 2011). Collectively these molecular and cellular processes are responsible for the chronic inflammation, bone erosion (peri-articular bone osteopenia) prevalent in RA (Hosseini *et al.* 2015; Wei *et al.* 2015), in addition to stimulating the autoimmune disease setting of RA (Jegalian, Facchetti and Jaffe 2009).

### 2.4.2 Hormonal factors

Hormonal homeostasis, albeit poorly understood, has an influential role in maintaining a healthy and competent immune system. Furthermore, an intimate, dynamic, and bidirectional connection exists between the immune system and neuro-endocrine system (Borba, Zandman-Goddard and Shoenfeld 2018). This is evidenced by the sex bias in RA prevalence towards female (with a female to male ratio of 1 to 7:1) (Tan, Peeva and Zandman-Goddard 2015). In support of this, females have an enhanced immunological reactivity through the enhanced capacity of antigen presentation (compared to males), an increased rate of synthesis of antibodies, elevated Ig levels, and quicker allograft rejection than males (Borba, Zandman-Goddard and Shoenfeld 2018). As a result, there is a possible role played by hormonal and reproductive factors in the development of RA (McInnes and Schett 2011; Alpízar-Rodríguez *et al.* 2017). This is reflected in the fluctuation of disease severity in female patients with RA with regard to the following factors (Hughes 2012; Tan, Peeva and Zandman-Goddard 2015; Alpízar-Rodríguez *et al.* 2017):

- menstrual cycle phases,
- contraceptive use,
- effect of pregnancy,
- post-partum period hormonal changes,
- menopausal status and
- hormonal replacement therapy.

However, the role the endocrine system plays in RA development remains unclear (Pedersen *et al.* 2006b; Alpízar-Rodríguez *et al.* 2017). The literature on the subject has proposed that oestrogen and prolactin (Pr) exert an immune stimulatory effect acting in a pro-inflammatory manner, while the progesterone (Prg) and androgens exert an immune suppressive effect acting in an anti-inflammatory manner (Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018) thereby affecting the expression of the RA.

### 2.4.2.1 Oestrogens:

As one of the factors considered to influence RA, oestrogens may exert an influence on different immune cells (Alpízar-Rodríguez *et al.* 2017). Oestrogens have different effects on different immune cells. This is related to the serum concentration, reproductive phase/menopausal status, medication supplementation (i.e., oral contraceptives and hormone replacement therapy), oestrogen receptor expression and intracellular metabolism (Alpízar-Rodríguez *et al.* 2017). For example, decreased oestrogen serum levels (between peri-ovulation and pregnancy) are able to stimulate both B-cell and T helper (Th) 2 cell response, support the survival of the auto-reactive clones (B and T cell lineages) and inhibit the differentiation of naïve T-cells into Th 17 cells (Noack and Miossec 2014; Alpízar-Rodríguez *et al.* 2017). These cells are thought to be responsible for the chronic inflammatory tissue destruction occurring in RA (Straub 2007).

In the synovial tissue of RA patients, oestrogen binds to oestrogen receptors (beta more so than alpha), which results in local hypoxia and inflammation stimulated by the expression of oestrogen receptor-β as well as an increase the production of Interleukin (IL)-6 and TNF-α by the local mononuclear cells (Alpízar-Rodríguez *et al.* 2017). Additionally, the TNF-α binds to specific receptors (e.g., p55) and drives the production of other cytokines, thus playing a role in the support of chronic inflammation. Additionally, IL-6 activates leukocytes within the synovium and enhances the production of antibodies. Collectively, the leukocytes and pro-inflammatory cytokines within the RA joint cause bone erosion and stimulate the osteoclastogenesis that results in the bone resorption (peri-articular bone osteopenia) prevalent in RA (Hosseini *et al.* 2015; Wei *et al.* 2015).

The above is also evidenced by the rapid menopausal drop in oestrogen serum levels, which sees an associated spontaneous rise in circulating pro-inflammatory cytokines (i.e., IL-1 and 6, and TNF- α). Introducing oestrogen therapeutically to peri-ovulatory levels in a post-menopausal woman has shown inhibitory effects on this cytokine release in a dose-dependent manner that is not noted in men and pre-menopausal women (Alpízar-Rodríguez *et al.* 2017). Therefore, the use of E2 oestrogen in the post-menopausal RA patient may contribute to the management of pro-inflammatory cytokines and prevent or lessen the resultant bone erosion and peri-articular osteoporosis in specific RA patient subgroups (Hosseini *et al.* 2015; Wei *et al.* 2015).

### 2.4.2.2 Androgens:

Present in both men and women, androgens are able to inhibit mononuclear cells, in particular the differentiation of the naïve T-cell into Th 1 and 17. By comparison with healthy, RA-free controls, both men and women with RA have been reported to have low concentrations of gonadal and adrenal androgens. In men with RA, not only are adrenal androgen levels low, but oestrogen levels are high (Alpízar-Rodríguez *et al.* 2017).

The protective effect of androgens due to the low serum concentration is negated in RA patients. This results from the inflammatory mediators stimulating an increased activity of the enzyme aromatase in peripheral tissue, reducing protective androgen levels. With low androgen concentrations there is no inhibition of the local mononuclear cells producing inflammatory cytokines (IL-6 and TNF- α) (Hosseini *et al.* 2015; Wei *et al.* 2015). Within the synovial fluid the ratio of oestrogens to androgens in both male and female RA patients is significantly altered, with oestrogen concentrations remaining significantly higher than that of androgen concentrations (Alpízar-Rodríguez *et al.* 2017). This predisposes the RA patient to an enhanced inflammatory component of RA (Section 2.4.2.1).

### 2.4.2.3 Progesterone (Prg)

Prg is a suppressive immunoregulatory hormone that exerts an unquantified influence over the immune pathways involved in RA (Hughes 2012; Borba, Zandman-Goddard and Shoenfeld 2018). Yet RA often falls into remission during pregnancy (a known state of immune-modulation), a time where oestrogen and Prg are physiologically extremely high systemically (Hughes 2012; Tan, Peeva and Zandman-Goddard 2015).

Prg mediates its effect on cellular function directly by means of a high specificity for membrane-bound and intracellular (cytosol and nuclear) receptors, inducing signalling cascades (Hughes 2012; Tan, Peeva and Zandman-Goddard 2015). Membrane-associated receptors include mPR-α / β or γ or the Prg receptor membrane component 1 (PGRNC1), which explain the rapid non-transcriptional action of Prg. Intracellular Prg receptors (iPRs) are characterised by a nuclear receptor related to a glucocorticoid receptor (GR), a highly expressed receptor on the membrane of most immune cell types, as well as a mineralocorticoid receptor. Prg has also shown interaction with ion channels in the plasma membrane, altering the cellular function.

Given the above receptors, the effect of Prg on specific immune cells is noted as follows (Hughes 2012; Tan, Peeva and Zandman-Goddard 2015):

- **Neutrophils**: (as will be explained further in the following section) have the capacity to inflict direct damage to the joint complex by the synthesis and release of proteases and reactive oxygen species (ROS), while participating in the induction and maintenance of inflammation via cytokine, prostaglandin, chemokine, and leukotriene synthesis, as well as the antigen presentation required to activate T-cell immune response. Unlike the peripheral blood neutrophil tissue, resident neutrophils in RA inflammatory sites undergo molecular changes that have a suppressed apoptosis extending their life and via this the action they impact on the joint (Wright *et al.* 2010). Prg suppresses ROS synthesis and apoptosis and hinders chemotaxis.
- **Natural killer cells (NK)**: Prg suppresses the production of interferon (INF)- γ (which “plays an important role in the innate and adaptive immune response, also playing a role in ROS

and cytokine production, antigen presentation, metabolic pathways, cellular differentiation, growth and survival of immune cells" (Kato 2020)) and induces apoptosis, in this way reducing the number of NK cells present and thus reducing the impact they have as a collective on the disease activity, which results in a reduction of RA related inflammation and damage.

- **Macrophages**: Prg. inhibits nitric-oxide and TNF- α synthesis. In the RA synovial joint Prg. directly suppresses the release of TNF- α, a factor potentially related to RA remission in pregnancy. Additionally, dysfunction related to cellular debris clearance by macrophages potentially results in a loss of immune-tolerance for nuclear-antigens.
- **Dendritic** (**myeloid** and **plasmacytoid**) **cells (DC)**: In the myeloid dendritic cells (mDC), Prg enhances the expression of IL-10 (an anti-inflammatory cytokine), thus inhibiting the expression of IL-1, 6 and 12 and TNF- α (all pro-inflammatory cytokines). It also inhibits the activation of mDC via TLRs (see Section 2.4.1: Genetic factors). This is further supported by the high Prg levels seen in pregnancy, resulting in mDC antigen presentation to CD4+ cells, in particular the Treg and Th2 cells, while Th1 cell function appears to be hindered. This results in RA suppression (see Section 2.5.2).
- **Plasmacytoid dendritic cells (pDC):** Here Prg inhibits TLR-9-signalled IFN-α production, which results in a decrease in chronic inflammation.
- **CD4+ T-cells**: Prg influences a Th 2 cell response by stimulating these cells, while it suppresses differentiation of naïve Th cells into Th 1 and 17 and enhances differentiation in Th 2 and Treg cells. The high concentrations of Prg noted in pregnancy enhance the IL-4 synthesis (a potent immune regulator) by CD4+ clones. These developments occur in response to a rapid loss of Prg which can stimulate expression of IL-1(β) and TNF- α, resulting in an RA flare-up (Alpízar-Rodríguez *et al.* 2017) (see Section 2.5.2).
- **CD8+ T-cells**: In autoimmune diseases CD8+ cells show irregular activity. During the menstrual cycle (in particular the luteal phase), Prg through CD8+ cells influence immune regulation via Treg cell stimulation and the suppression of NK cells. Prg also appears to suppress IFN-γ synthesis, resulting in a decrease in chronic inflammation.
- **B-cells**: Prg potentially suppresses B-cell class switch recombination and somatic hypermutation, both important mechanisms in B-cell differentiation. Furthermore, Prg stimulates antibodies' post-translation modification (by the mechanism of citrullination), influencing function to induce a remission of RA induced by pregnancy.

**With regard to the hormonal dynamics of pregnancy,** Prg specifically appears to facilitate remission of RA, provided the RA is active prior to the pregnancy. Often there is a flare-up post-partum (Hughes 2012). It is suggested that the Prg-induced changes in systemic immunity are involved in the RA remission (with increased IL-1RA induction of Tregs, suppressed Th1 or Th17 responses, increased Th2-related responses and increased levels of terminal galactose residues

on circulating Ig) (Hughes 2012). Additionally, factors released at the maternal-fetal interface may be active in pregnancy-induced remission of RA (Hughes 2012).

#### 2.4.2.4 Prolactin (Pr)

Pr, a hormone secreted primarily by the pituitary gland, is under the inhibitory control of dopamine via the hypothalamic-pituitary-adrenal axis. It may be produced in secondary sites (e.g., mammary glands, prostate, and immune cells). These sites do, however, produce Pr that has different molecular weights and differing bioactivity. Pr has an immuno-stimulatory effect which influences the regulation of immune cells and the communication between them (Borba, Zandman-Goddard and Shoenfeld 2018). The reported effect of Pr on specific immune cell lines includes the enhanced synthesis and release of pro-inflammatory cytokines by DC, T-cells, NK cells, monocytes and macrophages. DC cells are stimulated by Pr to upregulate the MHC and CD80 surface molecules. In granulocytes, Pr activates the mitogen-activated protein kinase (MAPK) and signal transducer and activator of transcription 1 (STAT1), imparting a stimulus on the transcription of inducible nitric oxide synthase and IFN regulatory factor 1 (Borba, Zandman-Goddard and Shoenfeld 2018). Pr has a dose-dependent influence on differentiating the NK cell into progesterone-activated killer cells (PAK) altering their cytotoxic and proliferator activity, which results in an increase in the chronic inflammation seen with RA.

In line with these findings, pregnant women with RA experience a period of remission or significant reduction in RA severity. This period overlaps with the period of hypercortisolism as cortisol inhibits the secretion of Pr by the endocrine system. However, the post-partum patient in due course commonly experiences flare-ups correlating with the Pr levels increasing again. Breast-feeding further stimulates the secretion of Pr and exacerbates the severity of the RA. Most female patients will relapse by the third month post-partum with nearly all experiencing flare-ups by the ninth month post-partum. By comparison with controls, the RA patient shows higher Pr levels both in serum and synovium. These findings are in line with the increased Pr synthesis and release both systemically by the endocrine system and locally (within the synovium) by the lymphocyte. These Pr concentrations relate to the disease activity. A genetic link has been postulated between Pr and RA. The Pr gene is located in close proximity to the HLA-DRB1 alleles in the sixth chromosome (Borba, Zandman-Goddard and Shoenfeld 2018). Furthermore, severity of the condition has been associated with the length and number of children breastfed. For this reason, RA patients should avoid breastfeeding (Borba, Zandman-Goddard and Shoenfeld 2018).

### 2.4.2.5 Female specific hormonal risk factors for RA

Given the dynamic changes of several hormones in the life of a female, when we consider the fluctuations over time of hormones like Prg, Pr, oestrogen and androgens and their effect on the immune system (as noted above), the collective effect is that the female patient carries an increased risk for the development of RA. The effects and the timing, both pre- and post-natal and also pre- and post-menopausal, and the resultant exposure to changes in hormones appear to play a role in the onset of RA. An example of this is the acute drop in ovarian function and / or bioavailability of oestrogen as seen in female patients who experience early onset of menopause, who are going through the post-partum period, who are using anti-oestrogen agents, or are post-menopausal, all of whom have an increased risk of RA onset. However, hormonal therapy (contraceptives and hormonal replacement therapy) is not yet clearly associated with the onset of RA (Alpízar-Rodríguez *et al.* 2017).

#### 2.4.2.5.1 Age of menarche

The age of menarche holds a significant risk for development of RA with regard to both subtypes, ACPA-positive and negative. When menarche occurs at the age of 15 or older, the risk of developing RA is doubled by comparison with women who experience menarche at 12 years old or younger (Alpízar-Rodríguez *et al.* 2017).

### 2.4.3 Infection factors

A long-standing link between RA and infective agents and the products of their infective processes has been noted. For example, the development of RA has been associated with periodontal diseases. One postulation is that *Porphyromonas gingivalis* expresses Protein-arginine deiminase (PADI) type-4 (PADI4, which holds a function in citrullination), and induces aberrant citrullination. In doing so, this provokes the loss of local tolerance for PADI4 (McInnes and Schett 2011; Derksen, Huizinga and van der Woude 2017) (see Section 2.4.1), which in turn results in increased inflammation and increases the risk of RA development. This links with immune complexes produced through the infection inducing RF action (through molecular mimicry), resulting in a high affinity binding with the Fc portion of immunoglobulin (McInnes and Schett 2011). However, research efforts have shown that *Parvovirus B19* and the Epstein-Barr virus both demonstrated a non-significant association with RA. In contrast, allergy-induced asthma diagnosed before the age of 45 (ACCP negative RA) and previous urinary tract infections with unspecified serology appear to have a protective effect, thus significantly reducing the risk of developing RA (Pedersen *et al.* 2006b). Therefore, the mechanism has been generally postulated but has not yet been substantiated (McInnes and Schett 2011).

### 2.4.4 Life events

Adverse life events have been associated with the onset of RA. A link between the hypothalamic-pituitary-adrenal axis with the production of cytokines has begun to provide potential insight into the underlying mechanisms at a molecular level. Molecular explanations for the impact of such phenomena are emerging from animal models of inflammation, which show a link between the hypothalamic-pituitary-adrenal axis and cytokine production (McInnes and Schett 2011). The central nervous system is normally involved in immune regulation and homeostasis, and neuro-immunologic interactions regulate disease development in rodent models of arthritis. Such effects may operate locally (several neurotransmitters are expressed in synovitis in RA) or centrally (cytokines are rapidly up-regulated in the hypothalamus during peripheral inflammation) (McInnes and Schett 2011). Looking into the dynamics of psychology, more specifically emotional state and stress and the role they play with regard to the RA symptoms experienced, among them pain and the severity of the disease, the evidence indicates that these interplaying factors can exacerbate inflammation in RA patients (Graham-Engeland *et al.* 2019). The broad category of negative emotions such as depression, hostility and stress has been linked to an increased level of inflammation as observed in circulatory markers. These findings associate negative emotional states with circulating IL-6 in a dose-dependent nature where higher negative emotion equates to higher levels of inflammation, while positive emotional states are associated with reduced circulating IL-6 in an inverted correlation: higher positive emotion equates to lower levels of inflammation. The chronicity of a negative emotional state imparts via cytokine state the chronicity of an inflammatory state, hence conferring the risk of triggering the onset of RA (Graham-Engeland *et al.* 2019). Nevertheless, the impact of individual emotional states on RA remains largely elusive and is poorly documented.

### 2.4.5 Environmental factors

#### 2.4.5.1 Pet ownership

Children growing up with pets in the household have shown no association with the risk of developing RA. Furthermore, adults with pets in the household have been associated with a reduced risk in both ACCP positive and negative serology subtypes of RA; it is as though the presence of pets contributes some protection against developing RA. In addition, the association holds a statistically significant protective role in the ACPA-positive subtype of RA (Pedersen *et al.* 2006b). This may have a connection with the presence of positive emotional states (Section 2.4.4), which have been associated with decreased levels of inflammation, thus decreasing the likelihood of RA (Graham-Engeland *et al.* 2019).

### 2.4.5.2 Marital status

It has been noted that RA patients are more commonly unmarried (Odds Ratio (OR) of 1.71); despite the lack of statistical significance the study comments on the trend of increased risk showing a stronger association with ACPA-positive RA (Pedersen *et al.* 2006b). This may be associated with the decreased negative emotional states (Section 2.4.4) associated with marriage (Sbarra 2009) that have been associated with decreased levels of inflammation, thus decreasing the likelihood of RA (Graham-Engeland *et al.* 2019). However, Kiecolt-Glaser, Gouin and Hantsoo (2010) suggest that marital conflict induces stress, thus enhancing an inflammatory state. In other words, marital emotional support is protective against RA, while marital conflict imposes risk.

### 2.4.5.3 Socioeconomic status and education

The case study by Pedersen *et al.* (2006b), reports a risk association between low socioeconomic status and RF-positive RA. This concurs with lack of employment showing an increase in risk of RA collectively (Ilar *et al.* 2018); however, the risk trend shows a stronger association with ACPA-positive RA. In a systematic review by López-Castillo *et al.* (2014), one of the cited studies (Olsson, Skogh and Wingren 2001) reports that a higher education level plays a protective role with regard to the development of RA, but another study (Smedstad *et al.* 1996) contradicts this, stating that the education level had an inverse role, particularly in RF-positive RA. Two additional studies cited in the systematic review (Bedi *et al.* 2005; Damjanović *et al.* 2009) report that low-level education significantly increases the risk of RA occurrence. This is supported by a Danish study (Pedersen *et al.* 2006a) cited in the systematic review that reports that a longer educational history provides a protective function with regard to the development of RA, while a shorter educational history increases risk (Pedersen *et al.* 2006b; Smolen, Aletaha and McInnes 2016).

### 2.4.5.4 Physical activity (PA)

PA has been associated with a number of biological and immunological effects. These include the following:

- The effect of leisure PA on RA may be affected by the following PA biologically imparting an immunological effect by regulating immune cell function and reducing pro-inflammatory cytokine production (Di Giuseppe *et al.* 2015). This is associated with reduced RA flare-ups as well as reduced disease activity (Kasapis and Thompson 2005; Kiecolt-Glaser, Gouin and Hantsoo 2010; Cooney *et al.* 2011).
- Also, the release of dopamine with exercise (Sutoo and Akiyama 2003) acts as a Pr inhibitor, thus preventing Pr pro-inflammatory effect.
- PA decreases high Body Mass Index (BMI) which is associated with the risk of developing RA (Liu *et al.* 2019).

- There is an association between increased intensity of PA which may suggest a cardiovascular protective mechanism (Liu *et al.* 2019).
- PA is associated with the release of endorphin and enkephalin, which has been connected to an elevated positive emotional state (Graham-Engeland *et al.* 2019) (Section 2.4.4).

Therefore, while literature reports that physically demanding jobs significantly increase the risk of RA development (Ilar *et al.* 2018), recreational PA outside of the working environment has been shown to provide a 33% reduction in risk (Pedersen *et al.* 2006b; Liu *et al.* 2019). Leisure inactivity, however, appears to increase RA risk (Di Giuseppe *et al.* 2015). Literature recommends PA so as to avoid the triggering of RA. It also recommends that RA patients increase their PA in order to reduce the risk of RA progression (Liu *et al.* 2019). However, a conundrum exists in the quantification of risk association between RA and PA in that the limitations associated with RA such as joint inflammation, pain and deformation may well place limitations on a patient's PA, which is stated as a reverse causation bias (Liu *et al.* 2019).

The one possible influencer regarding the outcome of whether PA is protective of or promotes the progression of RA lies in the details of the PA exercise programme: loads, load progression, volume of exercise, repetitions, frequency, intensity, sets, rest intervals, duration of programme, specificity of exercises, individualisation of the protocol to the participant (Heine *et al.* 2012). These vary between studies and have potentially differential effects at different points in the pathogenesis of RA (Hansen *et al.* 1993; Lyngberg *et al.* 1994; de Jong *et al.* 2004; Durcan, Wilson and Cunnane 2014; Williams *et al.* 2015).

### 2.4.5.5 Obesity

The WHO reported that the number of obesity cases in 1975 has nearly tripled, with 1.9 billion adults over the age of 18 of the world population of approximately 7.4 billion in 2016 being overweight (Hill 2018), which is similar in proportion to the 2020 statistics (Kaneda, Greenbaum and Kline 2020). This rise in obesity has been associated with an increase in the development of RA, specifically with the ACPA-negative subtype of RA. An obese female with a BMI of ≥ 30 kg/m$^2$ carries a reported three-fold increased risk by comparison with that of an individual with a normal BMI of 18.5 to < 25 kg/m$^2$ (Pedersen *et al.* 2006b; Myasoedova *et al.* 2020). This is supported by the findings of Liu *et al.* (2019) that BMI acted strongly as a facilitator between PA and RA, specifically sero-positive RA. However, to better understand this link as it stands, further evidence evaluating metabolic factors in pathophysiology and risk is required (Liu *et al.* 2019; Myasoedova *et al.* 2020), particularly as the inverse has been reported for BMI and ACPA-positive RA in male patients. Further research is also required to understand the link between BMI and serology phenotype (Myasoedova *et al.* 2020) as it may be linked to a common base of chronic inflammation because of the elevated plasma levels of IL-6, TNF-α, and C-reactive proteins (CRP) (Dandona, Aljada and Bandyopadhyay 2004) as well as to the fact that obesity exaggerates

inflammatory responses to stressors (Brydon *et al.* 2008; Wirtz *et al.* 2008; Kiecolt-Glaser, Gouin and Hantsoo 2010; Peled and Fisher 2014; Seo *et al.* 2015).

### 2.4.5.6 Tobacco smoking

Tobacco smoking is the only well-established environmental risk factor in the development of RA (Valesini *et al.* 2015; Ilar *et al.* 2018), and shows a selective profile for the increased risk of the subtype ACPA-positive RA. An OR of 1.65 has been reported with a 95% confidence interval (CI) between 1.03 to 2.64 when exposure to tobacco smoke is above twenty pack years compared to that of zero pack years (Pedersen *et al.* 2006b; McInnes and Schett 2011; Smolen, Aletaha and McInnes 2016; Safiri *et al.* 2019). According to the data from the period between 2000 and 2016, age-standardised use of tobacco has declined in terms of the prevalence rate in both male and female populations worldwide (Organization 2019). This is significant as the trends of tobacco use influence the prevalence of RA globally. The underlying mechanism is found in the pulmonary stress induced by the inhalation of tobacco smoke, silica, asbestos, and organic solvents. By contrast with tobacco, there is limited research supporting the association between the other stressors and the risk of developing RA that they introduce (McInnes and Schett 2011; Valesini *et al.* 2015; Ilar *et al.* 2018). The action of tobacco is believed to stimulate the inflammatory reaction in ACPA-positive RA since the stressor stimulates the PADI4 mediating the citrullination of proteins of the respiratory mucosal tissue (McInnes and Schett 2011; Valesini *et al.* 2015; Ilar *et al.* 2018). In addition, the stimulation of the antigen presentation by the immune system and the triggering of the loss of autoimmune tolerance induces the autoimmune response underlying the pathophysiology of RA (Ilar *et al.* 2018). These physiological mechanisms may be further enhanced by the association between the use of tobacco and the presence of negative emotion (Wang *et al.* 2016; Graham-Engeland *et al.* 2019) and levels of stress (Hamer and Stamatakis 2008; Kiecolt-Glaser, Gouin and Hantsoo 2010; Wang *et al.* 2016) (see Section 2.4.4).

### 2.4.5.7 Coffee intake

Coffee consumption has been associated with an increased risk of development of ACPA-positive RA (Martins and Fonseca 2019), with higher coffee consumption associated with greater RA risk. An OR was reported at 2.18 for more than ten cups vs zero cups of coffee drunk per day. However, when this was adjusted for tobacco smoking because of the habitual association between drinking coffee and smoking and the well-established risk associated with smoking the risk was diminished (Pedersen *et al.* 2006b), suggesting that coffee intake either has an associated risk when combined with other risk factors like smoking or plays a minor role in imparting the risk of RA development or, with reference to section 2.4.4 above, that it may be a proxy for an underlying factor such as a negative emotional state (Graham-Engeland *et al.* 2019). This provides an example of the complex combination of genetic and environmental risk factors

culminating in the loss of immune tolerance and predisposing to RA (McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019).

#### 2.4.5.8 Alcohol intake

An association between alcohol intake and the risk of developing RA has shown an inverse dosage risk, in that five or more drinks a week are shown to reduce the risk of developing RA, specifically the ACPA-positive subtype whereas no alcohol intake increased the risk. An OR of 1.96 has been reported for the risk associated with zero alcohol intake on the development of RA. Research has further associated a protective function of alcohol intake, especially wine, with regard to RA development. It is possible that the intake of alcohol down-regulates the production of pro-inflammatory cells (Di Giuseppe *et al.* 2012) and / or has an association with a positive emotional state (Graham-Engeland *et al.* 2019), thus suppressing immune mechanisms involved in the pathophysiology of RA. However, more research on the effects of alcohol with regard to RA is required (Pedersen *et al.* 2006b).

#### 2.4.5.9 First degree relative of a schizophrenic patient

Parents of psychiatric (i.e., schizophrenic) patients have had a significantly higher prevalence of ACPA-positive RA when compared to parents of non-psychiatric comparative subjects. An OR of 4.18 has been reported for risk of development of RA in first-degree relatives of a schizophrenic patient (Pedersen *et al.* 2006b). In addition, a reduced risk of a schizophrenic patient developing RA themselves has been noted. However, a large number of factors need to be considered as covariates, and thus further research into the topic is necessary (Pedersen *et al.* 2006b). There is literature contesting this relationship (Sellgren *et al.* 2014), and it may be a false association that is actually reflective of stress (Kiecolt-Glaser, Gouin and Hantsoo 2010; Graham-Engeland *et al.* 2019), the use of stress relievers such as tobacco (Valesini *et al.* 2015; Wang *et al.* 2016; Ilar *et al.* 2018) and the increased levels of physical stress associated with care of the psychiatric patient (Ilar *et al.* 2018).

#### 2.4.5.10 Gastrointestinal tract and the microbiome

The human digestive tract hosts an ecosystem of microorganisms and the system's genome (known as the microbiome), between which a symbiotic relationship has developed with many functions, among them nutrient metabolism and vitamin synthesis (Bodkhe, Balakrishnan and Taneja 2019). The microbiome develops, continuously undergoes changes and diversifies over a period starting pre-partum to around three years of age, at which time the microbiome is believed to stabilise. Thereafter, the host's diet and environmental exposure regulate the microbial composition (Bodkhe, Balakrishnan and Taneja 2019). Analysis of the microbiome and its effects has significant hurdles to overcome as only 1% can be cultured. DNA analysis has contributed to

understanding uncultivable microorganisms, but on-going research efforts are required (Bodkhe, Balakrishnan and Taneja 2019).

The ecosystem of microorganisms and the system's genome has attracted research focusing on the role of the microbiome in the pathophysiology with regard to its effect, risk association and the progression of the condition, of RA (McInnes and Schett 2011; Zhang *et al.* 2015; Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019). This is based on the action of the lymphoid tissue of the digestive tract that functions to maintain immune-homeostatic conditions so as to maintain host health status. The microbes show binding affinity with nucleotide-binding oligomerization domain receptors (i.e., (NOD)-like receptors) and TLRs through which they activate the immune system. Short-chain fatty acids, a metabolite product of the microbiome, can interact directly with the host. The composition and balance of the microbiome dictates the type of influence the fatty acids impart to the host's immune system. When well-balanced the microbiome helps maintain the immune tolerance of the symbiotic relationship. Disruption of these relationships and the microbiome's compositional balance has the capacity to act adversely on the mucosal-microbiome barrier through which the host's systemic immunity may be stimulated (Bodkhe, Balakrishnan and Taneja 2019).

Animal model studies present data that suggests an association between microbial dysbiosis and changes in disease activity and inflammatory markers in RA pathophysiology, and also that DMARD therapy is influenced by the microbiome (Bodkhe, Balakrishnan and Taneja 2019). Furthermore, initial human-based studies have implicated dysbiosis within the gastrointestinal system in the development of RA in the early RA patient (Smolen, Aletaha and McInnes 2016). One of these studies noted changes in the microbiome and associated CRP and the ACPA status (specifically ACPA-positive RA) of RA patients (McInnes and Schett 2011; Smolen, Aletaha and McInnes 2016). Not unlike the association between infection and RA (see Section 2.4.3), a comparative microbial composition study showed stimulation of a Th1 response with excess *Prevotella copri*, *Eggerthella lenta* and *Collinsella* in RA patients. By contrast, increased *Prevotella histicola* and reduced bacteroides have shown a suppressive action on arthritis (Bodkhe, Balakrishnan and Taneja 2019). Thus, dysbiosis influences inflammatory marker production and disease activity (Bodkhe, Balakrishnan and Taneja 2019).

The contribution of the gastro-intestinal microbiome to the development and then also the management of RA is further complicated by the impact of medication (DMARDs) (McInnes and Schett 2011; Smolen, Aletaha and McInnes 2016). The immunological modulatory properties of DMARDs, in addition to their primary function, they also have the capacity to act both directly and indirectly on the microbiome, providing modulating properties over it and therefore impacting the RA (Bodkhe, Balakrishnan and Taneja 2019). For example, sulfasalazine (SSZ) can only convert into its active form after the action of anaerobic bacteria that form part of the microbiome, but at

the same time SSZ also has an antibacterial action on the microbiome (Sayers, MacGregor and Carding 2018; Bodkhe, Balakrishnan and Taneja 2019).

Studies have shown that dysbiosis in RA patients was restored to normal, either partially or fully, or that the protective microbial population was increased with methotrexate (MTX) or hydroxychloroquine (HCQ) therapy, which impart a therapeutic change in RA activity (Sayers, MacGregor and Carding 2018; Bodkhe, Balakrishnan and Taneja 2019). Limited data is available to indicate the same for SSZ and chloroquine (CQ), despite their antibacterial properties (Bodkhe, Balakrishnan and Taneja 2019).

As can be seen in the preceding sections, it is important to understand RA not only in terms of the impact it has on patient's, their social-framework, the health sector, and the risk factors that predispose the patient to developing RA (see Sections 2.4), but also in terms of its pathogenesis and pathophysiology in order to provide RA patients with optimum treatment and management plans that are not only effective but also affordable. The next section of this chapter deals with the pathophysiology of RA.

## 2.5 The pathophysiology of RA

RA is an agnogenic disease, as noted in the definition (see Section 2.2). As research endeavours to clarify the pathophysiology, new aspects of RA come to light and offer insight, but the picture is as yet incomplete (McInnes and Schett 2011; Yang *et al.* 2017), reflecting the complexity and dynamic nature of the condition (McInnes and Schett 2011; Yang *et al.* 2017). With RA being an inflammatory autoimmune disease with a systemic impact primarily manifesting in the joint complex, synovial tissue, tendons, and bursae (Lewis and Battaglia 2019), it is often characterised by the infiltration of inflammatory cells, mainly macrophages and T-cells, and abnormal vasculature (e.g., the capillary network that surround the synovial membrane) (Yang *et al.* 2017; Chandrupatla *et al.* 2019). This synovitis develops for the following reasons:

- the infiltration of synovium by immune cells (B-cells, T-cells, DC, neutrophils and macrophages) and their immune products,
- the extravasation from the surrounding micro-vasculature,
- the activation of the micro-vascular endothelium increases expression of chemokines and adhesion molecules (such as alpha-v-beta-3-integrin (αvβ3) and E-selectin), and
- the increases in vascular permeability (McInnes and Schett 2011; Yang *et al.* 2017; Chandrupatla *et al.* 2019).

Even though research into RA continues, the emergence of several coincidental mechanisms relating to the presence of the auto-reactive immune cells primarily responsible for RA (i.e., cells that have lost immunological tolerance) shows immunological loss in the central and peripheral adaptive immune strategies. This leads to uncontrolled self-reactive B- and T-cell activation (Hosseini et al. 2015). The resultant autoimmunity theory is further enhanced by the innate

immune strategy (Hosseini et al. 2015). As it stands, a currently accepted hypothetical explanation of RA onset is extracted from the publication by Chandrupatla *et al.* (2019) which states that a two-stage trigger mechanism is responsible for the condition. Firstly, citrullination is accelerated in genetically susceptible patients (see Section 2.4.1: Genetic factors and Section 2.5.3: Citrullination) in extra-articular locations (examples can be found in Section 2.4.5.6: Tobacco smoking, Section 2.4.3: Infection factors, Section 2.4.5.10: Gastrointestinal tract and the microbiome). This citrullination induces production of RF, ACPA, and a-CarP. Following this, even years later, an unrelated episode of synovial inflammation and citrullination in the presence of RF, ACPA, and / or a-CarP may induce the chronic synovial inflammation that evolves into RA through the binding of antibodies and antigens within the synovial joint complex.

### 2.5.1 Innate immune system

#### 2.5.1.1 Macrophages

Macrophages are found within the synovial membrane and hold several known functions. These macrophages hold a central role in the development of synovitis, acting via cytokine synthesis such as ILs and TNF- α, ROS, nitrogen intermediates, prostanoids, matrix metalloproteases (MMPs), phagocytosis and antigen presentation of the T-cells (McInnes and Schett 2011; Chandrupatla *et al.* 2019). The predominant subtype of macrophages in RA tend to be M1 macrophages, which encourage inflammation as opposed to the M2 subtype that decreases inflammation and stimulates tissue repair (McInnes and Schett 2011; Mills 2012; Chandrupatla *et al.* 2019). The M1 macrophages therefore play a central role in maintaining local inflammation and the progression of the pathology (Peled and Fisher 2014; Chandrupatla *et al.* 2019).

Macrophage (M1) activation is facilitated by toll-like receptors (TLR), NOD-like receptors, folate receptors, scavenger receptors (SR), clusters of differentiation (CD)14 and 44 and VIP receptors (McInnes and Schett 2011; Chandrupatla *et al.* 2019). In addition, activation can be achieved by means of a number of mechanisms relating to cytokine stimulation, T-cell interaction, immune complexes (binding of antigen to antibody), lipoproteins, liver x-receptor agonists, proteases-rich microenvionment through protease-activated receptor 2, and microRNA species (e.g., MicroRNA-155) (McInnes and Schett 2011).

TLRs are transmembrane receptors that fall into the category of pattern recognition receptors (PRRs). PRRs are able to recognize both pathogen-associated molecules (PAMPs) and damage-associated molecular patterns (DAMPs). Therefore, TLRs recognize molecular patterns associated with pathogens, identifying them to the immune system. However, they also recognize self-proteins and endogenic nucleic acid, along with other signals such as DAMPs. Each TLR recognizes specific patterns. For example, TLR2 recognises bacterial membrane molecules, whereas TLR8 recognises single strand RNAs (Hosseini et al. 2015). Therefore, TLR signalling

holds a central role in the activation of the immune system, particularly the adaptive immune system. This is achieved by means of cytokine production and the stimulation of co-stimulatory molecules in the antigen-presenting cell. Under the construct that the TLR links the innate immune system to the adaptive immune system, activation of the TLR in an inappropriate fashion would contribute to or perpetuate the pathogenesis of RA (Hosseini et al. 2015). The noted effect of activation of the TLR on binding to ligands is to induce signalling cascades, resulting in the expression of immune response genes and thus the production and release of pro-inflammatory cytokines, ILs, and TNF- α, as well as the expression of chemokines and other co-stimulatory molecules. This collectively enhances the immune response (Hosseini et al. 2015). Supporting this, it is reported that endogenous ligands are present in arthritic joints. These ligands are able to act on TLRs and activate their cascade. Within a rheumatic joint, necrotic synovial tissue will release RNA, which activates TLR3 receptors on the synovial fibroblasts (Brentano *et al.* 2005). Therefore, endogenous ligands are present in arthritic joints (Roelofs *et al.* 2006). As is supported by the fact that, fibronectin fragments have been detected within the rheumatic synovium playing two roles: firstly, they induce the expression of MMPs via the MyD88 dependant TLR2 in chondrocytes and secondly, along with heat-shock protein (which can activate both TLR2 and 4), they act on fibroblast-like cells of the rheumatic synovium leading to the expression of additional domain A containing fibronectin (Hosseini et al. 2015).

Macrophage activation by lipoproteins is via either direct or indirect mechanisms (Peled and Fisher 2014). There are three mechanisms where direct activation can occur:

- The accumulation of lipoproteins, both intracellular and extracellular, forms crystals. Macrophage phagocytosis of the extracellular crystals and the formation of intracellular crystals activate NOD-, LRR- and pyrin domain-containing protein 3 (NLRP3) (an intracellular sensor detecting, among others environmental irritants), producing NLRP3 inflammasomes, and releasing pro-inflammatory cytokines (Peled and Fisher 2014; Swanson, Deng and Ting 2019).
- Oxidized LDL mediated by enzymatic activity (such as that of lipoxygenase and myeloperoxidase) and ROS, acts on SR and TLRs on the macrophage. The degree of oxidation of LDL has been proved to induce different responses (Dandona, Aljada and Bandyopadhyay 2004; Brydon *et al.* 2008; Wirtz *et al.* 2008; Kiecolt-Glaser, Gouin and Hantsoo 2010; Seo *et al.* 2015). Low-level oxidation (Low-oxLDL) acts on M1 subtype macrophages, inducing pro-inflammatory cascades, while high-level oxidation (Hi-oxLDL) acts on M2 subtype macrophages, inducing anti-inflammatory cascades (Peled and Fisher 2014; Seo *et al.* 2015; Chandrupatla *et al.* 2019).
- An increase in cholesterol in the plasma membrane of the macrophage influences the microenvironment of TLRs, inducing an increased sensitivity to ligands, heightening the inflammatory responses (Peled and Fisher 2014).

There are two mechanisms where indirect activation can occur;

- Local T-cell activation and the secretion of pro-inflammatory cytokines promote the activation of macrophages (Peled and Fisher 2014).
- The second indirect activation mechanism of macrophages can be achieved by means of oxidized LDL in the synovium and synovial fluid (as can be seen by the presence of foam cells in the synovium and synovial fluid) (Dandona, Aljada and Bandyopadhyay 2004; Brydon *et al.* 2008; Wirtz *et al.* 2008; Kiecolt-Glaser, Gouin and Hantsoo 2010). These changes, along with the hypoxic and acidic environment of the lesion, result in a shift towards heavy reliance on glycolysis to produce energy. M1 subtype macrophages tend to favour glycolysis while M2 subtype macrophages favour fatty acid metabolism in order to produce their required energy (McInnes and Schett 2011). This shift towards glycolysis promotes M1 over M2 macrophage activity, the former being pro-inflammatory (Peled and Fisher 2014).

#### 2.5.1.2 Neutrophils

Neutrophils release several products, among them proteolytic enzymes and oxidative products (Wright *et al.* 2010; Cassatella *et al.* 2019) that may play a role in the inflammatory synovitis seen in RA (Wright *et al.* 2010; O'Neil and Kaplan 2019). Neutrophils therefore feature significantly in the inflammatory response and are seen in large numbers within the RA joint's synovial fluid and tissue (Wright *et al.* 2010; O'Neil and Kaplan 2019). With appropriate activation, the neutrophil has the capacity to inflict damage on the joint complex (bone, cartilage and associated soft tissue) (Wright *et al.* 2010). The injury is affected by the synthesis and release of proteases and ROS, while there is continued induction and maintenance of inflammation via cytokine, prostaglandin, chemokine, and leukotriene synthesis, as well as the antigen presentation required to activate the T-cell immune response (Wright *et al.* 2010; Cassatella *et al.* 2019; O'Neil and Kaplan 2019).

#### 2.5.1.3 Natural killer cells (NK)

Subsets of NK such as the CD56 increase in the synovium of an inflamed joint, as is seen in RA. An increase in NK cells in the joint complex then results in an increased synthesis and release of IFN-γ resulting in an enhanced immune reaction as IFN-γ plays a major role in the stimulation of both the innate and adaptive immune system (Shegarfi, Naddafi and Mirshafiey 2012). This activates monocytes / macrophages and induces the expression of MHC class II on the cells. The IFN-γ plays a role in the production of ROS as well as in cytokine production, antigen presentation, creation of metabolic pathways, cellular differentiation including macrophage activation (polarization towards M1 subtype) and cell growth and survival (Shegarfi, Naddafi and Mirshafiey 2012). The fact that MHC class II variation (shared epitope) is the strongest genetic risk factor for RA raises a simple hypothesis, namely, that the MHC class II-inducing cytokine IFN-γ contributes to the development of RA (Kato 2020). Additionally, the NK cells have the capacity to direct

monocyte differentiation to a mature DC. TNF-α represents a potent osteoclastogenic cytokine that has been linked to RA. This cytokine induces receptor acquisition by NK cells and the combination with IL-15 enhances this effect. Furthermore, TNF-α plays a role in NK-dependant DC maturation. It is worth noting the macrophages and monocytes are the major producers of TNF-α in RA (Shegarfi, Naddafi and Mirshafiey 2012).

### 2.5.2 Adaptive immune system

If one considers the adaptive immune system and its function in the pathogenesis of RA within a joint complex, it is evident that numerous granulocytes, monocytes, macrophages, and DC (collectively known as myeloid cells) are present, along with the plasmacytoid dendritic cells (pDCs). These cells synthesize the necessary immune-active molecules to stimulate T-cell activation (McInnes and Schett 2011). Despite this and the fact that T-cells are present in the synovial joint, significant insight into their role is still lacking.

#### 2.5.2.1 T lymphocytes and their roles:

T lymphocytes play a role in immunomodulation under the influence of sex hormones. These Th cells are classified by cytokine production and include Th1, Th2, Th17, and T regulatory (Treg) cells. The complexity of the interaction between the Th cell lineages appears to demonstrate a degree of plasticity but this is yet to be fully elucidated (Tan, Peeva and Zandman-Goddard 2015). For example, the Th1 cells produce IL-2, TNF and IFN-γ and thus promote cellular immunity by driving T lymphocytes to fight viral and microbial infection and cancer and are also responsible for delaying the hypersensitivity reaction in skin (Tan, Peeva and Zandman-Goddard 2015). By contrast Th2 cells produce IL-4, 5 and 13, lymphocytes which also act on the B lymphocytes (humoral immunity) against the invasion of microbes through driving antibody production (Tan, Peeva and Zandman-Goddard 2015). The Th17 cells produce IL-17, a pro-inflammatory cytokine involved in chronic inflammation associated with RA (Straub 2007; Tan, Peeva and Zandman-Goddard 2015). Lastly, the Treg cells produce anti-inflammatory cytokines such as IL-10 and TGF-β, which play an important role in immune tolerance peripherally and act to prevent autoimmunity developing. Treg cells maintain tolerance by acting on the T lymphocytes (CD4+ and CD8+) by inhibiting their proliferation and blocking cytokine production. Furthermore, they act on B cell lymphocytes, inhibiting antibody production, on NK activation by means of cytokines, and on DC maturation (Tan, Peeva and Zandman-Goddard 2015). Autoreactive T-cells acting against citrullinated self-proteins have been identified in RA patients' synovium. Synovial T-cells oligoclonality, germinal-centered reactions, and B-cell hypermutation suggest ongoing local antigen-specific, T-cell-mediated B-cell activation (McInnes and Schett 2011).

### 2.5.2.2 The plasmacytoid dendritic cell (pDC)

The pDC is a rare sub-type of immune cell found in primary lymphoid tissue (i.e., bone marrow and thymus) and secondary lymphoid organs (i.e., lymph nodes and tonsils). These cells also account for an estimated 0.01 to 0.05% of the mononuclear cells in peripheral circulation; however, they are rarely detected in splenic and mucosal lymph tissue. The migration of pDCs into the lymph nodules is achieved by means of the E-selectin-driven mechanism utilizing cell adhesion molecules (Jegalian, Facchetti and Jaffe 2009). With regard to the lineage of pDC cells, they appear to lack markers specific to a myeloid, monocytic lineage as well as for the B-cells, T-cell and NK cell germline. This indicates that, although the pDC is classified as a mononuclear cell, it is distinct, as it remains inconclusive whether the cell is myelomonocytic-derived or lymphoid-derived (Jegalian, Facchetti and Jaffe 2009). The pDCs collect in clusters or less frequently in a scattered arrangement in inflammatory tissue and promote the inflammatory response. They are drawn to the site of inflammation by the action of chemokine-driven signalling along with the E-selectin-driven mechanism, utilizing cell adhesion molecules to migrate the pDCs from circulation into the correct tissue (Jegalian, Facchetti and Jaffe 2009). Stimulation of the pDC cell by means of CD40 ligands and IL-3 or by means of PAMPs causes the cells to produce dendritic projections assuming the mature DC morphology and functionality. This stimulates both cytotoxic T-cells (CD8+) and the T helper cell (CD4+) (Jegalian, Facchetti and Jaffe 2009). The pDC cell has two distinct and separate states, viz., the plasmacytoid state and the dendritic state. The plasmacytoid state produces and releases large quantities of IFN type 1 (which is associated with RA) (Muskardin and Niewold 2018). In an almost give and take relationship, the plasmacytoid state expresses little MHC type 2, and the dendritic state express high levels of MHC type 1 and 2 along with a collection of co-stimulatory molecules for the T-cell (Taneja and David 2001). In the active dendritic state, these cells may play a role in antigen presenting in the RA patient (Jegalian, Facchetti and Jaffe 2009). Stimulation of the pDC's TLR9 (recognition of bacterial and viral DNA with hypomethylation of the CpG sequence) and TLR7 (recognizing viral single strand RNA) results in the synthesis and release of large quantities of INF and cytokines (i.e., TNF-α, IL-6 and 8), which are cornerstone markers in RA inflammation. It is important to note that the TLRs may play a role in self-RNA and self-DNA detection, which may activate an immune cell like the pDC in an autoimmune disease setting (e.g., RA), with these cells accumulating in active sites of a variety of inflammatory conditions (Jegalian, Facchetti and Jaffe 2009). The molecules expressed because of pDC TLR9 and 7 activation will induce several responses:

- the recruitment of activated T-cells and NK cells to the site of inflammation,
- the activation of macrophages and NK cells, and the inducing of Th cell differentiation into Th cells releasing INF and IL10, or IL4.
- the inducing of the differentiation of plasma cells.

In addition to the INF-1's direct effect on viruses, it will enhance the function of DC of the myeloid germline (Jegalian, Facchetti and Jaffe 2009), therefore perpetuating the inflammatory process in RA. The pDCs hold an important role in linking the innate and adaptive immune systems and providing an immunomodulation function; they induce a wide range of influences on these immune divisions, with immune tolerance and modulation apparently among them. Research identifies an influence of this nature over the Treg cells in RA patients (Jegalian, Facchetti and Jaffe 2009).

### 2.5.3 Citrullination

Citrullination is the mechanism by which there is a post-translation modification of the protein-bound amino acid arginine into citrulline. Evidence links the role of citrullination with the synthesis of ACPA, which is strongly associated with RA pathophysiology. The modification is catalysed by the PADI enzyme. Each of the conversions into citrulline influences structural and the characteristics of the amino acid, including its acidity, iso-electric point, hydrogen bond capacity and interaction with others. These functional changes and the half-life of the protein they comprise generate novel epitopes and autoantigens which alter immune-tolerance, generating an autoimmune response typical of RA (Valesini *et al.* 2015). Five PADI isoforms have been identified in different tissue and exerting different functions. PADI1 is found in the epidermis and the tissue of the uterus. PADI2 is found in muscle tissue, CNS, hematopoietic cells and respiratory tissue. PADI3 is found in hair follicles. PADI4 is found in immune cells (neutrophils and eosinophils), respiratory, splenic and secretory gland tissue. Finally, PADI6 is found in reproductive, intestinal, splenic, respiratory, hepatic and muscle tissue (Lugli *et al.* 2015; Valesini *et al.* 2015). PADI4-mediated citrullination alters the chemokine physiological function and the neutrophil extracellular trap (NET) formation, both implicated in autoantigen synthesis related to RA. This is triggered by a number of stressors as described in the risk factors (see Section 2.4), leading to ACPA production in individuals that have shared epitope susceptibilities related to the HLA molecule and / or PTPN22. Evidence has shown that PADI4 can undergo auto-citrullination, which may inactivate the enzyme but can also modify the enzyme-enhancing auto-antibody (anti-PADI4-antibodies) recognition (Valesini *et al.* 2015). ACPA represents a collection of antibodies that recognise proteins and peptides that contain citrulline such as cyclic-citrullinated peptide (CCP), citrullinated fibrinogen and citrullinated vimentin. Evidence is currently inconclusive regarding the immune reactivity towards these structures by the ACPA in the inflamed joint; but there are 53 identified citrullinated proteins in the RA patient's serum and synovium (Valesini *et al.* 2015). Literature suggests that citrullination is a process that is inflammatory-dependant and not disease-dependant. This introduces another role that citrullination has in several chronic inflammatory conditions such as RA that is made evident by the presence of citrullinated protein in the RA synovium and is associated with a ACPA concentration in the synovium that is higher than that of serum concentration (Valesini *et al.* 2015).

### 2.5.4 Microvascular changes and abnormal vasculature

Increased vascular permeability is achieved by activation of the vascular endothelial cells (VECs) which induces a series of changes resulting in the formation of endothelial gaps. These gaps allow for exudate to leak into the affected joint, followed by monocyte migration and differentiation into macrophages (or DC, dependant on location), and the expression of inflammatory mediators (Yang *et al.* 2017). The inflammation, along with the acidic and hypoxic environment of the local tissue and the presence of cytokines, triggers neoangiogenesis of local capillary networks. This develops early in the pathogenesis of RA, enhancing the inflammatory mediator aggregation at the lesion site and the expression of adhesion molecules in the endothelium of the capillaries. The angiogenesis is crucial in the RA pathogenesis and core to this process are the VECs. The synovial tissue rapidly proliferates with the enhanced nutrient and oxygen delivery and develops the synovial pannus characteristic of RA. These synovial pannus formations will ultimately cause cartilage and subchondral bone erosions (McInnes and Schett 2011; Yang *et al.* 2017). The VECs' expression of αvβ3 and E-selectin (membrane proteins) holds a primary role in the process of angiogenesis and cellular migration in RA pathogenesis. The αvβ3 membrane protein has been shown to upregulate angiogenesis and is expressed in the RA synovium to a greater extent than seen in normal tissue, whereas the E-selectin membrane protein binds to leukocytes in the endothelium and is expressed in high concentrations at sites of inflammation while being nearly undetectable in normal tissue. Activated VECs in the RA lesion sites of inflammation are thus distinct from normal vasculature due to the phenotype and presence of αvβ3 and E-selectin (McInnes and Schett 2011; Yang *et al.* 2017). This angiogenesis is coupled with insufficient lymphangiogenesis which results in reduced lymphatic drainage and lymphedema (Yang *et al.* 2017). These changes to the RA lesion environment, combined with structural reorganization of the synovium architecture and local fibroblast activation, enhance, and enable the inflammatory state of the synovium denoted in RA (Yang *et al.* 2017; Chandrupatla *et al.* 2019).

### 2.5.5 Structural damage of RA

Yang *et al*. (2017) provide a general summation of RA by indicating that it is an "activation and recruitment of pro-inflammatory cytokines such as TNF-α, IL-1, IL-6 and transforming growth factor- beta (TGF-β)" which are involved in the pathological processes of RA by disrupting the immune balance. More specifically, TNF-α and IL-1β stimulate the release of tissue-degrading MMPs from synovial cells and TNF-α enhances the development of osteoclasts. In addition, the infiltration of macrophages, B-cells, T-cells, synoviocytes and fibrocytes in inflamed joints triggers the proliferation of synovial tissues and osteoclasts and the invasion of the synovium. These compounding processes ultimately lead to the erosion of joint cartilages and articular bone surfaces (Chandrupatla *et al.* 2019).

### 2.5.6 Phases of RA as a condition and clinical presentation

RA onset and progression has no specific character or estimated period. This variability means that RA can develop acutely and aggressively or develop and progress more slowly. Furthermore, as evidenced in the documented capacity of RA remission, periods of low disease activity and flare-ups, no strict designation can be given as to the disease's nature. For this reason, measures such as disease activity tools, of which the Clinical Disease Activity Index is one, and phases of disease have been devised, studied, and defined (Kaur, White and Bartold 2012; Walker *et al.* 2013).

A theoretical phase in RA development (preceding clinically evident arthritis) is considered a phase of systemic autoimmunity that could be present for several years. On average this period lasts five years prior to evident arthritis (Raza, Holers and Gerlag 2019). It may be detected through circulating autoantibodies (related to RA) and increased levels of inflammatory markers such as CRP, cytokines and chemokines, as noted in in the section dealing with those at risk (see Section 2.4). If the patient at risk presents with musculoskeletal symptoms accompanied by positive serology of ACPA and RF, they carry a 40% risk of a clinically diagnosed manifestations of RA in the subsequent two years (Chandrupatla *et al.* 2019; Raza, Holers and Gerlag 2019). This preclinical phase represents an opportunity to intervene on a preventative level by adjusting lifestyle risk factor exposures and using pharmaceuticals such as DMARDs pre-emptively in order to delay or prevent the onset of clinical manifestations (Raza, Holers and Gerlag 2019). This at-risk phase can be described by the European League Against Rheumatism (EULAR) phase nomenclature (Raza, Holers and Gerlag 2019):

- genetic risk,
- environmental risk,
- systemic autoimmunity,
- symptoms without signs of clinical arthritis,
- arthritis unclassified,
- RA diagnosis.

A patient can be described by one or more of the above, and progression from being at risk to the diagnosis of RA does not have to follow a linear path through all phases. From the point at which a patient presents with EULAR phase D, that patient is considered to have clinically suspected arthritis (CSA), and not only has associated circulating inflammatory markers but may already have circulating auto-antibodies. This phase of CSA may be the point at which the patient seeks medical advice for the musculoskeletal symptoms (Raza, Holers and Gerlag 2019). The patient with CSA can be differentiated from the clinically similarly presenting arthralgias by means of a 7-parameter criterion. If the patient presents with three of the following seven parameters with 90% sensitivity the criterion can identify it as CSA. If more than four parameters of the criterion are met, this can also identify CSA with 90% specificity. The seven parameters include:

- patient history of recurrent joint symptoms for at least a year,
- the symptom(s) located at the MCP joint,
- morning stiffness lasting at least an hour,
- symptom severity worst in mornings,
- a first-degree RA diagnosed relative,
- a positive squeeze test of MCP joints and / or
- difficulty clenching a fist revealed by a physical examination (Raza, Holers and Gerlag 2019).

Following the onset of RA, the patient will experience the following phases of disease progression (Steinbrocker, Traeger and Batterman 1949; Nakagawa *et al.* 2018):

- Phase 1 is considered early RA where joint pain, joint oedema and stiffness are present. However, there is no sign of the joint damage associated with RA.
- Phase 2 is considered moderate RA where synovitis (inflammation) begins to damage the joint, in particular the articular cartilage, with loss of joint mobility and range of motion (ROM) accompanied by the onset of pain.
- Phase 3 is considered progressive RA and symptoms tend to be severe. In this phase, the initial damage noted in phase 2 extends into the subchondral bone, inducing erosion of the bone and over time resulting in deformity. Pain and swelling are exacerbated in this phase.
- Phase 4 is considered end-stage RA. Joint inflammation has dissipated and, as a result of the damage to the joint complex, pain, swelling and loss of mobility increase and the joint itself may become ankylosed.

Typically, the RA patient will present with symmetrical polyarticular arthritis. The joints tend to be painful, swollen, and stiff. The stiffness is exacerbated in the morning – also known as “morning stiffness”, which improves throughout the day (Orange *et al.* 2020). This is most prevalent in small joints of the fingers, hands, wrists, and feet. However, large joint involvement is possible (Walker *et al.* 2013; Longmore *et al.* 2014). A less frequent set of disease presentation patterns in RA includes the sudden development of RA features (it is worth noting that acute onset RA is typically reported in older patients), with widespread distribution of articular inflammation. Another disease presentation pattern includes a palindromic or cyclic remission-relapsing pattern where there may be a mono-arthritic or poly-arthritic distribution. A final pattern is that of a persistent large joint mono-articular arthritis, most commonly the hip, knee, or shoulder joint (Walker *et al.* 2013; Longmore *et al.* 2014). Chronic pain associated with RA (in other words, pain experienced for three or more months) is a principal driver in the presentation of RA, and the patient’s medical seeking behaviour; it is also the prime cause of patient distress and functional impairment (Jadhav *et al.* 2011; Chen and Michalsen 2017). When joints are swollen and inflamed, they are tender to palpation and painful with active and passive movement (Walker *et al.* 2013). Generally, in early RA, patients tend to have inflammation associated with the MCP and proximal interphalangeal

(PIP) joints of the hands, the wrist and the metatarsal phalangeal (MTP) joints of the feet. At this stage, however, the joints do not exhibit articular damage. These findings tend to be accompanied by tenosynovitis and / or bursitis around these articulations (Longmore *et al.* 2014). In the later stages of RA, patients tend to exhibit articular damage that progresses to deformity. This is frequently seen in patients with long-standing RA that is not under appropriate control and may manifest in any of the following ways (Walker *et al.* 2013; Longmore *et al.* 2014):

- In the upper limb, the fingers may develop either Swan-neck and or Boutonniere (or buttonhole) deformities, and the thumbs could develop Z-deformities. The hand unit ulnar deviates at the wrist complex (most commonly radio-ulnar joint distally), potentially subluxating dorsally. Extensor tendons (most commonly the fourth and fifth) of the hand are commonly inflamed and may rupture (Longmore *et al.* 2014). Furthermore, the flexor tendons of the hand are prone to the development of nodules (Walker *et al.* 2013). Occurrence of adhesive capsulitis, carpal tunnel and De Quervain's (Walker *et al.* 2013; Longmore *et al.* 2014; Tokem *et al.* 2014), as well as extensor RA nodules around the elbow (Longmore *et al.* 2014).
- In the lower limb, the foot tends to develop cock-up toe deformities as a result of MTP joint dorsal subluxation. With weight-bearing, the exposed metatarsal heads are painful and are associated with adventitious bursae and formation of callus (Longmore *et al.* 2014). Eversion alignment deformity of the hindfoot is common and associated with damage to the subtalar joint, whereas collapse of the longitudinal arch is due to rupture of the tibialis posterior tendon. The knee/leg manifestations include synovitis and resultant popliteal cysts (Baker's cyst) (Walker *et al.* 2013), also Achilles tendon nodules (Longmore *et al.* 2014).
- Additionally, chronic spinal pain is a frequently reported feature. A particularly important manifestation is atlanto-axial subluxation/instability. This rare occurrence may present a threat to the spinal cord (Longmore *et al.* 2014).

### 2.5.7 Extra-articular and systemic manifestations of RA

RA is not limited to the articular system. Extra-articular and systemic manifestations include variable combinations of fatigue, weight loss, and loss of muscle mass (rheumatoid cachexia, reported in two thirds of all RA patients), as well as generalised lymphadenopathy, pericarditis, and polymyalgia rheumatica (Cooney *et al.* 2011; Walker *et al.* 2013; Longmore *et al.* 2014; Tokem *et al.* 2014). Some patients develop pyrexia / fever and / or chronic fatigue (Matsushita, Yamaji and Tamura 2020).

The majority of extra-articular manifestations are reported in the respiratory system. It should be noted that these pulmonary manifestations or diseases may introduce risk factors related to cardiac pathology, for example heart failure and pericarditis (Matsushita, Yamaji and Tamura 2020). Bronchiectasis triggered by bronchial inflammation occurs in 12 – 62 % of RA patients and

is associated with high levels of ACPA and RF (Wilczynska, Condliffe and McKeon 2013; Matsushita, Yamaji and Tamura 2020). Patients with genetic risk factors or who are long-term sufferers of RA and in addition smoke have an increased risk of this extra-articular manifestation. Another commonly reported respiratory manifestation in RA patients is organizing pneumonia as compared to interstitial pneumonia; the later occurs in about 50% of RA patients (Matsushita, Yamaji and Tamura 2020). Other respiratory manifestations that may occur include bronchiolitis, respiratory bronchiolitis / obliterative bronchiolitis, pulmonary rheumatoid nodules (Matsushita, Yamaji and Tamura 2020), fibrosing alveolitis, and pleural and pericardial effusion (Longmore *et al.* 2014). All the respiratory complaints tend to result in pleurisy in male RA patients. Additionally, the use of immunosuppressant drugs such as steroids and DMARDs for RA increases the risk of *Pseudomonas aerufinosa infection.*

Vasculitis, also known as rheumatoid vasculitis or malignant RA, occurs in up to 1% of RA patients and is the most common vascular manifestation. This develops in patients with high ACPA and RF levels and is associated with an elevated CRP and erythrocyte sedimentation rate (ESR). Vascular inflammation is associated with hyperglobulinemia and hypocomplementemia. Common vasculitis complications include:

- ocular manifestations such as scleritis, which is influenced by the severity of the RA and may remit spontaneously or necrotize and perforate in recurring cases, as well as episcleritis, scleritis, scleromalacia, keratoconjunctivitis sicca and or Sjogren's syndrome (Longmore *et al.* 2014; Matsushita, Yamaji and Tamura 2020),
- dermal manifestations (most commonly skin ulcers, but if erythema is present it suggests sepsis (Walker *et al.* 2013)),
- neurotrophic vasculitis, which may induce Raynaud’s phenomenon and / or a peripheral neuropathy (Longmore *et al.* 2014; Matsushita, Yamaji and Tamura 2020),
- systemic vascular inflammation, which is variably associated with pyrexia, fatigue, gastrointestinal bleeds, pericarditis, and loss of weight (Matsushita, Yamaji and Tamura 2020),
- microcytic-hypochromic or normocytic-normochromic anemia are other complications of RA secondary to the chronic inflammation, impairment of iron utilisation and hypersplenism. Blood tests reveal decreased serum iron levels, total iron binding, and increased ferritin (Longmore *et al.* 2014; Matsushita, Yamaji and Tamura 2020),
- the relative risk of myocardial infarction, estimated to be twice as common in women with RA relative to those without (Roubenoff *et al.* 1994), is one of the larger complications of RA. Furthermore, cardiovascular disease (CVD) events typically occur a decade earlier and to a greater extent in patients with RA relative to healthy controls (Del Rincón *et al.* 2001; Maradit-Kremers *et al.* 2004). This is supported by Avina-Zubieta *et al.* (2012), who indicate that there is a 50% increased risk of CVD deaths in patients with RA compared with the general population (Cooney *et al.* 2011).

Amyloidosis production and accumulation in organs (RA is associated with serum amyloid A) induce dysfunction resulting from chronic and severe inflammation; most frequently affected are the cardiac, renal, and gastrointestinal systems. It is essential that blood serum amyloid A levels be monitored (Longmore *et al.* 2014). A rare finding in RA patients is Felty's syndrome (Longmore *et al.* 2014) or pancytopenia, which may result from supression of the patient's bone-marrow by immunosuppressant therapy as in the case of MTX-lymphoproliferative disorder (Matsushita, Yamaji and Tamura 2020).

## 2.6 Diagnostics

When diagnosing, the physician should make use not only of the clinical history and findings of the physical examination but also the results of a number of special investigations and studies in order to aid the diagnosis and confirm not only that the patient has RA, but also the activity of the disease and the serological type.

Among the blood-based investigations physicians could include inflammatory markers (i.e., CRP and ESR), a full blood count (FBC) and, most importantly, an autoimmune panel including RF and ACPA. It is worth noting that the auto-antibodies detected in the autoimmune panel related to RA are clinically present in the patient's blood years before the disease would otherwise be diagnosed, indicating their expression and presence in the blood before manifestations of the condition are noticed (see Section 2.5.5 to 2.5.7) (Hughes 2012; Valesini *et al.* 2015; Chandrupatla *et al.* 2019).

- RF is positive in an estimated 70% of RA patients, while in Sjogren's syndrome and Felty's syndrome RF is positive in almost 100% of the cases. The higher the titre for this autoimmune serological marker the greater the association with a more severe form of RA with greater risk of bone erosion and extra-articular manifestations to occur (Longmore *et al.* 2014).
- ACPA has the highest specificity (98%) for RA when compared with other autoimmune markers (Longmore *et al.* 2014; Valesini *et al.* 2015). It has been reported that auto-reactive B-cells synthesise both RF and ACPA, with ACPA proving both sensitive and specific as a serological marker. Evidence shows that ACPA is associated with a more severe form of the disease, as well as with bone erosion, reduced remission rates and extra-articular manifestations (Valesini *et al.* 2015; Chandrupatla *et al.* 2019; Lorenzetti *et al.* 2019).

The FBC may show evidence of the anaemia of chronic disease, with raised platelets and raised inflammatory markers, including ESR and CRP, present (Longmore *et al.* 2014). Imaging investigations often include a bilateral hand x-ray series showing early signs of soft tissue swelling, loss in joint space, and juxta-articular osteopenia. As the diseases progresses, bone erosions develop with carpal bone destruction and / or subluxation may occur (Longmore *et al.* 2014). Additional studies may include ultrasound, Computerized Tomography and Magnetic

Resonance Imaging, where RA-related articular findings such as bone erosions may be seen at an earlier stage compared to conventional radiography in addition, synovitis in areas suspected of disease activity may be more easily detected (Longmore *et al.* 2014). When extra-articular and systemic manifestations of RA present, they require further special investigations such as chest x-rays and or other advanced imaging for respiratory manifestations like pneumonia and pleurisy (Wilczynska, Condliffe and McKeon 2013). The clinical diagnosis is determined by means of the diagnostic criterion developed and endorsed by the EULAR and the ACR as depicted in Table 2.2.

Table 2.2 Diagnostic criteria (EULAR/ACR RA criteria 2010)

| Criterion | Score |
|---|---|
| **Joints affected:** | |
| 1 large joint | 0 |
| 2 -10 large joints | 1 |
| 1 – 3 small joints | 2 |
| 4 – 10 small joints | 5 |
| **Serology:** | |
| Negative RF and ACPA | 0 |
| Low positive RF or ACPA | 2 |
| High positive RF or ACPA | 3 |
| **Duration of symptoms:** | |
| < 6 weeks | 0 |
| > 6 weeks | 1 |
| **Acute phase reactants:** | |
| Normal CRP or ESR | 0 |
| Abnormal CRP or ESR | 1 |
| Patients with a score of or more than 6 are considered to have definite RA. | |
| This diagnostic criterion is directly taken from the EULAR and ACR (Walker *et al.* 2013). | |

These criteria have been updated from the 1987 ACR (Mochan and Ebell 2008; Wasserman 2011) to the 2010 criteria in order to meet the most accurate available evidence and thus to ensure the most accurate diagnostic sensitivity possible for all serological types (Sakellariou *et al.* 2013; Zeidler 2013; Corrao *et al.* 2015; Myasoedova *et al.* 2020). This change in criteria has enabled the diagnostic tool to recognise RA at an earlier stage, with the resulting that earlier stages of RA can be treated more promptly, which decreases disease progression and increases the prognostic outcome for the patient (Myasoedova *et al.* 2020).

## 2.7 Treatment

This section provides an overview of the orthodox therapeutic efforts, among them pharmacological, surgical, and orthodox adjunctive interventions. This is followed by an overview of CAM, the prevalence of CAM use, perceptions of CAM, the gap in knowledge in this field and the potential use of manual therapies as an adjunct for RA treatment. Ultimately, the goal of RA intervention is to achieve remission, but if this is not possible, then at least to delay joint damage, reduce chronic pain and enhance the RA patient's QoL (Han *et al.* 2015). Despite the use of available therapeutic options and efforts to develop new and more effective therapeutic approaches (Matsushita, Yamaji and Tamura 2020), it has been stated that RA patients often do not obtain RA remission, improved health or pain relief (Han *et al.* 2015). Additionally, the pharmacological interventions have a profile of significant side effects, which impact on the patients QoL. Guidelines derived from the report by Combe *et al.* (2007) highlight the use of both non-pharmacological and pharmacological measures in the treatment of RA.

### 2.7.1 Orthodox therapy

For orthodox medicine to maximise the therapeutic efficacy of the treatment options offered to the RA patient, health policy and services need to adapt in order to meet the suggested global strategy laid out by the WHO (Safiri *et al.* 2019). This strategy aims at the development and maintenance of patients' functional ability in order to maintain their QoL and to depress disease activity. To achieve this access to specialized health care, early diagnosis, prompt initiation of therapeutic intervention and the patient's adherence to regime (based on treat-to-target principles) are essential in order to prevent the joint damage progression seen in a majority of patients and to increase the chance of remission (Safiri *et al.* 2019).

#### 2.7.1.1 Pharmacological therapy

As RA represents a major challenge to global health, the public increasingly needs to be educated about RA and to be made aware of risk factors as well as the importance that early diagnosis and intervention has for prognosis. Thus, the main goals of treatment are to prevent or control joint damage and loss of function whilst reducing pain (Christie *et al.* 2007) with the ultimate goal of remission. Furthermore, early diagnosis can prove cost-effective as the use of low-cost pharmaceuticals like MTX in early intervention may increase the efficacy of the treatment (Safiri *et al.* 2019).

In view of the fact that inflammation is the cornerstone of RA pathophysiology, controlling and treating it is a major therapeutic aim (Smolen, Aletaha and McInnes 2016). As stated in Chapter 1, the orthodox medicinal interventions include a combination of one or more of the following: NSAIDs, DMARDs and Steroids (Emery 2006). NSAIDs focus on pain control and stiffness, to improve the physical function of the patient. Unfortunately, they do not aid in the modification of

the disease, nor do they protect against structural damage. Similarly, steroids (such as glucocorticoids) are primarily used for rapid control of symptoms but have a role in disease modification (Smolen, Aletaha and McInnes 2016). DMARDs, on the other hand, target inflammation and attempt to control the structural damage synonymous with RA. Thus, the three main groups of DMARDs that are usually prescribed for RA act primarily by targeting the immune system, either through control of inflammatory mediators such as cytokines or non-specific or specific immune suppression. DMARDs are the principle pharmaceutical intervention for RA treatment as per recommendation by the ACR and the EULAR. Conventional or traditional synthetic DMARDs are considered the gold standard or first line of treatment for RA. These operate as potent suppressors of the immune response, controlling inflammation and disease activity in terms of joint destruction, clinical parameters and QoL by blocking the synthesis of immune active proteins, interfering with reactions critical to inflammatory cascades and in some instances acting as an antibacterial and directly or indirectly influencing the microbiome of the digestive system, thus affecting RA (see Section 2.4.5.10). Examples of this category of DMARDs include MTX, HCQ and SSZ (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019).

- MTX is the most widely used and preferred first-line oral pharmaceutical agent in the treatment of RA (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019). It is a drug derived from aminopterin, which is a compound capable of inhibiting proliferation of connective tissue. MTX has been shown to inhibit the immune system's capacity to respond to autoantigens through multiple pathways. Evidence shows MTX has the capacity to inhibit proliferation of T-cells (to obstruct the synthesis of purines and pyrimidines), to enhance the apoptosis of T-cells, to reduce cellular adhesion molecules (impeding signal transduction and activating an immune response), and to decrease the synthesis of cytokines (IL-1,4,6 and 10, TNF-α, IFN-γ and the granulocyte-monocyte-colony-stimulating factor (GM-CSF)) (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019).
- SSZ is an oral pro-drug that requires conversion activation by anaerobic bacteria that form part of the microbiome. Once converted to active form, the pharma-therapeutic action of this drug includes an anti-inflammatory role (reducing pain and inflammation), scavenging the synthesised ROS, enhancing the synthesis of prostaglandins, influencing cytokine synthesis by leukocytes, and inhibiting cellular proliferation (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019).
- HCQ is used in the treatment of RA for its anti-inflammatory action. This is induced by its interference with auto-antigen activity, together with its decreasing of CD4+ proliferation and the synthesis of cytokines; additionally it hindering TLR signal cascades and inhibits the synthesis of IL-6, 17 and 22 (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019). These various effects reduce both pain and joint swelling and may lessen the long-term disability that an RA patient is at risk of incurring.

Biological DMARDs such as TNF inhibitors, Anti-B-cell, Anti-T-cell co-stimulation, and Anti-IL 6R have four modes of action that represent the four subcategories of biologics. They are a viable alternative therapy for patients who show intolerance to the collective traditional DMARDs. When other biologics fail or are contraindicated, the drugs of choice in this situation are tocilizumab (an IL-6 inhibitor biologic) and rituximab (an anti-B-cell biologic) (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019). The four subcategories are as follows:

- The TNF inhibitors are anti-TNF antibodies such as infliximab and adalimumab which neutralize both extracellular and membrane forms of TNF, thus preventing upregulation of TNF-α in the synovium of RA-involved joints. Significant side effects include a risk of infection and malignancy complications (Bongartz *et al.* 2006; Yang *et al.* 2017).
- An anti-B-cell biologic such as rituximab is a monoclonal antibody that acts on CD20 receptors on the B-cell, resulting in B-cell induced cell death and thus immunosuppression (Smolen, Aletaha and McInnes 2016).
- Anti-T-cell co-stimulation biologics such as abatacept inhibit the antigen-presenting cell's capacity for co-stimulating the T-cell and thus its activation. Without activation, the T-cell cannot recognize self-antigens that would induce an upregulated response from the T-cell. Another pharmaceutical effect of the drug is that of inhibiting the function of the myeloid cells (Smolen, Aletaha and McInnes 2016; Lorenzetti *et al.* 2019).
- IL-6 inhibitors (e.g., tocilizumab) are monoclonal antibodies that block the CD68 receptor and IL-6 expression, thus collagen (type II) concentration and osteocalcin are downregulated (Smolen, Aletaha and McInnes 2016; Yang *et al.* 2017).

Research suggests that monotherapy with biologics offers no better clinical or functional improvement than MTX, but structurally appears to offer improved protection (Smolen, Aletaha and McInnes 2016). Therefore, combination therapy of any one of these biologics with MTX offers clinically significant enhanced therapy (Smolen, Aletaha and McInnes 2016). Targeted synthetic DMARDs (e.g., Janus kinase (JAK) inhibitors) such as tofacitinib achieve their clinical action by disrupting the signalling transduction associated with the stimulation by IL-6, 2 and 15, GM-CSF, and INF (i.e., a pan-JAK inhibitor). Clinical data suggests that tofacitinib has reported clinical outcomes superior to those achieved with MTX (Smolen *et al.* 2007; Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019).

The combination of tofacitinib with MTX has been reported to provide equivocal outcomes to biologics alone. Additionally, according to the ACR70, RA patients who failed to respond to TNF inhibitors have a response of roughly 15% to a combination therapy of JAK inhibitor therapy with MTX (Smolen, Aletaha and McInnes 2016; Bodkhe, Balakrishnan and Taneja 2019). The use of combinations of medication has led to the concept of polypharmacy. Generally, the extent of polypharmacy is proportionate to reduced response to and / or compliance with medication, which is of particular concern when adherence to orthodox RA therapy is already low because of the associated side effects. Many DMARDs are perceived to cause gastrointestinal side effects,

which results in patients choosing CAMs in preference to DMARDs, or misattributing side effects to DMARDs, and therefore stopping them (Zhao *et al.* 2017; Heidari *et al.* 2019). Additionally, the extra-articular manifestations of RA have an impact on the therapeutic regimen of choice and thus also on the clinical outcomes and on the effect on the ADL, QoL, and the prognosis associated with RA. An example of this is the use of MTX and / or JAK inhibitors for patients with extra-articular manifestations; however, these drugs increase the risk of infection in these patients (Matsushita, Yamaji and Tamura 2020). There is also a well-established relationship between adverse drug reactions and increasing age, which has been suggested to be a marker for polypharmacy and comorbidities (Zhao *et al.* 2017).

Given the complexity of pharmacological management of RA, three broad areas of limitations are noted in the literature:

- Treatment goals are based on the Clinical Disease Activity Index, Disease Activity Score in 28 joints (DAS28), Patient Activity Scale II, Routine Assessment of Patient Index Data 3 and Simplified Disease Activity Index (England *et al.* 2019). The efficacy of the intervention may also be evaluated by the ACR response criteria (ACR20/50/70). These guidelines offer a means to report and monitor treatment effectiveness with disease suppression and the ultimate goal of disease remission. Many RA patients never achieve remission, while other patients who achieve remission have to continue therapy to maintain it. This implies the need for new therapeutic agents (Smolen, Aletaha and McInnes 2016). Additionally, if treatment goals are not met by the sixth month of therapy, the protocol needs to be reviewed, exposing the patient to a cocktail of pharmacological interventions over time. This flexibility in providing pharmacological interventions is limited by balancing the choice of treatment protocol against many patient health factors and the risks associated with the agents used (Smolen, Aletaha and McInnes 2016; Heidari *et al.* 2019).
- Co-morbidities, or the presence of either related pathologies or unrelated health care ailments influence the pharmacological treatment options available to the RA patient. For example, rituximab is the drug of choice when the RA patient has co-existing MS, because of its efficacy in treating RA and MS, while the use of TNF inhibitors may elicit flares of MS (Smolen, Aletaha and McInnes 2016).
- Reduced drug efficacy, effectiveness and the lack of patient compliance suggest that patients may with time become less responsive to the treatment regimen implemented, either through decreased drug efficacy / immunogenicity or lack of patient compliance. The resultant loss of pharmaceutical efficacy requires regular evaluation of disease activity (Smolen, Aletaha and McInnes 2016) and regime adaptation. By contrast, a lack of patient compliance requires improved education on side effect profiles and the toxicity of the agents (Breedveld and Kalden 2004) and how these are best dealt with. Chandrashekara (2011) summarises orthodox medicine treatment strategies in treating RA as limited to a

below 70% variable response rate, with 30% or more of those treated according to this regimen suffering adverse effects.

Notwithstanding the above limitations, RA patients are still required to consume significant amounts of medication to control the progression of RA. These carry with them notable side effects, both positive and negative, and possibilities of toxicity (Breedveld and Kalden 2004). Additionally, as stated in Chapter 1, despite the use of orthodox medical therapy RA progresses over the patient's life, resulting in increased morbidity, a significant decrease in the patient's QoL, forcing the patient to adapt daily activities accordingly, and causing premature mortality. In addition, research shows an additional and significant burden on patients, with the annual costs (medical and disability related) escalating and correlating closely with degree of disability caused by RA (Breedveld and Kalden 2004). These factors are the basis for many RA patients seeking care outside of the recommended polypharmacy options that are available to them (Heidari *et al.* 2019).

#### 2.7.1.2 Surgical intervention and orthodox adjunctive therapy

In addition to surgical intervention, orthodox adjunctive therapy includes but is not limited to physical therapy, training, and weight loss as tools to aid symptom management. Unfortunately, these methods are effective in some cases, but are unable to completely control the symptoms or prevent the condition from progressing (Tokem *et al.* 2014). Surgical intervention is generally a last resort when medical interventions fail to achieve sufficient management of the disease and RA progression manifests in structural changes. The focus is on stabilizing lax joints, correcting deformities, and structuring pain management to facilitate improved QoL and ADL. Among these interventions, tendon sheath synovectomy may facilitate pain relief and prevent tendon rupture. When joint damage results in deformities requiring surgical correction osteotomy, arthrodesis and arthroplasty are among the techniques used to intervene (Walker *et al.* 2013).

### 2.7.2 Complementary and Alternative Medicine

#### 2.7.2.1 Definition of CAM

CAM has been defined by the WHO as “a broad set of healthcare practices that are not part of the country’s own tradition and are not integrated into the dominant healthcare system” (Macfarlane *et al.* 2012; Phang *et al.* 2018). CAM is thus a system of diverse medical practices and products excluded from what is considered orthodox medicine, in that these are not part of standard medical school curriculum nor are they taught to be used in the standard care of patients (Fernandez-Llanio Comella, Fernandez Matilla and Castellano Cuesta 2016; Zhao *et al.* 2017). Researchers view CAM as an alternative medicine often used in place of orthodox medicine, as opposed to alongside it (Fernandez-Llanio Comella, Fernandez Matilla and Castellano Cuesta 2016). In this study CAM represents a broad spectrum of healing practices and products that reflect all health care systems, modalities and philosophies of practice outside of what is taught and practiced by orthodox western medical schools and practitioners (Bhalerao *et al.* 2013; Fernandez-Llanio Comella, Fernandez Matilla and Castellano Cuesta 2016; Zhao *et al.* 2017). For the purposes of overview, it is possible to classify CAM therapies in five groups (Michalsen 2013; Chen and Michalsen 2017; Phang *et al.* 2018):

- **Biology-based therapies**
  Diet, natural products supplements, and herb therapy.
- **Manipulative and body-based practices**
  Massage, manual therapy, chiropractic, osteopathy, reflexology, and movement therapies, such as the Feldenkrais method or the Alexander technique, Tai chi, and dance therapy.
- **Mind-body interventions**
  Meditation, yoga, relaxation therapy, breathing techniques, biofeedback, and hypnosis, cognitive-behavioural therapy, and music therapy.
- **Energy therapies**
  Therapeutic touch, prayer, reiki, and qigong.
- **Whole medical systems**
  Traditional Chinese medicine (herbs, acupuncture, moxibustion, etc.), traditional Indian medicine (Ayurveda, Siddha) as well as anthroposophical medicine, naturopathy, homeopathy, European traditional medicine, Korean traditional medicine, Arabian traditional medicine, and Kampo traditional medicine.

A renewed approach is needed towards CAM as a growing movement in the world of medical science, underpinned by patient demand for CAM interventions (Bhalerao *et al.* 2013), improved access to information, the increased need for informed consent (McQuoid-Mason and Dada 2011; Allied Health Professions of South Africa 2015; Boucher, Brousseau and Chahine 2016), and the demand for the use of safe interventions that have acceptable clinical outcomes (Ryan *et al.* 2014). This requires integrating CAM into the orthodox medical system, shifting the paradigm of exclusion and segregation into a paradigm of integrative medicine (Chen and Michalsen 2017).

An example of this integration is the management of chronic pain using a multidimensional, multidisciplinary, and integrated approach (Chen and Michalsen 2017).

### 2.7.2.2 Prevalence of the use of CAM and patient seeking behaviour

According to Bhalerao *et al.* (2013), an estimated two-thirds of the world's population seeks alternative health care options. Cross *et al.* (2014) ascribed a 0.24% global burden to RA, while Ramos-Remus, Gutierrez-Ureña and Davis (1999) reported a minimum of 18% of rheumatic disease patients utilizing CAM therapy, indicating that about 3,4 million RA patients utilise CAM as an adjunctive or to replace orthodox medicine. Some RA patients choose to self-medicate while a large majority seek care from qualified practitioners, which results in rheumatological conditions being among the most frequent conditions treated by CAM practitioners (Phang *et al.* 2018; Heidari *et al.* 2019). In several studies, estimates of CAM use for rheumatic conditions have been reported (Table 2.3), but no differentiation was made as to whether this was a principal intervention or a co-intervention with orthodox polypharmacy (Heidari *et al.* 2019).

Table 2.3: Epidemiology of CAM use

| Country | Author (Year) | Prevalence of CAM use |
|---|---|---|
| Australia | Jadhav *et al.* (2011) | 40 – 82% prevalence |
| Canada | Jadhav *et al.* (2011) | 60 – 91% prevalence |
| England | Zhao *et al.* (2017) | 38% prevalence (arthritis in general) |
| Germany | Jadhav *et al.* (2011) | 78% prevalence |
| India | Jadhav *et al.* (2011) | 43 – 72% prevalence |
| | Bhalerao *et al.* (2013) | 42.7% (38.25 – 47.21 %) prevalence in RA patients |
| Israel | Breuer *et al.* (2006) | 42% prevalence |
| | Shmueli and Shuval (2004) | 15% of CAM users (for joint disease)<br>29% of CAM users (for back pain) |
| Japan | Kajiyama *et al.* (2006) | 35% prevalence |
| Mexico | Jadhav *et al.* (2011) | 56 – 83% prevalence |
| Turkey | Tokem *et al.* (2014) | 46.2 – 76.0% prevalence of CAM use in pts with arthritis |
| USA | Jadhav *et al.* (2011) | 18 -94% prevalence |
| | Eisenberg *et al.* (1998) | 42% prevalence |
| | Efthimiou and Kukar (2010) | 38% estimated prevalence of CAM us by RA patients |
| | Ernst (2004) | 33% prevalence of all RA patients in preceding year |
| | Tokem *et al.* (2014) | 28 – 90 % prevalence |
| General | Jadhav *et al.* (2011) | 60 – 90% of persons with arthritis |
| | Tokem *et al.* (2014) | 16 – 60% of persons with rheumatic disease |
| | Breuer *et al.* (2006); Alvarez-Nemegyei, Bautista-Botello and Dávila-Velázquez (2009) | 18 – 94% of persons with rheumatic disease |

As seen from Table 2.3, surveys of patients with rheumatic diseases show a wide range of prevalence of CAM use, from 18–94% (Ramos-Remus, Gutierrez-Ureña and Davis 1999). This demonstrates a steady worldwide increase in the use of CAM for the management of symptoms such as pain and also for the effects CAM has on RA and patients' QoL (Macfarlane *et al.* 2012; Bhalerao *et al.* 2013; Fernandez-Llanio Comella, Fernandez Matilla and Castellano Cuesta 2016; Chen and Michalsen 2017). The use of CAMs is driven by younger, active female patients who are more highly educated (Alaaeddine *et al.* 2012; Bhalerao *et al.* 2013). According to Zhao *et al.* (2017), patients report seeking CAM therapies because they feel that CAM:

- offers additional pain management (Rao *et al.* 1998; Lee *et al.* 2008; Jadhav *et al.* 2011; Michalsen 2013; Han *et al.* 2015),
- addresses costs. The cost of polypharmacy therapy may be excessive and thus accessibility becomes a concern, allowing CAM to provide potentially cheaper pain control alternatives (Bhalerao *et al.* 2013; Zhao *et al.* 2017) via the cost-effective, convenient options offered by CAM therapies (Jadhav *et al.* 2011; Bhalerao *et al.* 2013),
- applies a holistic and individualized approach (Jadhav *et al.* 2011; Bhalerao *et al.* 2013; Michalsen 2013). The bio-psycho-social approach allows the patient to explore with the practitioner the dynamic multifactorial character of their RA and gives them a role and responsibility in the treatment. In doing so, patients achieve a sense of control over their condition and the treatment going forward (Han *et al.* 2015),
- addresses issues in the relationship between patients and those orthodox medical practitioners (general practitioner, rheumatology specialists and nursing staff) who appear to lack confidence in and to feel dissatisfaction with orthodox healthcare and treatment options (Tokem *et al.* 2014). In addition, a perceived lack of efficacy, a large side-effect profile, the occasional serious adverse events and the associated substantial morbidity and mortality of the polypharmacy approach have been reported (Bhalerao *et al.* 2013; Michalsen 2013; Han *et al.* 2015; Chen and Michalsen 2017; Zhao *et al.* 2017),
- is perceived to be effective and safe, with the additional benefit that it addresses improved QoL for RA patients (Jadhav *et al.* 2011; Bhalerao *et al.* 2013). When its safety was further explored, most patients using CAM therapies felt that they were safe, and that the therapies did not cause side effects (Jadhav *et al.* 2011). Alaaeddine *et al.* (2012) report that 63.8% of patients in their study stated that CAM was beneficial, and they saw a reduction in intensity of pain, as well as improved sleeping patterns and levels of activity. Nevertheless, this does not imply that CAM therapies are not without side effects and adverse effects (Bhalerao *et al.* 2013). Some studies have stated that patients are under the misperception that CAM is natural, safe, and “soft”, and therefore must be healthier with fewer side effects (Alaaeddine *et al.* 2012; Macfarlane *et al.* 2012; Michalsen 2013; Zhao *et al.* 2017). However, Tokem *et al.* (2014) state firmly “no CAM modalities are harmless, and it is important to select the appropriate CAM”.

The above discussion supports the outcomes reported by Bhalerao *et al.* (2013) reported, where 27% of their sample took part in two or more CAMs. Furthermore, according to the study conducted by Tokem *et al.* (2014), almost half of their sample had used CAMs. Additionally, Jadhav *et al.* (2011) claimed that of the CAM users in their study, 57% underwent regular therapy, while the remaining 43% only used CAM on an irregular basis. This use of CAM is based on the following sources of CAM therapy information:

- advice from family, friends, and acquaintances (Tokem *et al.* 2014),
- internet, social media, and mass media (Alaaeddine *et al.* 2012; Tokem *et al.* 2014; Zhao *et al.* 2017),
- health care providers – It is regarded as the duty of the health care providers to offer the patients sufficient and reliable information on the pros and cons of all treatment methods.

Regardless of the limited scientific backing and orthodox medical approval of CAM, patients appear willing to try it (Alaaeddine *et al.* 2012; Zhao *et al.* 2017). Patients approach orthodox medical practitioners for CAM therapy information (Tokem *et al.* 2014), thus Tokem *et al.* (2014) explain that the orthodox medical professionals are under obligation to provide sufficient, objective, and reliable information on the pros and cons of all therapy options, both orthodox and alternative. With specific regard to the rheumatic patient and the global trends for CAM use, the most commonly used CAMs for the pain related to rheumatic conditions include thermotherapy, acupuncture / acupressure, massage and spinal manipulation (chiropractic, osteopathic and chuna (tuina)), together with herbal supplements from the disciplines of naturopathy and homeopathy (Jadhav *et al.* 2011). These choices are driven by consumer demand, geographic location, and cultural influencers (Jadhav *et al.* 2011). The patient demand in the sphere of rheumatic conditions, which are difficult to treat and manage, requires that both CAM practitioners and orthodox medical practitioners need to better understand the safety, efficacy, and relative effectiveness profiles of different CAMs in order to provide their patients with objective and clear information (Ernst and Posadzki 2011). This is important as the literature concedes that the level of implied effectiveness affirmed by both the CAM practitioner and patient indicates that the CAMs contribute to addressing patient concerns around RA (Chandrashekara 2011; Chen and Michalsen 2017). Therefore, even with a significant research effort over the last 30 years (Macfarlane *et al.* 2012), there remains a need for further scientific evaluation that targets prioritized CAM research areas (e.g., physical therapy, psychotherapy, balneology and rehabilitation). Literature does, however, communicate that while evidence builds, the lack of evidence does not imply absence of efficacy; it rather implies that objective caution should be practiced (Michalsen 2013).

#### 2.7.2.3 Perceptions of clinical RA care and their implications

The need for objective caution in clinical practice is driven by several clinical and regulatory imperatives that strongly suggest that health care practitioners treating RA make purposeful enquiry into the use of CAMs (Bhalerao *et al.* 2013). According to Bhalerao *et al.* (2013) these imperatives include:

- the need for informed consent (McQuoid-Mason and Dada 2011),
- the right of patients to discuss care options and seek second opinions – Patient Right's Charter (Department of Health of South Africa 2020),

- the freedom for patients to be able to disclose to their medical health practitioner that they are utilising CAMs, whether in conjunction with conventional prescribed therapy or in place of conventional therapy (Zhao *et al.* 2017). Bhalerao *et al.* (2013) report that 38.5% of patients disclosed CAM use, while Jadhav *et al.* (2011) reported that 29% of patients did. This may be related to the facts that 70-84% of health care practitioners were not enquiring about the use of CAMs (Bhalerao *et al.* 2013) and / or 75% of health care practitioners were expressing disapproval of the use of CAMs in some jurisdictions (Jadhav *et al.* 2011; Alaaeddine *et al.* 2012). By contrast, 60% of American and Israeli primary care physicians made referrals to CAM providers at least once in the preceding year (Borkan *et al.* 1994; Breuer *et al.* 2006).
- the need for improved safety in terms of drug interactions, particularly herbal or medication-based CAMs, especially with regard to elderly RA patients because of possible altered pharmaco-kinetics and -dynamics, comorbidities, and polypharmacy, all of which influence potential drug toxicity and compliance to the prescribed treatment (Zhao *et al.* 2017).

With increased patient awareness of their right to access information regarding CAM (Chen and Michalsen 2017) and the need for objective and clear information around CAM therapy, research is necessary to address the patients' predicament to the greatest extent possible (Tokem *et al.* 2014; Zhao *et al.* 2017).

## 2.8 Manipulation, mobilisation, massage, and exercise

According to a study by Manek *et al.* (2010), rheumatologists practising in the United States who were recruited into the study provided responses to the perceived benefit of each CAM therapy as shown in the figure extracted from this study (see Figures 2.1 and 2.2). From both figures it can be seen that the highest perceived benefit was derived from massage with 70**%** of respondents indicating that these practices were either "very beneficial" or "moderately beneficial". This was followed by meditation (63%), acupuncture (54%) and spinal manipulation (52%), all considered to be either "very" or "moderately" beneficial. Additionally, it was noted that rheumatologists in institutional practice settings were more likely to perceive the benefit of spinal manipulation and bodywork practices than their counterparts in other practice settings (Manek *et al.* 2010). This suggests that there are a growing number of studies being conducted on bodywork practices (massage, manual therapy, chiropractic, osteopathy, reflexology, and movement therapies, Tai chi, and dance therapy) for the purpose of evaluating their capacity to treat rheumatic conditions (Phang *et al.* 2018).

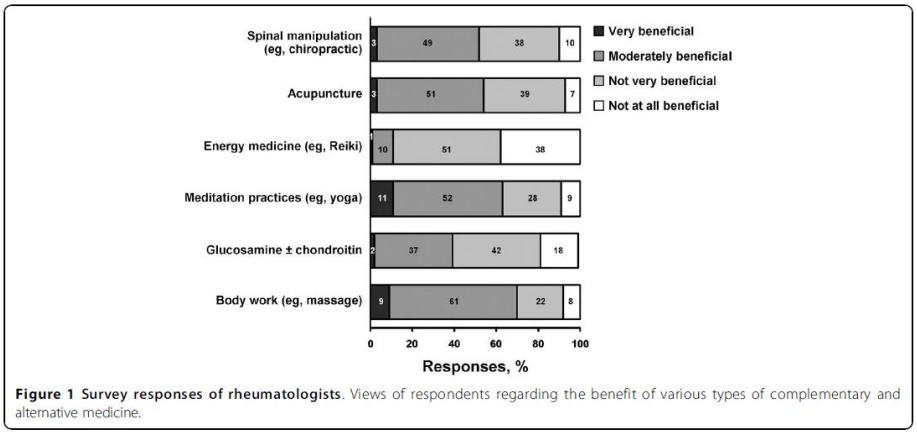



Figure 2.1: Extracted from Manek *et al.* (2010)

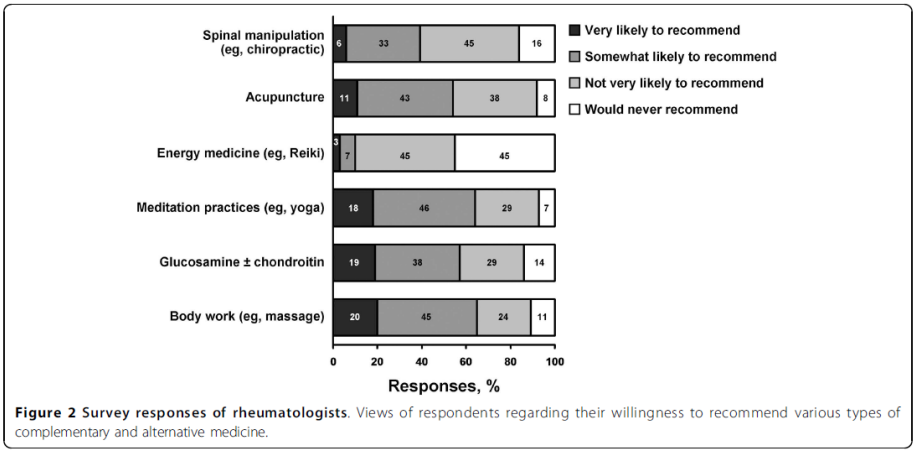



Figure 2.2: Extracted from Manek *et al.* (2010)

### 2.8.1 Manipulation and Mobilisation

Manual therapy in the form of joint mobilization or manipulation (graded as I through V) (Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011) is a modality used by several disciplines, among them chiropractic, physiotherapy and osteopathy. It has typically been employed to restore mobility in joints and alleviate pain in patients with musculoskeletal disorders (Levitsky *et al.* 2019). These disorders include joint dysfunction in the form of hyper- or hypomobility that stems from either an inflammatory or a non-inflammatory (mechanical or neurological) process (Mooney and Robertson 1976; Leach 2004; Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011). Thus, an ankle sprain results from a mechanical disruption of the ankle joint complex (Coetzer 1999), while, by contrast, systemic conditions such as gout, RA and psoriatic arthritis change the normal joint's internal environment through processes of crystalline deposits, excess pannus formation or inflammatory processes (Walker *et al.* 2013; Longmore *et al.* 2014). The manual therapist is therefore required to assess the joints for the causative agent prior to determining the intervention type, for example, manipulation and

mobilisation. This is because manipulation and mobilization are principally directed at the hypomobile joints secondary to mechanical joint locking, adhesion formation, meniscoid entrapment, muscle spasm/splinting and / or irritation of the nervous system that results in muscle spasm, causing the joint/s to lock (Mooney and Robertson 1976; Leach 2004; Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011). Additionally, hypomobile joints may result from regions of hypermobility and / or alterations in the kinematic chain (Chung and Mior 2015). Manipulation addresses these hypomobile joints by imparting a high-velocity low-amplitude thrust (HVLA thrust) into the joint at the point of restriction (Sandoz 1976; Vernon and Mrozek 2005; Bergmann and Peterson 2010). The HVLA thrust accomplishes the following:

- it separates the joint surfaces, thus reducing mechanical joint locking (Cramer *et al.* 2000; Bergmann and Peterson 2010),
- it breaks intra-articular adhesions in order to restore normal joint movement (Vernon and Mrozek 2005),
- it allows for entrapped meniscoids to be freed (Leach 2004),
- it reduces muscle spasm through reflex neurological control mechanisms of muscle function (Pickar 1999, 2002; Leach 2004; Grindstaff *et al.* 2009; Botelho and Andrade 2012),
- it reduces pain by mechanoreceptive stimulation (Melzack and Wall 1965) which results in endorphin release (Vernon *et al.* 1986; Chung and Mior 2015) as well as addressing muscle splinting (Pickar 2002; Leach 2004; DeVocht, Pickar and Wilder 2005; Lehman 2012), reducing oedema (Coetzer 1999; Sambajon *et al.* 2003), restoring normal movement and thus normal mechanoreceptive activity (Melzack and Wall 1965) and causing an attendant serotonin release that comes from the patient's participation in normal PA (Chung and Mior 2015),
- and, additionally, it reduces pain by stimulating activating segmental inhibitory pathways or descending pain inhibitory systems (Haldeman 1984; Wall, McMahon and Koltzenburg 2006).

Similarly, mobilizations involve positioning a joint at its restrictive barrier (i.e., the end of passive ROM) after which the practitioner uses a series of gentle repetitive movements towards and through the restriction in order to increase the ROM available at the joint (Atchison 2000). This is often done to treat painful joints and typically involves the application of a rhythmic oscillatory motion of the joint surfaces within the normal joint range (Lehman 2012). Studies have suggested that mobilization reduces pain and improves function while increasing the ROM (Pickar 2002). Thus, passive mobilizations (grade I to IV) are generally utilised to maintain or restore ROM (Haldeman 2005). In the context of RA, where there are swollen, stiff and painful joints (Orange *et al.* 2020) associated with increased nociceptive impulse activity (leading to peripheral and central sensitization) and increased sympathetic nervous system activity (Slosberg 1990; Konttinen *et al.* 1992; Schaible and Grubb 1993; Konttinen *et al.* 1994; Hogeweg *et al.* 1995), it

has been suggested that the following aspects of the RA pathophysiological processes are addressed by manipulation and mobilisation:

- managing the inflammatory component of RA (Droz-Georget 1980),
- addressing abnormal compensatory kinematics (Droz-Georget 1980),
- decreasing the RA-associated hyperalgesia / pain (Dhondt 1999; Meeus *et al.* 2012; Chung and Mior 2015),
- reducing muscle spasm (Chung and Mior 2015),
- reducing adhesions through the chronic RA inflammation and tissue damage (Weinberg, Pippen and Greenberg 1991),
- ensuring improved circulation (Zegarra-Parodi *et al.* 2015),
- and collectively assisting in addressing patient distress and functional impairment by improving strength and self-efficacy (Jadhav *et al.* 2011; Chung and Mior 2015; Chen and Michalsen 2017).

The above effects would need to be differentially achieved depending on the stage of the RA and the presence of flare-ups or variances in the presentation of the RA. Thus, manipulation and mobilisation need to individually match to the patient need, especially in RA patients who might have variable presentation of synovial tissue proliferation, persistent inflammation, cartilage degradation, bone erosion, and damage to the adjacent soft tissue and neural structures (Chung and Mior 2015; Romanowski *et al.* 2020). This is further complicated by the timing of delivery of therapeutic interventions because of the unpredictable pattern of changing symptoms in rheumatologic patients, can differ from day to day (Home and Carr 2009). Thus, clinicians are faced with the challenge of sifting through the many grey areas of RA presentation and the evidence associated with manipulation and mobilisation as well as the patient's needs and requests (Sackett *et al.* 2007; Kelly 2009). In contrast to the potential benefits outlined above and patients' increasing use of CAM and in particular manipulation and mobilisation (Jadhav *et al.* 2011; Alaaeddine *et al.* 2012; Zhao *et al.* 2017), the WHO advises that manipulation and mobilisation of RA patients is absolutely contraindicated in regions of the skeletal structure that have active RA involvement (Sweaney 2005). This advisory was based on clinical findings of instability at the atlanto-axial complex in RA patients. Yet, despite the rare occurrence of this manifestation, and the limited anatomical location of the instability (Longmore *et al.* 2014), it forms the basis for a guideline that is holistic in its approach. Chung and Mior (2015) have proposed that the WHO guideline be more joint- or region- specific so as to indicate where the risk of manipulation or mobilisation outweighs the benefit to the patient. However, the recommendation by Chung and Mior (2015) requires evidence to support the development of this guideline, and there is a significant shortage of evidence contributing to current clinical guidelines regarding manipulation and mobilisation based on expert opinion rather than empirical evidence (Forestier *et al.* 2009). Additionally, evidence that does exist in the form of published peer-reviewed literature is contradictory and often recommendations are not contextualised in the dynamic characteristics

of RA that enable practitioners to identify which RA patients are more likely to benefit from manipulation and / or mobilisation (Dhondt 1999; Chung and Mior 2015; Romanowski *et al.* 2018; Levitsky *et al.* 2019).

There is, therefore, a need to review the studies available on manipulation and mobilisation in order to better understand the evidence supporting or not supporting of manipulation and mobilisation in the treatment of RA patients, but also to understand the reported benefit gained from manipulative and mobilization intervention for the various stages of RA. This would make it possible to identify the safety of the interventions and draw guidelines with a view to recommending research in areas that are as yet not fully addressed.

### 2.8.2 Massage:

In the same way as manipulation and mobilisation impart movement to joints, massage therapy imparts flexibility to soft tissue by manipulating this tissue in various ways. These include but are not limited to effleurage, petrissage, hacking, stroking and cross frictions, which are principally done by hand but may also include the use of a handheld device to achieve one or more of the following effects in muscle, fascia, and tendon / ligamentous structures:

- the breaking of muscular, fascial and / or tendon sheath adhesions in order to restore flexibility in these structures (Beck 2012),
- a reduction of muscular entrapment syndrome (Romanowski, Romanowska and Grzeskowiak 2012),
- a reduction of muscle spasm through the stimulation of the muscle spindle receptors and Golgi tendon by means of pressure, stretch and positioning which results in spinal reflex-induced muscle lengthening and relaxation (Beck 2012) and also reduces pain by mechanoreceptive stimulation (Melzack and Wall 1965). This not only results in endorphin release (Beck 2012) but also addresses muscle splinting (Fitzcharles 2002), reducing oedema and lymphatic obstruction (Beck 2012) and restoring normal movement and thus normal mechanoreceptive activity (Melzack and Wall 1965) with the attendant endorphin and enkephalin release that come from the patients' participation in normal PA (Chung and Mior 2015),
- a reduction of pain through activating segmental inhibitory pathways or descending pain inhibitory systems (Meeus *et al.* 2012). Beck (2012) furthers this construct by reporting the elevated endorphin and enkephalin levels in the CNS as a result of massage therapy while Vigotsky and Bruhns (2015) in their review looking at the role of descending modulation in manual therapy report that oxytocin has a role in the analgesic effects of massage.
- an increase in the feeling of well-being deriving from the touch therapy that massage imparts to the patient (Lindgren *et al.* 2012),
- indirect effects arising out of the use of particular oils, as in aromatherapy, or the specific placement of pressure, for example, on pressure points or reflexology points, which may

achieve more systemic effects in terms of pain control and relaxation with beneficial results in terms of patient stress and anxiety (Steflitsch and Steflitsch 2008; Yip and Tam 2008; Basaran 2009; Beck 2012).

Given that RA patients have reduced muscle strength, flexibility and the increased likelihood of contractures and adhesions (Weinberg, Pippen and Greenberg 1991), the use of massage could potentially address these considerations. Field, Diego and Hernandez-Reif (2007); Field *et al.* (2007) report that the use of massage techniques such as wringing and skin rolling, as well as circular and friction type strokes on a group of RA patients produced a significant improvement in grip strength as well as in mobility and function by comparison with a control group that did not receive massage therapy. Furthermore, other techniques such as effleurage may also benefit RA patients who have lymphatic blockage in lymph nodes proximal to affected joints (Osborn 2005). Additionally, the benefits of the touch and stimulation of mechanoreceptors assist in reducing the discomfort and pain experienced by RA patients. This reduction in pain and discomfort is further enhanced by circular friction techniques that have been shown to increase oxygen and nutrient delivery and assist in the removal of waste products in the soft tissue surrounding the affected joints (Osborn 2005). Collectively these effects would be expected to reduce pain (Field *et al.* 2004; Field, Diego and Hernandez-Reif 2007; Field *et al.* 2007; Harris and Richards 2010), and thus have a positive effect with regard to self-efficacy and reducing the psychological impacts of RA (stress, depression, anxiety) (Field *et al.* 2011; Beck 2012).

All the above massage techniques and their effectiveness may be augmented with the use of aromatherapy oils (Metin and Ozdemir 2016), the effects of which are mediated by physiological and psychological means (sense of smell or absorption through the skin) (Steflitsch and Steflitsch 2008). Aromatherapy is credited with having the following systemic effects, depending on the oil mix (antiseptic, antibacterial, analgesic, anti-inflammatory, anti-spasmolytic, antitoxic, immune stimulatory), as well as promoting relaxation and pain reduction (Steflitsch and Steflitsch 2008; Yip and Tam 2008; Basaran 2009). Evidence for aromatherapy massage for treating in juvenile RA suggests that it relieves pain, fatigue, morning stiffness, and anxiety (Field *et al.* 1997; Kim, Nam and Paik 2005; Field *et al.* 2007). In addition, an increase in physical and mental well-being have also been noted (Brownfield 1998; Ovayolu and Ovayolu 2013), principally on the strength of decreased pain scores for patients with RA (Kim, Nam and Paik 2005; Han *et al.* 2010). Massage applied on reflexology points showed significantly decreased pain scores after six one-hour reflexology sessions (Khan, Otter and Springett 2006) and significantly decreased fatigue after a weekly 45-minute reflexology intervention over six weeks (Khan, Otter and Springett 2006; Otter *et al.* 2010).

Nevertheless, even with the benefits noted above, there are limitations on massage application, and it has been suggested that massage is contraindicated during the acute inflammatory stage of RA (Beck 2012), but when RA is in remission, massage can effectively manage symptoms, prevent inflammation, and reduce joint damage (Lowe 2006).

### 2.8.3 Exercise:

Exercise comes in many forms but the benefits for specific patients lie in a properly designed physical exercise programme that improves cardiorespiratory fitness and cardiovascular health (Thompson *et al.* 2003; Lavie *et al.* 2009), increases muscle mass, reduces adiposity, improves strength and aids in improving physical functioning. Collectively these factors are protective of chronic lifestyle diseases related to inactivity (e.g., diabetes mellitus) and diseases of old age (e.g., osteoporosis). Moreover, they produce a feeling of well-being and independence in the patient, which contributes to assisting patients being active members of society (Cooney *et al.* 2011). Thus, in the context of RA, exercise seems to be designed to:

- manage the inflammation (James *et al.* 1994),
- increase aerobic capacity (Van den Ende *et al.* 1996; Van den Ende *et al.* 2000; Lemmey *et al.* 2009),
- increase muscle strength and reduce muscle loss (Van den Ende *et al.* 1996; Van den Ende *et al.* 2000; Lemmey *et al.* 2009; Matschke *et al.* 2010a; Matschke *et al.* 2010b),
- improve joint mobility (Van den Ende *et al.* 1996; Van den Ende *et al.* 2000; Lemmey *et al.* 2009) and the preservation of the joint cartilage (Van den Ende *et al.* 2000; de Jong *et al.* 2003; de Jong *et al.* 2004; de Jong and Vlieland 2005; de Jong *et al.* 2008; de Jong *et al.* 2009) and to decrease the number of joints involved in RA (Lyngberg et al. 1994),
- increase tendon mobility and flexibility (Buchanan and Marsh 2002; Reeves, Narici and Maganaris 2006),
- strengthen ligaments (Benjamin and Ralphs 1996, 1997),
- improve balance and proprioceptive ability, reducing risk of falls and other injuries that may occur otherwise (Peyron 1986; Sharma and Pai 1997; Hurley and Scott 1998; Van den Ende *et al.* 2000),
- improve the cardiovascular capacity compromised in RA (Van den Ende *et al.* 1996; Van den Ende *et al.* 2000; Lemmey *et al.* 2009; Metsios *et al.* 2009), thus decreasing the likelihood of a reduced life expectancy associated with RA (Cooney et al. 2011),
- decrease the rate of adiposity that usually accumulates with RA secondary to inactivity (Van den Ende *et al.* 1996; Van den Ende *et al.* 2000; Lemmey *et al.* 2009),
- improve functioning (especially of the hands) (Van den Ende *et al.* 1996; Van den Ende *et al.* 2000; Lemmey *et al.* 2009), improving the self-efficacy of the patient (de Jong *et al.* 2003; Hakkinen *et al.* 2005; Marcora, Lemmey and Maddison 2005; Lemmey *et al.* 2009),

- improve psychological wellbeing (Häkkinen 2004; Marcora, Lemmey and Maddison 2005; Neill, Belan and Ried 2006; Lemmey *et al.* 2009),
- decrease the effects of inactivity such as lethargy and fatigue (Roubenoff *et al.* 1994; Neill, Belan and Ried 2006; Lemmey *et al.* 2009; Metsios *et al.* 2009),
- and decrease the risk of co-morbid conditions (e.g., diabetes mellitus, osteoporosis) (Roubenoff *et al.* 1994; de Jong *et al.* 2004; Sinigaglia *et al.* 2006; Franck and Gottwalt 2009; Lemmey *et al.* 2009; Metsios *et al.* 2009).

Unlike mobilisation, manipulation and massage, the use of exercise to control RA is less subject to being contraindicated by reason of RA. Exercise may, however, be relatively contraindicated in the case of RA patients with conditions regarded as comorbid (e.g., CVD). This does not mean that exercises are completely contraindicated but rather that they require tailoring to the patient's ability at the time that they are started and must be modified as the patient's capacity to participate in exercise increases (Peyron 1986). In summary, the strength and endurance of the muscles together with tone and elasticity of the connective tissues promote optimal joint stability, alignment and attenuation of impact and compressive forces (Bland 1988; Minor 1991), helping the RA patient achieve a normal life. However, the application of exercise is limited as the intensity, frequency, exercise type, and goal of the exercise are for the most part undefined for the different stages of RA. Factors to be considered here are loads, load progression, volume of exercise, repetitions, frequency, intensity, sets, rest intervals, duration of programme, specificity of exercises and individualisation of the protocol to the participant (Heine *et al.* 2012). This is also seen in the differing outcomes in studies with differently staged RA patients and / or different exercise interventions (Buchanan and Marsh 2002; Cooney *et al.* 2011). This limits the ability to provide the patient with optimal exercise programmes that are linked to their stage of RA development and that will enable them to achieve optimal clinical outcomes. The situation therefore requires a review of the literature to indicate the status of the current evidence available for different types of exercise and the parameters related to exercise that would be beneficial to the RA patient. Such a review can also highlight areas that require future studies in order to allow practitioners to identify what is best for the patient in their specific clinical context.

A further limitation to the use of manipulation, mobilisation, massage and exercise for the treatment of RA is the fact that only certain regions of the spine have been addressed in the literature (Atchison 2000; Chung and Mior 2015). Additionally, the focus has been on the limitations that the above interventions present with regard to treating patients' wrists and hands to the exclusion of other regions affected by RA (Levitsky *et al.* 2019).

### 2.8.4 The implications for the CAM treatment options

Based on the limitations that are found with regard to each of the manual therapies outlined above (manipulation, mobilisation, massage and exercise), it is evident that although there is support for the use of the interventions based on their physiological and clinical effects, the actual clinical application has yet to be fully understood. This is further compounded by the contradictory published data (Christie *et al.* 2007; Furlan *et al.* 2012), which highlights the need for a review of what is currently known in order to be able to address the translation of this knowledge into clinical practice (Tierney 2001; Ryan *et al.* 2014; Munn *et al.* 2018a). This process will, however, also highlight areas of future research. These two steps as part of the outcomes of a review of the literature (in this study a systematic review of the literature) will address patient, practitioner, and guideline arenas in clinical practice in the following respects (Chandrashekara 2011):

- a review of the relevant literature will enable patients to give informed consent regarding their treatment as a result of understanding the implications of these manual therapies for their RA (Allied Health Professions of South Africa 2015; Boucher, Brousseau and Chahine 2016),
- a review of this nature will facilitate evidence-informed practice (Joseph 2004; Sackett *et al.* 2007), increasing the cultural authority and assisting medical professionals to employ evidence-informed practices as well as supporting the use of these practices in treating RA (Gambrill 2003; Sackett *et al.* 2007; Chandrashekara 2011),
- such a review will also enhance inter-professional practice interactions, with a positive impact on RA patient care (Makaram 1995),
- such a review will also make a positive contribution to evidence-informed practice guidelines (Tranfield, Denyer and Smart 2003; Sackett *et al.* 2007).

Additionally, the above will then be supportive of the growing CAM seeking behaviour of RA patients (Breedveld and Kalden 2004; Bhalerao *et al.* 2013; Phang *et al.* 2018) while also endorsing the efficacy and safety profiles of manipulation, mobilisation, massage and exercise in the treatment of these patients (Michalsen 2013).

For all the reasons cited above this systematic review aims to identify, evaluate and summarise the findings of all relevant studies pertaining to the efficacy of non-invasive manual therapy (manipulation, mobilisation, massage, and exercise) in the management of RA.

## 2.9 Systematic reviews

The value of a systematic review is found in its critical evaluation of available literature in order to provide a concise, accurate and reliable source of evidence in a given field of interest (Liberati *et al.* 2009), the goal being to provide a source from which the medical scientific community may create a risk / benefit analysis of the intervention to the patient when implementing such an intervention to treat the focus condition (Liberati *et al.* 2009). This is achieved by means of a clearly defined, rigorous methodology implemented in the clinical appraisal of the available evidence (Moher *et al.* 2009b).

The well-defined protocol of a systematic review is an important attribute of the study methodology and allows for a clear understanding among the researchers involved as to the process of each step in the study and what is required. This transparency and accountability provides the study protection from the risk bias introduced by the researchers (Moher *et al.* 2009b).

### 2.9.1 How is a systematic review implemented?

The systematic review aims to methodologically identify and retrieve available evidence to address the research question, which in this study is whether non-invasive manual therapy (manipulation, mobilization, massage, and exercise) is efficacious in the management of RA. The evidence is then critically reviewed in a manner that reduces bias, thus providing a succinct, reliable, and objective set of findings, that allowing both evidence-based and informed conclusions together with clinical decision-making around the management of RA (Pearson 2004; Higgins, Green and Collaboration 2011; Steinberg *et al*. 2011; Pearson, Jordan and Munn 2012). Furthermore, the systematic review provides insight into the quality of the evidence and the conflict between studies that supports or refutes their outcomes (Edoardo and Alan 2014; Munn *et al.* 2018a; Munn *et al.* 2018b). This provides a synthesised research base to inform practice and clinical guidelines, to guide policy and to direct further research efforts towards gaps in evidence (Pearson 2004; Liberati *et al.* 2009; Edoardo and Alan 2014; Munn *et al.* 2018a; Munn *et al.* 2018b). However, for the systematic review to be most effective, practitioners need to utilise systematic reviews, which also offer an effective and efficient summary of the literature in a methodical, critical, and organised manner (Centre for reviews and dissemination (CRD) guideline (Dissemination 2009)). The use of a transparent and rigorous approach as to the synthesis of the data in this study, as seen in Chapter 3, not only reduces the bias that is often present in narrative reviews but also increases the validity of the study for future reproducibility (Mallett *et al.* 2012). In keeping with systematic review practice, this systematic review follows the Cochrane handbook (CRD guideline – (Dissemination 2009)) which lays out the eight steps a systematic review should follow:

- develop a defined question for the review with inclusion criteria by which the available evidence is included,
- search the literature using keywords and both electronic and manual search engines,
- select appropriate evaluation tools and data collection,
- apply the evaluation tools (internal and external validity and associated level of bias),
- analyse the data retrieved from the study,
- identify and address the bias observed,
- summarise the findings of the individual studies and provide a description of their respective strengths and weaknesses and the results published,
- interpret the results of the review and draw conclusions**.**

In this study, the evaluation of evidence is achieved by means of an evidence-based, non-biased critical analysis of studies including RCTs, nRCTs, and CS/CR (Roopchand 2014). As displayed in Figure 2-3, extracted from the publication of Evans (2003), study designs fall into a hierarchical framework with systematic reviews and meta-analyses holding the highest level of clinical appraisal.

| | Effectiveness | Appropriateness | Feasibility |
|---|---|---|---|
| **Excellent** | • Systematic review<br>• Multi-centre studies | • Systematic review<br>• Multi-centre studies | • Systematic review<br>• Multi-centre studies |
| **Good** | • RCT<br>• Observational studies | • RCT<br>• Observational studies<br>• Interpretive studies | • RCT<br>• Observational studies<br>• Interpretive studies |
| **Fair** | • Uncontrolled trials with dramatic results<br>• Before and after studies<br>• Non-randomized controlled trials | • Descriptive studies<br>• Focus groups | • Descriptive studies<br>• Action research<br>• Before and after studies<br>• Focus groups |
| **Poor** | • Descriptive studies<br>• Case studies<br>• Expert opinion<br>• Studies of poor methodological quality | • Expert opinion<br>• Case studies<br>• Studies of poor methodological quality | • Expert opinion<br>• Case studies<br>• Studies of poor methodological quality |

Figure 2.3: Hierarchy of evidence: ranking of research evidence evaluating health care interventions.

This hierarchy provides a structure within which an appraisal of available evidence can be completed and guidelines for clinical practice can be developed (Blunt 2015). The framework thus provides an important role for the development of evidence-based medicine (Blunt 2015). It is, however, important that this framework is not seen as a ridged hierarchy and that study designs are chosen appropriate to the question/s that they can address, not based on the position they happen to occupy in the framework (Cook and Thigpen 2019). Additionally, clinical practice is best emulated in the lower sections of the framework and the further one departs from this point the greater the disjuncture between clinical practice and the outcomes of the trial. This limits the application of evidence in clinical practice (Blunt 2015).

Because systematic reviews serve a vital role in clinical decision-making and resource allocation, decision-makers should expect consistent and unbiased standards across topics covered in their review (Whiting *et al.* 2016). However, even though a systematic review is methodological and systematic in terms of the procedure and processes that it follows, like all other studies on the framework it is subject to the influence of bias. This bias can develop at any stage of the review and readers should keep this in mind when interpreting the findings of a systematic review (Whiting *et al.* 2016).

These biases include but may not be limited to:

- **Publication bias**

  Publication bias represents a concern as positive outcome studies tend to have preferential submission for publication (Callaham *et al.* 1998), and tend to appear in English journals more frequently (Egger *et al.* 1997), and the lack of access to grey literature (defined as research not formally peer-reviewed), whether such literature be unpublished or published, in non-electronic form or not accessible due to licencing concerns (Knobloch, Yoon and Vogt 2011), contributes to publication bias which then positively biases a systematic review, allowing the reviewers to favour outcomes reporting. This tends to result in "a systematic failure of interpretation and clinical decision-making" from the review findings (Knobloch, Yoon and Vogt 2011). Thus, publication bias creates possible difficulty with regard to identifying, detecting and accessing studies for the purposes of the systematic search.

  Given the risk of publication bias, the selection of articles for the systematic review needs to be meticulously considered and to follow a transparent and clearly stated method in order to offer the best protection against such bias, even though this may not fully eliminate bias (Knobloch, Yoon and Vogt 2011). The PRISMA-statement and checklist (Appendix K) were followed in this systematic review so as to offer the reader a clear understanding of this process (Moher *et al.* 2009b).

- **Language bias**

  The limiting of a systematic review in terms of language is a common practice, with English publications being predominantly included (Stern and Kleijnen 2020), which introduces language bias (the risk of excluding key studies and the data from non-English studies) as well as the loss of cultural context (the risk of losing the value and applicability of the review findings to varying populations) that could result from a translation in to English necessary in order to allow article inclusion (Jackson *et al.* 2019; Walpole 2019; Aromataris 2020). The language bias is further complicated by geographical location, economic and regional cultural influencers that may impact on levels of access to publications and also differential access or full access to publications (Stern and Kleijnen 2020). Exclusion of studies as a

result of translation practices or translation outcomes and the associated bias related to the translation processes of a study may also impact negatively on the interpretation of the study even if such studies were to be included, which potentially results in the emphasis on or diminishing of the outcomes, thus skewing the interpretation of the data available around the research question.

To avoid publication and language bias a systematic review should search multiple databases and multiple languages of publication in the process of attempting to identify publications (Aromataris 2020). It should be noted that Cochrane advocates that linguistic restrictions should not be implemented (Higgins *et al.* 2019). This is reasoned for by a statement made by Stern and Kleijnen (2020): “by limiting a review to one language from the outset, we are violating the very essence of what a systematic review is and its purpose in assisting in making informed decisions from the best available evidence”. While it is true that limitations faced by the research team are inevitably the factors that define the scope of the study, efforts must be made to continue to push for the inclusion of diverse languages in order to provide a true global evidence (Stern and Kleijnen 2020). Systematic reviews should outline the restriction of language with appropriate reasoning and also consider the impact of such a restriction on the overall outcome of the review and its application in clinical settings, decision-making, and generalisability to other cultures in the discussion (Stern and Kleijnen 2020). The above should also be reflected in the PRISMA (Item 6: Eligability criteria and Item 25: Limitations) (Moher *et al.* 2009b) in order to provide the reader with the context from which to read the review and make judgements in full knowledge of the associated bias (Stern and Kleijnen 2020).

- **Selection of studies (inclusion criteria-based bias)**

  In order to include a full set of available research on the study’s defined question as well as to produce the highest value from the review, it has been suggested that grey literature should be included (Hemingway and Brereton 2009; Liberati *et al.* 2009; Moher *et al.* 2009b). The study should attempt to source and include in the search grey literature that is relevant to the research question and meets the inclusion criteria but may not be within the peer-reviewed published literature. This may be achieved by the use of a single MESH term or a combination of several of these terms in the search, which is not limited to traditional search engines. The inclusion and exclusion of certain terms has the potential to introduce bias into the study but incorporating a larger set of search terms may offer the study some degree of control over the bias it imparts.

- **Reviewer bias**

  This bias here may develop from such circumstances as the reviewers reviewing their own article or reviewing an article based on their positive or negative relationship with the author/s of the article or any of the other factors that result in review variability between examiners (Belur *et al.* 2018). Additionally, there are those factors that may affect the group of reviewers such as the timing of their receiving reviews or different personal circumstances among the reviewers. This is particularly prevalent in systematic reviews utilising multiple reviewers (Belur *et al.* 2018), as is recommended (Higgins, Green and Collaboration 2011). Hemingway and Brereton (2009), supported by Liberati *et al.* (2009), reiterate this recommendation, stating that the review process with two or more reviewers provides a control for and reduces the risk of inclusion bias. Further details on the implemented methodology can be found in Chapter 3.

  In designing a systematic review numerous decisions have to be made, among them the scope of review, the selection of search terms and the inclusion and exclusion criteria as well as the specific methods used. The analysis tools and decisions around these are made by more than one individual. For this reason, the use of a detailed reporting tool (like the PEDro scale) accompanied by explanations and instructions as to its use (see Appendix G), establishes clear guidelines and providing a common starting point, thus keeping information consistent between reviewers (Lombard, Snyder-Duch and Bracken 2002; Belur *et al.* 2018). Methods used in achieving this consistency may include reviewer training in respect of how coding is to be achieved, how decision-making is to be conducted, also with regard to conflict reasoning and the process by which resolution is to be achieved. These factors are relevant to a study's capacity to replicated (Belur *et al.* 2018).

  Additionally, disagreement resolution protocols need to be structured and applied consistently when there are disagreements (Lombard, Snyder-Duch and Bracken 2002; Krippendorff 2004; Belur *et al.* 2018).

- **Mismatch between reviewing tool and study**

  Any study needs to use the most contextually applicable tools. All study designs have several tools to choose from. The present study has opted to employ the PEDro scale for RCTs (Olivo *et al.* 2008; Zeng *et al.* 2015), the Newcastle-Ottowa scale (NOS) for nRCTs (Deeks *et al.* 2003; Zeng *et al.* 2015), and the Liddle scale for CS/CR (Deeks *et al.* 2003). A detailed explanation is available in Chapter 3 (Section 3.9).

- **Reporting bias**

  In the reporting of the outcomes the data needs to be presented consistently from the data collection tools and the analyses need to be consistent in their application of the principles associated with study limitations (Kirkham *et al.* 2010). Thereafter, the overall grading of the evidence in support of an intervention requires the consistent application of a ranking system, such as the GRADE (Brożek *et al.* 2009) or AGREE (Dagenais and Haldeman 2011) or through other systems such as Foley *et al.* (2003). These data tools rank evidence per intervention (thus, via an aggregate of articles that test the same treatment). A conclusion based on these grading systems describes the strength of evidence for or against a clinical intervention grouping in terms of the modality and treatment site. In some grading systems the conclusion also indicates whether the evidence is homogenous or heterogenous in terms of the evidence available. A detailed explanation is available in Chapter 3 (Section 3.10).

## 2.10 Conclusion

Following the overview of efforts to address RA pathology (see Section 2.5) by means of a detailed evaluation of the mainstay of treatment efforts, namely, pharmaceutical interventions (see Section 2.7.1.1), and the proposed impact that manual therapy (manipulation, massage, and exercise) can provide for patients suffering the effects of RA (see sub-sections of 2.8). Chapter 2 then explains the gaps in knowledge and the role the systematic review can play in order to address them. The following chapter (Chapter 3) provides a clearly outlined and well-defined protocol / methodology utilised in producing, analysing, and contextualizing the available data in order to draw conclusions answering the study questions.

# CHAPTER 3: Research methodology

## 3.1 Introduction:

This chapter lays out the methodology used in this study. Starting with the research and covering the issues of Permissions, Location and Advertising, from Sample, Chapter 3 proceeds firstly to outline the procedure used to search for the literature included in the study and then to outline the means by which the researcher selected the material for the study and defined the inclusion and exclusion criteria. Chapter 3 also covers the steps taken to set up the study, viz., the selection of reviewers, the review process, the measurement tools, and system of evaluation employed, the analysis of the data obtained, the issues of ethics and bias and, lastly, the synthesis of the results are presented in Chapters 4 and 5.

## 3.2 Research design:

This study follows a mixed methods methodology: it is a systematic review incorporating both qualitative and quantitative data analysis (Harden *et al.* 2018).

The quantitative portion of the study evaluates the methodological rigour (the level of evidence provided by an individual article) of the selected material, while the qualitative portion of the study is the contextualization of the methodological rigour within the parameters of the study, measured not by the scales but as it affects the outcomes within each article as well as the clinical outcomes achieved by the study (the contextual limitations of the research). This is in line with the requirements of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) and Cochrane Collaboration guidelines for systematic reviews.

## 3.3 Permissions:

- Permission for this study to proceed was obtained from the Faculty Research Committee (FRC) at the Durban University of Technology (DUT) (Appendix A).
- Registration of the systematic review with the international prospective register of systematic review (PROSPERO) (PROSPERO ID: CRD42020205554) (Appendix B).
- DALRO approval for copyright permission in the use and distribution of included articles in the study (Appendix C).
- Agreement from the reviewers to participate in the study was formalized through the Memorandum of Agreement (MoA) (Appendix F).

## 3.4 Location of study

This systematic review was conducted at the DUT using journal articles acquired from online databases, with reviewers located at their respective sites.

## 3.5 Advertising:

No patient recruitment was required for this study. Therefore, no advertising was undertaken. The participants, namely, the reviewers, were selected (by means of searching academic and research platforms for practitioners and research fellows, along with referral by potential candidates and supervisor recommendation that meet the necessary criteria stipulated below) and approached by the researcher, after a search for candidates meeting the requirements stipulated in the reviewer inclusion criteria (Section 3.7).

## 3.6 Sample:

To obtain a data sample a systematic article search, and review were undertaken. This process is laid out in the following steps and in the PRISMA flow diagram (Appendix D).

### 3.6.1 Literature Search

A literature search was conducted on three search engines/databases, viz: Pubmed, Google Scholar and EBSCO - The Cumulative Index to Nursing and Allied Health Literature - (CINAHL). The search was conducted on the following key words: “rheumatoid arthritis” and a combination of one or more of the following: “RA”, “treatment”, “management”, “chiropractic”, “osteopathy”, “physiotherapy”, “physical therapy”, “massage”, “mobilization”, “manipulation”, “joint oscillation” and “exercise”. These search terms were systematically combined in a manner that provided a broad search net, offering a thorough consideration for the potential key terms used in the literature synthesised from the search. This was in accordance with the inclusion/exclusion criteria, the search produced 1 991 citations (Google Scholar (1 730 citations), EBSCO (CINAHL) (179 citations), and Pubmed (82 citations)).

### 3.6.2 Sample characteristics:

#### 3.6.2.1 Inclusion criteria:

- The articles were required to be in English or translated into English, in order to reduce translation errors that could affect the outcomes of the review process (Harris 2014).
- Articles were required to be in an electronic format or convertible into an electronic format in order to access their citations, abstracts and full texts easily as well as to enable efficient distribution to reviewers.
- RCT, nRCT, and CR/CS were included in this study, as well as grey literature that conformed to one of these designs (Hemingway and Brereton 2009; Liberati *et al.* 2009; Moher *et al.* 2009b).
- Articles were required to include the following key search terms: “rheumatoid arthritis” and a combination of one or more of the following: “RA”, “treatment”, “management”,

"chiropractic", "osteopathy", "physiotherapy", "physical therapy", "massage", "mobilization", "manipulation", "joint oscillation", and "exercise".

- Only articles where the intervention was applied to adult (18 years and older), human subjects were included in the study.

#### 3.6.2.2 Exclusion Criteria

- Previously conducted systematic reviews and literature reviews, or articles containing any expert opinions, narrative reviews, blogs, gazettes, and guidelines (Harris 2014).
- Studies involving the treatment of RA by means of injectables, medication or invasive manual therapies such as dry needling or acupuncture, surgical intervention, or therapies not delineated in the definitions section of this proposal.
- Studies involving hydrotherapy exercises.
- Studies with biological, serological, and non-clinical outcomes.
- Studies conducted on juvenile RA.
- Studies conducted on animal models.
- Articles that were not available online or in paper versions (which could not be found through the university inter-library system for scanning into the database).
- Non-English language articles.

Additional citations were added through other sources, such as those found in literature article searches, through inter-library loans system and the input of the Co-investigator/Supervisor (Dr C Korporaal). In total 87 such citations were included. Totalling 2078 citations.

### 3.6.3 Steps involved in setting up this systematic review

#### 3.6.3.1 Citation selection

The citations, with abstracts attached, were up-loaded onto EndNote (the citation manager used during this study). Duplication search was then conducted on the uploaded citations which resulted in 102 citations being excluded, bringing the total number of citations still included to 1 976.

#### 3.6.3.2 Screening of selected citations

Screening of full titles and abstracts of citations according to the inclusion/exclusion criteria resulted in 1 890 citations excluded for not meeting the criteria (among them animal-based studies, pharmacological studies, validation studies, clinical guidelines, editorials, reviews, and analyses). Ten citations were excluded because no English version or translation was available.

As a result, 1 890 citations were excluded, leaving 86 citations still included in the search.

#### 3.6.3.3 Articles obtained

Full articles were obtained for 86 citations. With reference list hand searching an additional 250 articles were added if carrying plausible eligibility (Lefebvre *et al.* 2019). Of these 336 articles, 12 articles were requested via inter-library access, 29 were not accessible in any form, leaving 307 (of which 12 were pending). The search was conducted through EndNote online full text finding function, manual online searching via DUT library SUMMONS, DUT library access to Google Scholar and the assistance of the post-graduate librarian.

#### 3.6.3.4 Selecting the final number of articles included

Two hundred and ninety-five full text articles assessed and 12 articles pending brought, the total to 307 articles. Eighteen articles were excluded because no English version or translation was available, with a further 228 articles excluded on grounds of falling outside the scope of the study. Items were excluded from the master list if they were posters or conference proceedings that did not have published / associated publications (articles) or if they focussed on systemic outcomes (other than musculoskeletal) and were post-surgically or surgically RA patient-related. In other words, articles included in the study therefore needed to have had a musculoskeletal outcome.

The remaining 61 articles of these 12 articles remained pending. Pending articles were included if they became available prior to FRC approval for the data collection period to start. These were added to the master list according to the categorisation seen in Tables 3.1 and 3.2.

#### 3.6.3.5 The Faculty Research Committee's approval of this study

On 21 August 2020 FRC gave approval for this study (Appendix A). This represented the starting point of PROSPERO**:** registration. As of this date, therefore, one of 12 pending articles had been obtained, rendering the other 11 excluded purely owing to access difficulty in any format of the article. The master list thus included 50 articles at the start of the study.

Once the reviewers signed the MoA (Appendix F) and officially participated in the study, they had the option to advocate for the inclusion or exclusion of particular articles within the study. Based on their rationale and the meeting of the inclusion/exclusion criteria, these articles were then either included in or excluded from the study accordingly.

### 3.6.3.6 Sample size:

This study initially incorporated 50 articles. Tables 3.1 and 3.2 below indicate the number of articles per article type included.

Table 3.1: Article numbers characterized by study design.

| Article type | Number of articles |
|---|---|
| RCT | 30 *(1 grey lit)* |
| nRCT | 10 |
| CR/CS | 10 *(1 grey lit)* |
| Grey literature – dissertations or other | 2 |
| Total | 50 |

Table 3.2: Article numbers characterized by intervention

| | Manipulation / mobilisation / oscillation | Massage | Exercise |
|---|---|---|---|
| RCT | 2 | 1 | 27 *(1 grey lit)* |
| nRCT | 0 | 1 | 9 |
| CR/CS | 1 | 1 | 8 *(1 grey lit)* |
| Grey literature | 0 | 0 | 2 |
| ***Total*** | ***3*** | ***3*** | **44** |
| Total | **50** articles were included | | |

*Articles were categorised by most prominent intervention category. Many studies were multidisciplinary, and this will be stratified in Chapter 5.*

## *3.7* The reviewers

The first step was to identify reviewers and obtain approval from FRC (Appendix A). Next, the reviewers identified were contacted in order to establish their eligibility (they were required to submit a short bio-sketch) and to obtain their agreement to participate in the study. The MoA (Appendix F) was concluded with each of the reviewers. They then received a copy of the articles as per their allocated group together with a briefing on the study explaining its requirements, aims and objectives as well as the methods used to evaluate the articles (utilising standard reviewing scales in the form of the PEDro, NOS and Liddle scales (Liddle, Williamson and Irwig 1996a; *PEDro Scale* 1999; Wells *et al.* 2003). They were then given a period of eight weeks (only) from the point at which all reviewers within the allocated grouping had signed the MoA. The results of each article from each reviewer were emailed to the research supervisor, Dr CM Korporaal, for retaining until such time as the researcher had completed his review of the articles.

In order to achieve a diverse set of reviewers, they were selected according to the following criteria per review group supported by Higgins, Green and Collaboration (2011):

- There must be a minimum of two external reviewers plus the principal researcher to review each article and preferably an uneven number for decision making purposes.
- Reviewers must hold a master's or doctorate in one of the following areas: chiropractic, physiotherapy, physical therapy, biokinetics, osteopathy, naturopathy, and medicine (in the fields of rheumatology, orthopaedic surgery, or any other specialisation involved in the management of or research into RA).
- For the purposes of range and diversity reviewers must have experience in one principal category but preferably other categories as well.
- Ideally, reviewers should be from different geographic locations so as to limit inadvertent inter-reviewer communication, to reduce bias, to increase the significance of agreement on scale scores for each article reviewed and to strengthen the methodological rigour.

The study used 50 articles for the systematic review. Six groups (see Table 3.3) were created from the available participants (with two reviewers and the principal researcher in each group). Each group was required to review between seven and 11 articles over an eight-week period.

### 3.7.1 Reviewers

The researcher approached multiple potential reviewers meeting the study requirements, so of whom accepting and some declining; this process continued until the required number of 12 reviewers was met. Further details of those accepting are supplied in Table 3.3 and 3.4.

Table 3.3: Template example of the reviewers (including the researcher)

| Article review group | | Masters Qualification | PhD Qualification | Academic | Clinical | Research |
|---|---|---|---|---|---|---|
| A1 | Reviewer 1 | | X | X | | X |
| A2 | Reviewer 2 | X | | | X | |
| B1 | Reviewer 3 | | X | X | X | X |
| B2 | Reviewer 4 | X | | | X | |
| C1 | Reviewer 5 | | X | X | X | X |
| C2 | Reviewer 6 | X | | X | X | X |
| D1 | Reviewer 7 | | X | X | X | X |
| D2 | Reviewer 8 | X | | | X | |
| E1 | Reviewer 9 | | X | X | X | X |
| E2 | Reviewer 10 | X | | X | X | |
| F1 | Reviewer 11 | | X | | X | X |
| F2 | Reviewer 12 | X | | | X | |
| All | Reviewer 13 (researcher) | - | - | - | - | - |

The profession of the reviewers could be chiropractic, physiotherapy, physical therapy, osteopathy, medicine, or naturopathy (but the field was not limited to these, and participation was dependent on reviewer agreement based on the MoA (Appendix F)).

Table 3.4: Reviewer qualifications

| Reviewer | Qualification | Country / Region |
|---|---|---|
| A1 | Neurophysiologist with PhD | United Kingdom |
| B1 | Physiotherapist with PhD | South Africa |
| C1 | Physiotherapist with PhD | South Africa |
| D1 | Physiotherapist with PhD | South Africa |
| E1 | Biokineticist with PhD | South Africa |
| F1 | Chiropractor with PhD | Norway |
| A2 | Physiotherapist with Masters | South Africa |
| B2 | Chiropractor with Masters | United Kingdom |
| C2 | Chiropractor with Masters | South Africa |
| D2 | Chiropractor with Masters | Netherlands |
| E2 | Chiropractor with Masters | United Kingdom |
| F2 | Chiropractor with Masters | South Africa |

Once the reviewer selection process had been completed the articles were sent to the external reviewers for analysis based on the scales for each type of study as set out in Table 3.7 below. The individual articles were then critically analysed to identify the individual strengths and weaknesses of each study.

## 3.8 Review process

The review process covered RCTs, nRCT, and CR/CS, selected through the process outlined in Section 3.6. The articles were then emailed to the respective external reviewers, who had been familiarised with the inclusion and exclusion criteria used in this systematic review. Contact was made at periodic intervals with regard to the progress of the reviews (weeks four, six and a final follow-up in week seven for completion by week eight).

The reviewers recruited into the study were afforded the opportunity to add articles (see MoA – Appendix F) meeting the selection criteria into the review. Reviewers offered up 15 articles not already included in the study and meeting the criteria as seen in Table 3.5 and 3.6:

Table 3.5: Reviewer input article numbers characterized by study design.

| Article type | Number of articles |
|---|---|
| RCT | 10* |
| nRCT | 5* |
| CR/CS | 0 |
| Grey literature – dissertations or other | 0 |
| Total | 15 |

** Scholten et al. (1999) published results of an nRCT and an RCT in the same publication. For the purpose of this table the article is counted in both categories.*

Table 3.6: Reviewer input article numbers characterized by intervention

| | Manipulation / mobilisation / oscillation | Massage | Exercise |
|---|---|---|---|
| RCT | 0 | 0 | 10* |
| nRCT | 0 | 0 | 5* |
| CR/CS | 0 | 0 | 0 |
| Grey literature | 0 | 0 | 0 |
| Total | 0 | 0 | 15 |
| Total | 15 articles were included | | |

** Scholten et al. (1999) published results of an nRCT and an RCT in the same publication. For the purpose of this table the article is counted in both categories.*

These articles were reviewed as a second round of analysis by groups constructed from the same recruited reviewers (Table 3.3) according to the same characteristics and following the same review process outlined in Section 3.6.

As a result of the reviewer article input the final sample included 40 RCTs, 15 nRCTs, 10 CS/CR and two grey literature (one CS/CR and one RCT), totalling 65 articles.

## 3.9 Measurement tools and systems of evaluation:

### 3.9.1 Measurement tools:

This systematic review makes use of various sets of measurement tools. The first set are used in Chapter 4 to evaluate the internal validity of the included studies, with the second set of tools used in the grading of the evidence available for a given intervention (as seen in Chapter 5).

#### 3.9.1.1 Reliability and validity of measurement tools for RCTs considered for this study:

RCTs are usually evaluated by means of contextually appropriate tools, namely, scales of evaluation as seen in Section 2.9.1. Of these scales the best recognised and most frequently utilised include the Jadad scale (Clark *et al.* 1999), the PEDro scale (Maher *et al.* 2003), the Delphi list (Olivo *et al.* 2008) and the Downs and Blacks scale (Olivo *et al.* 2008).

According to Clark *et al.* (1999), the Jadad scale has been shown to carry low reliability ratings. It was developed for RCTs with pain as a primary outcome measure rather than as physical therapy research with various clinical outcome measures. Additionally, the scale evaluates methodological rigour based on randomisation, blinding and attrition (withdrawals and dropouts). Given that physical therapy inherently fails to achieve double blinding (Olivo *et al.* 2008), this scale frequently only measures two parameters of methodological quality. The scale therefore does not account for bias from the effect of treatment protocol specification or treatment adherence by the patient and consistent intervention administration by the therapist. All of these are considered important attributes of the physical therapy trials. Lastly, the findings of Herbison,

Hay-Smith and Gillespie (2006) that the Jadad scale fails to adequately divide the trials evaluated into high and low quality combine with the shortcomings given above to render the Jadad scale inappropriate for the present study as it is incapable of distinguishing between different levels of methodological quality.

O'Connor *et al.* (2015) indicate that the Down and Black scale has been shown to carry a reliability rating ranging between substantial to excellent. However, this scale was developed in the field of public health so, despite its reliability rating, it is unsuitable for this systematic review which focusses on physical therapy. In other words, the scale could potentially fail to recognize bias in the focus area of physical therapy, thus putting the results of the systematic review at risk of skewing (Olivo *et al.* 2008). For this reason, Down and Black was deemed inappropriate for our study and excluded.

Like the Jadad scale, the Delphi list has been shown to carry a low reliability rating (Olivo *et al.* 2008). This scale was developed to assess RCTs without a specific area of interest. Nevertheless, it is considered within scope for evaluating physical therapy practice (Olivo *et al.* 2008). It has been the base for the development of a number of other scales, among them the PEDro scale (Olivo *et al.* 2008). In view of the unique requirements and limitations of the field of physical research, however, its low rating and lack of specificity for physical therapy research, together with the unique requirements, and limitations of the field of physical therapy research, put the Delphi scale and the accuracy of its evaluation of methodological quality at risk of bias in itself. This is a bias that could skew the results of a systematic review such as the present one.

The PEDro scale remains the most frequently used tool in evaluating the methodological quality of physical therapy RCTs (Olivo *et al.* 2008) and, according to Maher *et al.* (2003), carries a reliability rating of fair to good. The use of the PEDro scale is not only supported by Maher *et al.* (2003) but also by Macedo *et al.* (2010), Bhogal *et al.* (2005), Sherrington *et al.* (2000) and Zeng *et al.* (2015). For example, the PEDro scale, which was developed from the Delphi List within the context of physical therapy, offers a more comprehensive measure of methodological quality in stroke rehabilitation literature than the Jadad scale (Bhogal *et al.* 2005). This scale takes into consideration blinding, concealment of allocation, intention-to-treat analysis and follow-up assessments (Olivo *et al.* 2008). In light of the above, the PEDro scale is most appropriate for the systematic review and is also the most appropriate for a systematic review in the given field of physical therapy, as such, it is better suited to the focus of the current study than the Jadad scale, the Down and Black scale and the Delphi list.

#### 3.9.1.2 Reliability and Validity of measurement tools for nRCTs considered for this study:

Despite the variety of available scales to evaluate the methodological quality of nRCTs, few are frequently implemented in research endeavours in health and physical therapy interventions (Quigley *et al.* 2019). Among these scales are the NOS (Wells *et al.* 2003), ROBINS-I (Sterne *et al.* 2016), and MINORS (Slim *et al.* 2003).

According to Quigley *et al.* (2019), the reliability and validity ratings for both MINORS and ROBINS-I are unclear. On these grounds they have been excluded from use in the present study.

The NOS, by contrast, is reported by Hartling *et al.* (2013) and Quigley *et al.* (2019) as carrying a reliability and validity rating of fair. It should be noted, however, that conflicting evidence regarding the usefulness of the NOS is to be found in systematic reviews. On the one hand Stang (2010) states that the NOS lacks the supporting evidence for reliability and produces little valuable data for systematic review use, while according to Hartling *et al.* (2013) the NOS has limited ability to identify bias and requires revision. Deeks *et al.* (2003), on the other hand, find that NOS is both user-friendly and suitable for systematic review use, which is supported by the Cochrane group's statement that they both accept the NOS scale and endorse its use (Higgins, Green and Collaboration 2011). Furthermore, (Wells *et al.* 2003) evaluate the scale as carrying a high inter-rater reliability in the context of cohort study design and report not only good reliability but also high validity in this regard.

#### 3.9.1.3 Reliability and validity of measurement tools for Case studies / reports / series

CS/CR have habitually been overlooked as a source of knowledge with regard to evaluating the efficacy of physical therapy. Reinforcing this attitude are the weak inferences of treatment efficacy and effectiveness as well as the higher risk of bias. Nonetheless, CS/CR have influenced medical literature and the continued advancement of medical research (Murad *et al.* 2018). The reporting of observations and clinical pattern recognition, among other things, has contributed to this. It must be noted that anecdotal observations, as a CS/CR would be considered to be, particularly when limited to a single case, represent the lowest form of scientific evidence and consequently do not substantiate implementation into clinical practice (Pierson 2009).

Huth (1999) offers four scenarios where CS/CR can offer valuable insights for medical literature. These are:

1. unique/unknown syndromes or diseases;
2. unreported associations or suggesting relationships;
3. outliers to the known presentation;
4. an unexpected response or course which suggests a previously undocumented effect, either therapeutic or adverse, of a particular intervention (Pierson 2009).

Because CS/CR have tended to be overlooked, limited efforts have been made to develop tools for evaluating the evidence derived from and the quality of methodological rigour of their evidence (Murad *et al.* 2018). This kind of evaluation is of particular importance in systematic reviews. Three such tools utilised in evaluation include the Pierson criteria (Pierson 2009), the Bradford Hill criteria (Hill 1965) and the Liddle scale (Liddle, Williamson and Irwig 1996a).

The Pierson criteria have rarely been used and have been said to "conflate methodological quality with other constructs" (Murad *et al.* 2018). The limited use of these criteria, the paucity of clear validity and reliability evaluation and the risk of conflating results combine to exclude them from the present study.

As noted in the study by Hill (1965), the Bradford Hill criteria have been utilised in the evaluation of CS/CR of adverse drug reactions and are considered a standard tool of assessment in the impact of broad-based public health interventions (McDonald and Strang 2016). The lack of evaluation both with regard to validity and reliability, particularly in the field of physical therapy-based research, imparts the risk of skewing the outcomes of the systematic review. For this reason, the Bradford Hill evaluation criteria have been excluded from the present study.

The Liddle scale was developed with a standardised method of approach in order to evaluate the quality of evidence in a consistent manner. The Liddle scale provides an overall assessment of quality by evaluating separate criteria of each source of bias and the extent to which each influences the results of the study. When combined, the collective criteria provide an overall assessment establishing the methodological quality of the study according to sound epidemiological principles (Liddle, Williamson and Irwig 1996a; Harris 2014).

In terms of the above survey of the tools available for evaluating the interventions of interest for the present study (manipulation, massage, and exercise), the scales best suited to this systematic review, namely, the PEDro scale, the NOS and the Liddle Scale, have been selected as producing the highest quality of data. This is summarised in Table 3.7 below which uses these three scales to review the various types of studies and additionally indicates the most appropriate evidence grading system applicable for the purposes of the present study and systematic review.

Table 3.7: Summary of the utilised scales in this review.

| Type of study | Scale | Reliability and Validity of the scale | |
|---|---|---|---|
| RCT | PEDro Scale<br>(*PEDro Scale* 1999) | Fair to Good<br>(Sherrington *et al.* 2000; Maher *et al.* 2003; Bhogal *et al.* 2005; Macedo *et al.* 2010; Zeng *et al.* 2015) | Appendix G |
| nRCT | Newcastle-Ottawa Scale (NOS)<br>(Wells *et al.* 2003) | Fair<br>(Zeng *et al.* 2015; Quigley *et al.* 2019) | Appendix H |
| Case Series/Studies (CS) | Liddle Scale<br>(Liddle, Williamson and Irwig 1996a) | Validity and clinical reliability are present<br>(Vlayen *et al.* 2005; Siering *et al.* 2013) | Appendix I |
| Case Reports (CR) | Liddle Scale<br>(Liddle, Williamson and Irwig 1996a) | Validity and clinical reliability are present<br>(Vlayen *et al.* 2005; Siering *et al.* 2013) | |

Once the external reviewers had reviewed all the allocated articles, a summary was drawn up based on the findings of the included articles and following the structure of the PRISMA guidelines (Moher *et al.* 2009a). After identifying any discrepancies between reviewer feedback with regard to any article had been identified, the areas in question and commentary of the reviewers were collected and sent to the relevant reviewers to enquire if any re-evaluation would be considered. Based on the feedback, the data for each article was closed off and analysed.

### 3.9.2 Description of chosen measurement tools selected for the purposes of this systematic review:

#### 3.9.2.1 Randomised Controled Trials:

For the RCTs, as previously indicated, the most appropriate scale to use for assessment purposes is the PEDro Scale (Appendix G). This scale focuses on 11 criteria, each of which is used to evaluate the strength of the methodology used in the studies to which the scale is applied. Each of the criteria that is met contributes a single point, with a maximum of 11 and a minimum of zero points. An average of the three submissions in total, was calculated. The mean score correlates with the strength of the methodological rigour. If an article receives the highest score, namely, 11 across the three reviews, the highest methodological rigour can be assumed. In contrast, should the article receive, the lowest score, namely, zero across the three reviews, the lowest methodological rigour can be assumed.

#### 3.9.2.2 Non-Randomised controlled trials:

For the assessment of nRCTs, the NOS (Appendix H) is the most appropriate scale to use. The scale contains three headings: Selection, Comparability and Exposure. These headings have a variable number of questions (four, one and three respectively). Each question comes with a set of answers to which the participant must select the answer most appropriate to the article being reviewed. Points are awarded accordingly. The headings “Selection” and “Exposure” can each be awarded a maximum of one point per question. The heading “Comparability” may obtain a

maximum of two points. With all headings and questions completed a maximum total of nine points may be awarded. An average between the three submissions (from the researcher and the two participants) per article was calculated. As with RCTs above, the mean score correlates with the strength of the methodological rigour. Should the article have received the highest score, namely, nine across the three reviews, the highest methodological rigour can be assumed. In contrast, should the article have received the lowest score, namely, zero across the three reviews, the lowest methodological rigour can be assumed.

#### 3.9.2.3 Case studies / series / reports:

For CS/CR studies the Liddle Scale (Appendix I) was used as it offers clinical flexibility and affords clinical applicability while maintaining validity as a scale measuring methodological rigour (Vlayen *et al.* 2005). The scale comprises a set of criteria that are constructed in a question format. These questions enable the reviewer to identify and assess the possible areas of opportunity where bias can be introduced into the design and methodology of the review / study's. The scale comprises two criteria, the first of these being the evaluation criteria for the study and the second the overall assessment of the study. The evaluation criteria comprised eight questions that focus on an area which the reviewer must critique, while the overall assessment criteria comprise four questions that focus on an area which the reviewer must critique.

The scale facilitates the assessment of the appropriateness of the study's procedures to ensure that guidelines and recommendations benefit the study's intended population. The scale utilises a grading system, that is subject to the reviewer's perspective on fulfilment of criterion questions.

The grading system for the first criteria is determined by fulfilment of the question (A = entirely fulfilled; B1 = mostly fulfilled; B2 = mostly not fulfilled; C = not fulfilled; I = not described and n/a = not applicable). The grading system for the second criteria determines the risk of bias in the reviewed article (A = low risk of bias; B1 = low-medium risk of bias; B2 = moderate to high risk of bias and C = high risk of bias).

## 3.10 Data reduction and analysis:

### 3.10.1 Quantitative analysis:

The methodology here involves analysing the data by recording reviewer responses as well as analysing discordance between the reviewers (if any). The data is then then be formulated in Chapter 4 (Quantitative analysis tables).

### 3.10.2 Qualitative analysis:

An analysis of the contextual limitations of the study and their impact on the outcome / generalisation of the results as well as the concluded outcomes is provided in Chapter 4 (Qualitative analysis tables).

### 3.10.3 Outcome evaluation:

Outcome evaluation derived from the above two analysis (Section 3.10.1 and 3.10.2):

Table 3.8: Review outcome matrix

| **Review outcome matrix** | Good clinical outcomes | Poor clinical outcomes |
|---|---|---|
| High methodological rigour | High level of evidence in support of the clinical use of the intervention. | High level of evidence not in support of the clinical use of the intervention. |
| Low methodological rigour | Low level of evidence in support of the clinical use of the intervention. | Low level of evidence not in support of the clinical use of the intervention. |

### 3.10.4 Grading of available evidence

In Chapter 5, after the review process, articles are grouped into clinical categories, in other words, per intervention, whether mobilization, manipulation, massage, and exercise. The level of evidence in support of these interventions on RA is first defined by the review outcome matrix, as is depicted in Table 3.8 above, which defines the articles by methodological rigour and clinical outcomes dichotomously. The nature of the evidence then ranked according to the levels of evidence as proposed by Foley *et al.* (2003).

In addition to this, Foley *et al.* (2003) and Dagenais and Haldeman (2011) introduce a further three successive categories which show whether the evidence presented in the prior paragraph can be ranked as strong, moderate, limited or no evidence in terms of the quality of the evidence reported.

By comparison with other systems, the GRADE system is significantly complex and therefore difficult to apply correctly; also it focuses on the evidence provided by RCT studies alone (Blunt 2015). Given that the present systematic review utilises evidence from a mix of study designs (RCTs, nRCTs, CS and CR) the GRADE system of evaluation has not been utilised in this study.

The AGREE system (Dagenais and Haldeman 2011) and the Foley system (Foley *et al.* 2003) of evidence evaluation were each developed on a clinical model (manual therapy and shoulder rehabilitation in stroke patients respectively), giving these systems an appropriate foundation with regard to the present study. These evidence evaluation systems (AGREE and GRADE systems) provide a standardised method for guideline development in health evidence-based models.

By comparison with the AGREE system, the Foley system of evidence evaluation incorporates additional categories (Consensus and Conflicting). This enables the Foley system to evaluate not only the strength of the evidence available as being strong, moderate, limited, or non-existent but also the degree of agreement between these studies. For the above reasons, the Foley system of evidence evaluation, and the AGREE ranking system are believed to be best suited to this study. In using both these ranking scales, the systematic review is able to report on the strength of the evidence provided and the degree of consensus between items of evidence generated for the intervention from the vantage point of all evidence compiled both for and against the modality, and thus guide future research efforts towards areas requiring evidence generation and factors needing to be considered in order to limit bias and confounding in future studies.

## 3.11 Ethics and Bias:

The nature of the study, a systematic review, rendered the approval from the Institutional Research Ethics committee (IREC) unnecessary owing to the fact that the FRC had been delegated the responsibility of systematic reviews as a level one study, which means that studies not involving human participants do not require IREC approval. Although reviewers were necessarily involved and played a crucial role in the study, they were not seen as research participants (Cuppusamy 2014; Izudi *et al.* 2018).

In order to protect both the reviewers and the rigour of the study all the reviewers, including the researcher, were denied knowledge of who the other reviewers of a particular article were. Reviewers were also required to maintain confidentiality and not to discuss the study until it was concluded, in order to reduce any bias in the feedback of results. For the purpose of confidentiality and anonymity the participants' identifying credentials were not disclosed so as to make it impossible to trace a submission (i.e., the review of an article) back to a participant. This was achieved by means of a central submission point to which the researcher and the participants submitted their evaluations (reviews). The supervisor, undertook this role, withholding the documents from the researcher until all submissions were accounted for and all identifying data was removed. In addition, the identities of the participants have not been revealed in the dissertation, though, as evidenced by Table 3.4, and noted in the methodology and the critical review, their qualifications have been specified.

### 3.11.1 Compliance with and protection of the four pillars of ethical medical research, in accordance with those defined by Gillon (1994):

#### 3.11.1.1 Autonomy:

Reviewers were not obligated to participate in this systematic review and maintained the right to leave at any point during the study. In the event that a reviewer was to withdraw from the study, an equivalent replacement would be found.

#### 3.11.1.2 Beneficence:

All reviewers benefited by way of an honorarium as well as the opportunity to be credited with co-authorship should the study be published.

#### 3.11.1.3 Non-maleficence and Utility:

Not applicable as the study did not involve any intervention or withholding of any care.

#### 3.11.1.4 Justice:

All reviewers were treated in a similar manner, based on a MoA (Appendix F).

The recommendations from this study may provide guidance with regard to the management and treatment of patients with RA. They are therefore couched within the specific contexts of the studies reviewed. In addition, issues around article selection bias, reviewer bias and publication bias have been addressed (Roopchand 2014).

## 3.12 Conclusion

As indicated in the opening paragraph, Chapter 3 formulates the methodology employed not only in setting the parameters of this study but also in searching for, identifying, and selecting the appropriate literature, as well as in extracting, analysing, and synthesising the data that emerges from this systematic review. As such, this chapter lays the foundation on which the following chapters build and in addition maps out how this will be done.

# CHAPTER 4: RESULTS

## 4.1 Introduction

The purpose of Chapter 4 is to report and synthesise the data collected from 12 reviewers who evaluated the 65 articles included in the systematic review with regard to the level of evidence they provide for the non-invasive manual therapies (viz., massage, manual therapies, and exercise prescription) used in the treatment of RA. The chapter begins by defining the primary and secondary data, after which a list of abbreviations frequently used in the chapter is provided. Next, to facilitate a concise reporting of the results, there is an analysis of common factors imparting risk of bias and confounding across the study designs. The chapter then provides an analysis of each incorporated article, which is presented in two tables. The first table, titled "Feedback of reviewer data", presents the reviewer scores, the overall reviewer agreement percentage, and the ranking according to the relevant scale utilised for that particular study. This table evaluates the internal validity of the study. The second table, titled "Properties of the study, outcome, and discussion", presents an analysis of the particular study's external validity, incorporating information such as the form and frequency of evaluation (outcome measures), the duration of the intervention, the number of participants and dropouts, the blinding of evaluators and the use or not of randomization as well as an exploration of the limitations, strengths, outcomes and conclusions of the study.

The article data is presented in groupings according to study design, namely, RCT-designed articles (as listed in Table 4.1), nRCT-designed articles (as listed in Table 4.41), and CS/CR-designed articles (as seen in Table 4.58).

## 4.2 Data

### 4.2.1 Primary data

The primary data in this study was produced from the reviews of the articles comprising the master list (a description of the search method used, and the articles incorporated in the study can be found in Chapter 3 – see Section 3.6 and 3.8). The data enabled the calculation of a ranking score and an agreement percentage representing the degree of agreement between reviewers, in terms of which the overall percentage agreement is calculated.

Based on the percentage agreement, scoring an overall agreement percentage was calculated as an average. A study obtaining an overall percentage agreement score of 70% or above is considered to have homogenous agreement among all reviewers (Liberati *et al.* 2009).

### 4.2.2 Secondary data search

This refers to the literature that has assisted in the analysis and reporting of the information as presented in Chapter 2. The collective is derived from textbooks, published articles, and grey literature, sourced through the DUT library (online platform) and search engines.

## 4.3 Commonly used abbreviations in Chapter 4

CONSORT : Consolidated Standard of Reporting Trials
DUT : Durban University of Technology
MCID : Minimal clinically important differences
nRCT : Non-Randomised clinical/controlled trials
RCT : Randomised clinical/controlled trials
CS/CR : Case Series/Studies/Reports
VAS : Visual Analogue Scale
ROM : Range of Motion
HAQ : Health Assessment Questionnaire Disability Index.
DAS : Disease Activity Score
ADL : Activities of Daily Living
AIMS2 : Arthritis Impact Measurement Scales
QoL : Quality of Life

## 4.4 Discussion of the articles

### 4.4.1 General discussion of commonalities in factors imparting bias and confounding on the articles reviewed in this systematic review.

In terms of the study design and the imparted risk it carries, RCTs are considered the gold standard of research evidence, but their inherent structure risks diminishing external validity of the study by strict sampling criteria, that may reflect a sub-population of the condition rather than the true population in its entirety and thus compromise the generalisability, of the results to patients in general and typically those patients who present to clinical practices with multiple comorbidities (Smith *et al.* 2013). This tension between the clinical relevance of RCT outcomes as opposed the effectiveness of tested interventions provides a means to critique the RCT (Kendall 2003; Wells *et al.* 2012). This requires researchers and clinicians not only to evaluate the internal rigour of the study through appropriate scales (PEDro 1999; Smith *et al.* 2013), but also to consider the context in which the clinical outcomes are generated as these factors influence the generalisability of the outcomes to patients in clinical practice (Dagenais and Haldeman 2011; Smith *et al.* 2013). The inverse applies for nRCTs and CS/CRs have reduced internal validity but better represent the clinical application despite the context of the application being narrowed to patients presenting similarly to the patients of the sample used in a specific study design.

Determining the external validity / generalisability of a study is more complex than determining the internal validity. Factors that influence external validity can be divided into patient and systemic factors. Patient factors include, among others, patient selection and baseline differences between trial patient groups and the general patient population (Wells *et al.* 2012; Smith *et al.* 2013). Systemic factors or confounding biases include the trial setting, differences between trial protocol (patient selection) and routine practice, observer bias and the impact of the intervention on the condition and / or costs for patients and providers (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013). The evaluation of the RCTs in this systematic review therefore critiques these two aspects of internal and external validity in order to determine the effect of bias or chance on the concluded outcomes for mobilisation, manipulation, massage and exercise (Kendall 2003).

Accordingly, the following discussion looks at the impact of the various possible points of bias in the RCTs applicable to manipulation, mobilisation, massage, and exercise as a basis for the review of the subsequent articles (for internal and external validity).

#### 4.4.1.1 External influencers:

These, if not taken into consideration and controlled for within the methodological designs of RA RCT, introduce confounding biases that influence the clinical conclusions (Wells *et al.* 2012) and may include the following:

- Medication monitoring (adherence and duration to prescription, changes to the prescription prior to or during the study, types of medication (see Section 2.7.1), dosage and administration methods, pharmacokinetic and pharmacodynamic characteristics, drug interactions, and drug wash-out periods) must be taken into account as the above medication-related factors allow for variance in response to manual interventions. Differential reporting of medication variables (such as pain control) is necessary in order to control for these variables and is achieved either through stringent inclusion criteria or patient stratification (Zhao *et al.* 2017; Heidari *et al.* 2019).
- Occupation/work may affect outcomes when comparing participants whose occupations are restricted by RA deformities (e.g., hand deformities) with those whose occupations are not similarly affected. In such cases, it is possible that the non-manual participants may improve to a greater extent than those participants that are required to complete manual work (Naqvi, Hassali and Aftab 2019).
- Other interventions used simultaneously by patients will negate the possibility of measuring the effect of the tested intervention (Mouton 2006).

- Surgical intervention, whether during a period prior to the study or planned for a time during the study period, will either negate measuring the effect of the intervention and / or modify participant response based on the prognosis of the surgery and not the actual tested intervention (Naqvi, Hassali and Aftab 2019).
- Hand or side-dominance of the participants may influence response to treatment, particularly on the dominant hand side and especially in the context where hand symptomatology is the main clinical outcome of the study (Lee, Gandevia and Carroll 2009; Dragert and Zehr 2013).
- Comorbidities of RA with documented presence (e.g., osteopenia, osteoporosis, asthma, depression (see Section 2.5.6 and 2.5.7) (Kłodziński and Wisłowska 2018)) must be controlled for because of their impact on participant response to care. These comorbidities may be stratified in the participant sample or specifically identified as an inclusion if subgroups of RA patients were considered for particular interventions (England *et al.* 2019; Levitsky *et al.* 2019; Myasoedova *et al.* 2020).
- Concomitant diagnosed medical conditions documented as present that may influence response to treatment (e.g., diabetic induced peripheral neuropathy), need to be considered so as to mitigate their impact on the clinical outcomes of RA RCT testing manual therapies (Bergmann and Peterson 2010; Dagenais and Haldeman 2011).
- The impact of excluding patients who have contraindications to the tested interventions even though they have RA necessitates that the participant population be explicitly defined in order to ensure that outcomes can be contextualised for patients in clinical practice (Durcan, Wilson and Cunnane 2014).
- Lifestyle factors, among them smoking/nicotine status, alcohol intake, coffee intake, and parity, need to be considered as these are factors that may influence the onset and rate of progression of RA, thus potentially confounding the outcomes related to the testing of an intervention (Pedersen *et al.* 2006a; McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019; Myasoedova *et al.* 2020).
- The endocrine status in female patients (menopausal status, pregnancy, and one-year post-natal) is not dissimilar in effect to the lifestyle factors (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018).
- General factors related to ethnicity, personality, social / cultural background, age, sex, and the severity of the condition (Bahreini *et al.* 2020). These are known influencers of, for example, how participants respond to pain (Bartley and Fillingim 2013). The pain severity of the condition is important as it relates to the possibility of outcomes being hampered by the floor or the ceiling effects if the severity of RA is not comparable between the patient groups.

These influencers can give rise to differences between the groups on which a particular intervention is tested, resulting in either favourable or unfavourable clinical outcomes (implying an acceptance of perceived outcome or a type II error) (Bartley and Fillingim 2013; de Boer *et al.* 2015). In addition, there are likely to be differences between the RA research participants and the general population seeking care for RA.

It is therefore important that the context of recruitment be considered in terms of the recruited patients (Wells *et al.* 2012). For example, single site recruitment (a single location such as a hospital setting or a single geographic area) decreases the likelihood that the patients in the RCT are similar to the general RA population on a global scale. To limit the impact of this recruitment process it is therefore critical to consider the inclusion criteria in order to control for confounding variables and to align the research patient population with the general RA population as well as to ensure that there is homogeneity between the groups receiving the tested interventions (Wells *et al.* 2012). For the RA patients the criteria could include the diagnostic criteria used, serology, functional class, disease duration and disease activity (tool for analysis example: DAS28), as well as age, pain severity, imaging studies, medication, other physical modalities in use and comorbidities (England *et al.* 2019; Myasoedova *et al.* 2020). Consideration of these complexities would better allow RCT outcomes to be contextualised within specific RA subpopulations and thus enable greater applicability to clinical practice (Kendall 2003). Careful consideration also needs to be given to the exclusion criteria applied as patients may present with contraindications to the intervention. In addition, patients may have difficulty in complying with the required regimens (Kendall 2003).

The sample size and power of the study needed to draw valid conclusions must be calculated prior to the start of the study in order to ensure that the sample size of the groups and the overall sample of patients in the study will be sufficient for the clinical outcomes and conclusions drawn to be better than chance alone (Kendall 2003; Christley 2010). This is usually done with reference to the minimum clinically important differences (MCID) that are necessary for clinically effective change to be experienced by patients (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020). Without this baseline starting point it is often seen as unethical to have patients participate in an underpowered study that will from the outset not be able to draw definitive clinical conclusions with regard to the tested interventions (Mouton 2006; Christley 2010). Linked to this, the recruitment process needs to ensure that there is adequate provision for lack of patient adherence or dropout (Fogg and Gross 2000).

Randomization in an RCT is mandatory. However, the method used and the degree of concealment vary and present different bias challenges (Mouton 2006). The purpose of randomization is to prevent bias resulting from non-comparability between groups (de Boer *et al.* 2015), as this implies that the clinical outcomes are not necessarily only due to the interventions (Fogg and Gross 2000; Mansournia *et al.* 2017). This bias in clinical outcomes also imparts a

negative effect on the intention-to-treat analysis and post hoc analyses (de Boer *et al.* 2015; Unbiased Research 2017), thus potentially enhancing favourable or unfavourable outcomes (Fogg and Gross 2000; Kendall 2003; de Boer *et al.* 2015).

To ensure that randomization is effective and achieved without interference from the researchers the allocation is best done through computer-generated randomisation tables and applied through concealed allocation. To eliminate researcher bias in the group allocation process it is also essential that placement of the patient into the study groups is done after they have been identified as meeting the inclusion criteria and have agreed to participate in the study (Kendall 2003). Other considerations that may be important in the context of the many external influencers in the RA population are stratified randomization, block randomization and cluster randomization of the patients even though they are allocated randomly. This is to minimise the effect of the external influencers (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019).

After randomization has been completed and prior to the analysis, baseline comparison between the groups needs to be assessed for variance between the groups (de Boer *et al.* 2015). This baseline analysis should include external influencer factors as well as the repeated measures of clinical outcomes (Pocock *et al.* 2002). Deviation between the groups may impact the testing of the interventions and introduce type II error into the study (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).

#### 4.4.1.2 Patient-related effects that may influence the outcomes of the RCT include the following:

- The placebo effect or the inherent non-intervention treatment effect. This needs to be considered for all intervention groups. This may be dictated by doctor-patient interaction dynamics not being equal between the intervention groups, thus introducing bias. It may be influenced by patient perception of the intervention and / or the expectation that the patient has of the research process and the way the researchers emphasise the possibility for a clinical outcome. This is of particular importance where manual therapy interventions are unable to achieve double blinding; the placebo effect can influence studies with pain as a clinical outcome since it the placebo effect has an impact on how pain is reported (Kaptchuk 1998; Bialosky *et al.* 2011).
- The Hawthorne effect or the effect that the implied status of the practitioner has on the patient and how the patient anticipates that they should react to the questions by the practitioner / researcher. Like the placebo effect, this effect is something that needs careful monitoring in the context of clinical trials in which only subjective clinical outcomes are completed by the patient (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015).

- Patient naivety or the degree to which the patient has been or not been exposed to the intervention being tested. Those patients with preconceived ideas about the intervention can abnormally enhance or detract from the clinical outcomes. The optimal situation for testing an intervention and its clinical effect is when patients are completely naïve to the interventions (Mouton 2006). This may be difficult to achieve with chronic RA patients in that they are likely to have considered a wider option of treatment interventions as compared to recent onset RA suffers who have not explored the intervention options open to them.
- Patient blinding or not informing patients as to which group, they have been allocated. This relates to the issue of patient naivety and is also important in controlling the degree to which the patient responds to the knowledge of being in the normal treatment / control / placebo group versus the group on which the intervention of interest is being tested. Generally, people in the normal treatment group tend to underreport findings by comparison with those in the novel treatment group (Moustgaard *et al.* 2014).
- Memory recall and memory decay. Both these factors can influence the patient's reporting of clinical improvement. If subjective clinical outcomes are administered too often patients are more likely to be able to recall previous reported outcomes and to manage their response in terms of what they think they should be reporting as opposed to the clinical reality. Memory decay, on the other hand, influences outcomes when measurements are not taken at the time and patients are asked at some later stage to recall clinical outcomes retrospectively (Mouton 2006; Ricker, Vergauwe and Cowan 2016).
- Touch therapy. Many manual interventions involve some form of physical contact, which imparts a touch effect. Thus, if the RCT compares a manual to a non-contact intervention, the manual intervention patients are more likely to be satisfied with the treatment as a result of perceived interaction by comparison with the non-contact intervention patients (Senderovich *et al.* 2016). The work of Melzack and Wall (1965) and Moayedi and Davis (2013) offers further support for this concept.
- Pain threshold and pain tolerance also need to be considered in terms of the outcomes of RA studies since many of the RA RCTs use pain as an outcome measure. The differential reporting of pain has been linked to sex, age, and chronicity of the condition. Additionally, this reporting affects outcome measures like pain pressure thresholds (Dhondt 1999; Melia *et al.* 2019).

#### 4.4.1.3 Researcher related effects that may influence the outcomes of the RCT include the following:

- Blinding of the patient assessor in terms of participant allocation to their groups decreases the possibility of assessor bias in allocating patients to specific groups in order to show artificially that the intervention has benefit / no benefit (Moustgaard *et al.* 2014).
- Blinding of the outcome/s measurer. It is important that the person who assesses repeated outcomes is not made privy to the groups that the patients have been allocated to in order to ensure that this person cannot manipulate or influence Hawthorne or placebo effects or, by actual changes in recorded findings, the outcomes attained and reported (Kendall 2003; Moustgaard *et al.* 2014).
- Blinding of the person providing the intervention is a crucial in the double-blinded RCT that is often utilised in pharmacological RCTs. In manual therapy, however, it is impossible to blind the therapists as they are applying the intervention in person. According to the literature on the subject, one method by which this may be addressed is to have two different therapists applying therapy to the intervention groups so that neither knows what intervention the other is applying (Moustgaard *et al.* 2014). The concern with this RCT structure is that the interpersonal effects affecting the placebo effect, the Hawthorne effect and the interaction based on touch therapy cannot be as easily controlled for as when using only one therapist. The argument against this is that the single therapist structure may also perpetuate differences between the groups as the inherent perception of the intervention by the therapist may be inadvertently conveyed to the patient, thus resulting in the same outcomes as having two therapists applying the two interventions (Kendall 2003). One way to control the impact of blinding is by asking assessors to guess treatment assignments; should a significant proportion correctly guess the assignment, then the potential for this as a source of bias needs to be considered (Kendall 2003; Moustgaard *et al.* 2014).
- Protocol consistency or its absence among multiple researchers and research assistants participating in a study can be a cause of patient choice variability, therapist variability or assessor / outcomes measure variability. The manner in which protocol consistency is maintained is often unreported in many RCTs, but could easily be achieved through operation manuals, research meetings and / or role specific training within the context of the research in order to standardise actions and increase the quality and integrity of the results so as to reflect true clinical outcomes (Kendall 2003).

#### 4.4.1.4 Outcome measure / data collection tool-related effects that may influence the outcomes of the RCT include the following:

- Choice of the outcome measures. It is important that outcome measures directly measure the intended effects of the intervention. For example, the use of ROM outcomes for massage would be less sensitive and intervention-specific by comparison with manipulation. Using pain pressure threshold over tender points in the muscle as a measure is more appropriate when measuring effects of massage (directed at muscle) by comparison with mobilisation (targeting joints) (Coster 2013).
- Validity, reliability, and suitability of the outcome tool. These concepts relate to the how well the outcome measure actually measures change and how this is interpreted clinically (Coster 2013; Bahreini *et al.* 2020).
- Sensitivity and specificity. These are additional concepts needing to be considered. The tool must be sensitive enough to detect change and specific enough to the condition under study to ensure that the outcome is accurate (Coster 2013; Bahreini *et al.* 2020).
- Alignment with the objectives of the study. The outcome measures need to align with the study's objectives for the study to be able to draw appropriate conclusions about the effectiveness of the intervention/s. This also requires that the outcome measures are comprehensive in reporting various aspects of the RA and the changes the intervention is expected to achieve (Coster 2013).
- Objective and subjective outcome measures. There should be a balance of these (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Performance bias implies a systematic difference between groups that is shown in evaluating the different outcome measures, other than those which could be attributed to the interventions. The mechanism of performance bias is multifaceted which may involve differences in the quality and content of the participant-researcher interaction as with the Hawthorne effect, different degrees of the placebo effect, and differences in the basic care (Moustgaard *et al.* 2014).
- Ceiling and floor effect (Liang 2000; Coster 2013). This issue relates to the properties of sensitivity and responsiveness of the measurement tool. Sensitivity is an indicator of the instrument's ability to detect change in a population beyond measurement error (Liang 2000). Multiple factors can influence a measure's sensitivity the presence of ceiling effect / restriction of range. If most of the participants are already performing well at baseline (i.e., their scores are close to the top of the scale), then the scale will not be able to capture much change in that group. Because statistical analyses often use sample variability, the results of analysis may be misleading if little variance is seen. The same effects are possible in the floor effect but on the opposite end of the scale (i.e., when the scale measure decreases are limited by baseline measures). By contrast, responsiveness is evaluated by showing a relationship between change on the measure and change as measured by a gold standard (Coster, Ludlow and Mancini 1999).

- Calibration of the outcome measure tools. This is an important consideration in studies that have objective outcomes requiring maintenance and servicing in order to provide accurate clinical data as outcome measures for clinical trials, the reason being that the reliability and validity of the outcome data underpin the conclusions made in the study in terms of the intervention's efficacy or effectiveness (Kimberlin and Winterstein 2008).

#### 4.4.1.5 Details related to the intervention that may affect the outcomes of the RCT include the following:

- Study duration. This needs to be designed around the hypothesised clinical effect of the intervention. If the intervention is suggested to cause immediate changes, then a pre-post study design is appropriate, whereas if the study requires changes in fascial or collagen structures (tendons) that are measured, then a period of at least three months is needed before changes can be clinically detected (Fraser 2008).
- Treatment protocols. These need to be clearly delineated. For example, exercise protocols must detail loads, load progression, volume of exercise, repetitions, frequency, intensity, sets, rest intervals, duration of programme, specificity of exercises and individualisation of the protocol to the participant/s. This detail is important as variances in these applications can be possible reasons for the outcomes attained and may impact on the application of the exercise protocols for RA patients in practice, therefore specific protocol weaknesses need to be identifiable (Heine *et al.* 2012).
- Programmes used in the treatment of patients. Within these, the clinical outcomes attained by the patient may inadvertently be impacted by the synergistic effects of interventions on one another, thus amplifying the outcomes which do not reflect in individual intervention applications (Robins, Hernán and Siebert 2004) or the antagonistic effects of interventions on one another (decreasing the outcomes which do not reflect in individual intervention applications) (Robins, Hernán and Siebert 2004). This complicates identifying the causal clinical effects of interventions.
- Wash out periods. These are defined as periods in time that are utilised to minimise the effects of an intervention (not the tested intervention) on the outcome of the study (Evans 2010). Washout periods are important in terms of medication changes in the period before and during an RA study. The pharmacological action of the medication may continue into the study period, thus, impacting on the measured outcomes and as a result reflecting the change in medication and not the effect due to the intervention. This is also true of manipulative, massage and exercise therapies that may have effects requiring a wash out period to negate their impact on the study. This is particularly true in crossover studies where the first intervention arm may have an impact on the second intervention arm and thus colour the conclusions drawn in terms of the effectiveness of the interventions (Evans 2010).

#### 4.4.1.6 The relationship between the intervention and measurement / data collection points that may influence the outcomes of the RCT includes the following factors:

- The timing of measures in relation to the treatment procedures. This is important as measurements taken directly after an intervention, that is, on the same day and time so as to avoid the diurnal variation of RA (Nandgude, Hasabe and Kolsure 2018) only denote the immediate action of the intervention irrespective of whether this happens at the first or the last visit in the research process. No comments or conclusions can therefore be drawn with regard to short term or long-term success or failure (Pater *et al.* 1998).
- The condition in which measures are taken. This requires that the measurements cannot initially be taken at the research centre and then have the same follow-up measures taken at the patient's home or telephonically (Richardson 2007). Each of these different means of collecting data has a different perceived relationship to the intervention and is often ranked by the patient on a scale from important (in the research centre) to less important (over the phone). The reporting may thus be more indicative of this perception than the clinical reality of the patient at that time, which impacts the outcomes negatively.
- The period for the clinical effect of the intervention must be manifest. For example, literature suggests (as has been stated elsewhere) that it takes a minimum of six to eight weeks for muscle strength to improve in response to an intervention directed at muscle strengthening (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). This implies that measuring an outcome of muscle strength for a given programme will only really measure effect after this period of time has passed, a factor that may explain the difference in the outcomes attained by Bell (1998), Breedland *et al.* (2011), and Williams *et al.* (2015).
- Choice of control. The use of unilateral versus bilateral measures in the context of single-side interventions also impacts on the outcomes as there have been several mechanisms that suggest that the treatment of one side may affect both sides based on the systemic effects of the intervention or through a neurological link (Dragert and Zehr 2013; Villafañe, Cleland and Fernandez-de-Las-Peñas 2013). Additionally, the use of the untreated opposite side as the control may decrease the effectiveness of the intervention since both sides respond to the invention but at different magnitudes.
- Time-modified confounding occurs when there is a time-fixed or time-varying cause of disease that influences subsequent treatment and when the effect of this confounder on either the treatment or outcome changes over time (Platt, Schisterman and Cole 2009) This can be seen in the effect of timing of the outcome measures on a given day and on the condition. Another time-modified confounder could be an unpredictable flare-up of RA during the study period (Kaur, White and Bartold 2012; Walker *et al.* 2013).

#### 4.4.1.7 Statistical considerations that may influence the outcomes of the RCT include:

- Post hoc analysis. Statistical significance such as the use of post hoc analysis to control for this bias is not recommended (Fogg and Gross 2000; Kendall 2003). There is an argument that the use of baseline testing is not appropriate in the reporting of RCTs. The argument that the use of baseline testing to determine whether post hoc analysis and / or confounder variance controlling are applicable is not appropriate in analysing the outcomes of the study as it "serves no purpose and can be misleading" (de Boer *et al.* 2015). This is especially true when this process does not acknowledge the prognostic strength of the covariant. It is thus more important to utilise appropriate sampling techniques and more rigid inclusion criteria to improve baseline similarity between test groups in order to more effectively understand the intervention than to control for differences statistically after the fact. Thus, there is a tension between the requirements of the CONSORT (Begg *et al.* 1996; Kendall 2003; Moher *et al.* 2012) and the impact of the practice/intervention being contextualised (de Boer *et al.* 2015).
- Intention-to-treat analysis. The use of this analysis suggests a particular outcome based either on a calculated clinical trajectory (in terms of improvement, regression, or the plateauing of the outcome measures) or on the average changes seen in the remainder of the group to which the participant belongs. It may be that neither of these scenarios adequately reflects the actual clinical outcomes that the participant would have attained which thus biases the outcome of the study (Gupta 2011).
- Dropouts. Some study designs do not have adequate data to compute an intention-to-treat analysis, thus resulting in dropouts. These dropout participants and their clinical status often go unreported, which has the potential to bias conclusions as it is usually the more clinically severe RA participants who opt to step out of a study in favour of personalised care. This leaves the relatively clinically less severe participants making the call as to whether an intervention is indeed appropriate (Heritier, Gebski and Keech 2003).
- Minimal clinically important difference. While commonly used outcome measures are usually able to detect small changes which may be statistically significant, this amount of change may not necessarily be clinically significant or attain the MCID (Bahreini *et al.* 2020).
- Parametric statistical procedures. These rely on assumptions about not only the distribution (i.e., they assume a normal distribution) in the underlying population but also the form or parameters (i.e., means and standard deviations) of the assumed distribution. Nonparametric statistical procedures rely on no or few assumptions about the shape or parameters of the population distribution from which the sample was drawn (Hoskin 2012; Bahreini *et al.* 2020). This concurs with the "within subjects" design, where less variance is associated with individual differences, thus providing more statistical power. There is, however, a chance of "carry-over effects" such as learning / practice and fatigue effects skewing results (Bordens and Abbott 2002).

- Reporting of Studies. The details related to the reporting of studies may affect the outcomes of the RCT (Moher *et al.* 2012).

### 4.4.2 The analysis of articles:

The analyses of articles were tabulated below according to their order and study design (sections; 4.4.2.1: RCT; 4.4.2.2: nRCT; 4.4.2.3: CS/CR) with relevant table numbering for reference.

### 4.4.2.1 RCTs

Table 4.1: List of table numbers for RCT feedback and analysis.

| Tabulated feedback of reviewer data | Properties of the study, outcome, and discussion | Author/s: | Year of publication: | Title of research paper: |
|---|---|---|---|---|
| Table 4.2a | Table 4.2b | Baillet, Payraud, Niderprim, Nissen, Allenet, Francois, Grange, Casez, Juvin and Gaudin | 2009 | A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. |
| Table 4.3a | Table 4.3b | Bell, Lineker, Wilkins, Goldsmith and Badley | 1998 | A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis. |
| Table 4.4a | Table 4.4b | Breedland, van Scheppingen, Leijsma, Verheij-Jansen and van Weert | 2011 | Effects of a group-based exercise and educational program on physical performance and disease self-management in rheumatoid arthritis: a randomized controlled study. |
| Table 4.5a | Table 4.5b | Brodin, Eurenius, Jensen, Nisell, Opava and Group | 2008 | Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. |
| Table 4.6a | Table 4.6b | Buljina, Taljanovic, Avdic and Hunter | 2001 | Physical and exercise therapy for treatment of the rheumatoid hand. |
| Table 4.7a | Table 4.7b | Cima, Barone, Porto and de Abreu | 2013 | Strengthening exercises to improve hand strength and functionality in rheumatoid arthritis with hand deformities: a randomized, controlled trial. |
| Table 4.8a | Table 4.8b | de Jong, Munneke, Lems, Zwinderman, Kroon, Pauwels, Jansen, Ronday, Dijkmans and Breedveld | 2004 | Slowing of bone loss in patients with rheumatoid arthritis by long-term high-intensity exercise: results of a randomized, controlled trial. |
| Table 4.9a | Table 4.9b | de Jong, Munneke, Zwinderman, Kroon, Jansen, Ronday, van Schaardenburg, Dijkmans, Van den Ende and Breedveld | 2003 | Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis?: Results of a randomized controlled trial. |
| Table 4.10a | Table 4.10b | Dellhag, Wollersjö and Bjelle | 1992 | Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. |
| Table 4.11a | Table 4.11b | Dhondt, Verbruggen, Oostendorp and Duquet | 1999 | Pain threshold in patients with rheumatoid arthritis and effect of manual oscillations. |
| Table 4.12a | Table 4.12b | Durcan, Wilson and Cunnane | 2014 | The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. |

Table 4.1: List of table numbers for RCT feedback and analysis (*cont.)*.

| Tabulated feedback of reviewer data | Properties of the study, outcome, and discussion | Author/s: | Year of publication: | Title of research paper: |
|---|---|---|---|---|
| Table 4.13a | Table 4.13b | Eversden, Maggs, Nightingale and Jobanputra | 2007 | A pragmatic randomised controlled trial of hydrotherapy and land exercises on overall well-being and quality of life in rheumatoid arthritis. |
| Table 4.14a | Table 4.14b | Garner, Fenton, Martin, Creaser, Johns and Barnabe | 2018 | Personalized diet and exercise recommendations in early rheumatoid arthritis: A feasibility trial. |
| Table 4.15a | Table 4.15b | Häkkinen, Mälkiä, Häkkinen, Jäppinen, Laitinen and Hannonen | 1997 | Effects of detraining subsequent to strength training on neuromuscular function in patients with inflammatory arthritis. |
| Table 4.16a | Table 4.16b | Häkkinen, Sokka, Kotaniemi and Hannonen | 2001 | A randomized two-year study of the effects of dynamic strength training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis. |
| Table 4.17a | Table 4.17b | Häkkinen, Sokka, Kautiainen, Kotaniemi and Hannonen | 2004 | Sustained maintenance of exercise induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5 year follow up. |
| Table 4.18a | Table 4.18b | Hamilton, Bywaters and Please | 1959 | A controlled trial of various forms of physiotherapy in arthritis. |
| Table 4.19a | Table 4.19b | Hansen, Hansen, Langgaard and Rasmussen | 1993 | Long term physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. |
| Table 4.20a | Table 4.20b | Harkcom, Lampman, Banwell and Castor | 1985 | Therapeutic value of graded aerobic exercise training in rheumatoid arthritis. |
| Table 4.21a | Table 4.21b | Hoenig, Groff, Pratt, Goldberg and Franck | 1993 | A randomized controlled trial of home exercise on the rheumatoid hand. |
| Table 4.22a | Table 4.22b | Levitsky, Kisten, Lind, Nordström, Hultholm, Lyander, Hammelin, Gentline, Giannakou and Faustini | 2019 | Joint mobilization of the hands of patients with rheumatoid arthritis: results from an assessor-blinded, randomized crossover trial. |
| Table 4.23a | Table 4.23b | Lineker, Bell, Wilkins and Badley | 2001 | Improvements following short term home based physical therapy are maintained at one year in people with moderate to severe rheumatoid arthritis. |
| Table 4.24a | Table 4.24b | Lyngberg, Harreby, Bentzen, Frost and Danneskiold-Samsøe | 1994 | Elderly rheumatoid arthritis patients on steroid treatment tolerate physical training without an increase in disease activity. |
| Table 4.25a | Table 4.25b | Mannerkorpi and Bjelle | 1994 | Evaluation of a home training programme to improve shoulder function in rheumatoid arthritis patients. |
| Table 4.26a | Table 4.26b | Manning | 2013 | Exercise-based upper limb rehabilitation in rheumatoid arthritis. |
| Table 4.27a | Table 4.27b | McManus, Visker, and Cox | 2015 | Effect of an arthritis foundation exercise program on sleep quality/sleep disturbance in seniors with rheumatoid arthritis: A pilot study. |
| Table 4.28a | Table 4.28b | McMeeken, Stillman, Story, Kent and Smith | 1999 | The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis. |
| Table 4.29a | Table 4.29b | Meeus, Hermans, Ickmans, Struyf, Van Cauwenbergh, Bronckaerts, De Clerck, Moorken, Hans and Groseman | 2015 | Endogenous pain modulation in response to exercise in patients with rheumatoid arthritis, patients with chronic fatigue syndrome and comorbid fibromyalgia, and healthy controls: a double-blind randomized controlled trial. |

Table 4.1: List of table numbers for RCT feedback and analysis (*cont.)*

| Tabulated feedback of reviewer data | Properties of the study, outcome, and discussion | Author/s: | Year of publication: | Title of research paper: |
|---|---|---|---|---|
| Table 4.30a | Table 4.30b | Metin and Ozdemir | 2016 | The effects of aromatherapy massage and reflexology on pain and fatigue in patients with rheumatoid arthritis: a randomized controlled trial. |
| Table 4.31a | Table 4.31b | O'Brien, Jones, Mullis, Mulherin and Dziedzic | 2006 | Conservative hand therapy treatments in rheumatoid arthritis — a randomized controlled trial. |
| Table 4.32a | Table 4.32b | Piva, Khoja, Toledo, Chester-Wasko, Fitzgerald, Goodpaster, Smith and Delitto | 2019 | Neuromuscular electrical stimulation compared to volitional exercise for improving muscle function in rheumatoid arthritis: A randomized pilot study. |
| Table 4.33a | Table 4.33b | Romanowski, Špiritović, Romanowski and Straburzyńska-Lupa | 2020 | Manual Therapy (Post isometric Relaxation and Joint Mobilization) in Knee Pain and Function Experienced by Patients with Rheumatoid Arthritis: A Randomized Clinical Pilot Study. |
| Table 4.34a | Table 4.34c | Scholten, Brodowicz, Graninger, Gardavsky, Pils, Pesau, Eggl-Tyl, Wanivenhaus and Zielinski | 1999 | Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. |
| Table 4.35a | Table 4.35b | Shinde and Varadharajulu | 2017 | Effect of therapeutic exercise programme in adults with early rheumatoid arthritis. |
| Table 4.36a | Table 4.36b | Strasser, Leeb, Strehblow, Schobersberger, Haber and Cauza | 2011 | The effects of strength and endurance training in patients with rheumatoid arthritis. |
| Table 4.37a | Table 4.37b | Van den Ende, Breedveld, Le Cessie, Dijkmans, De Mug, and Hazes | 2000 | Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial. |
| Table 4.38a | Table 4.38b | Van den Ende, Hazes, le Cessie, Mulder, Belfor, Breedveld and Dijkmans | 1996 | Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomised clinical trial. |
| Table 4.39a | Table 4.39b | Vljeland, Zwinderman, Vandenbroucke, Breedveld and Hazes | 1996 | A randomized clinical trial of in-patient multidisciplinary treatment versus routine out-patient care in active rheumatoid arthritis. |
| Table 4.40a | Table 4.40b | Westby, Wade, Rangno and Berkowitz | 2000 | A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. |
| Table 4.41a | Table 4.41b | Williams, Williamson, Heine, Nichols, Glover, Dritsaki, Adams, Dosanjh, Underwood and Rahman | 2015 | Strengthening And stretching for Rheumatoid Arthritis of the Hand (SARAH). A randomised controlled trial and economic evaluation. |

Table 4.2a: Table 4.2a Tabulated feedback of reviewer data

| Author/s | Baillet, A., Payraud, E., Niderprim, V.-A., Nissen, M. J., Allenet, B., Francois, P., Grange, L., Casez, P., Juvin, R. and Gaudin, P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2009 | | | | | |
| Title | A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | No | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention-to-treat" | Yes | No | Yes | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 9/11 | 8/11 | 9/11 | 9/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.2b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Baillet A., Payraud E., Niderprim V., Nissen M.J., Allenet B., Francois P., Grange L., Casez P., Juvin R., and Gaudin P. |
| Year | 2009 |
| Title | A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| HAQ (primary outcome), AIMS2-SF, Nottingham Health Profile (NHP), Sequential Occupational Dexterity Assessment (SODA), Duruoz Hand Index (DHI), Simple Narrowing Erosion Score (SENS), DAS28.<br>Roeder manipulative aptitude test<br>Five-handle position grip test<br>Grip and pinch strength test | Baseline, month one, six and 12 (Except SENS = baseline and month 12 and questionnaires through email every 3 months)<br>Baseline for confounding control (dexterity). | 12-months | Sample recruited of 50 randomised into 25 per group with 2 controls dropping out immediately after randomisation. Thus, a control of 23 and dynamic exercise programmes (DEP) of 25. participated in intervention/s and 80% of control and all DEP completed. | Yes | Yes (Conventional joint rehab group) | Yes | 9/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | Recruitment was carried out in various rheumatology departments of the Grenoble University Teaching Hospital in France. This raises the perception that the sample consisted of patients who had previously sought specialist care, implying that their RA disease activity and severity may be greater than in the case of the average RA patient (Wells *et al.* 2012). Treatment novelty and participant naivety, factors accompanying prior exposure to a range of interventions that could affect the outcomes of this study (Mouton 2006), also have to be considered. Participants were required to have diagnosed RA (ACR 1987) (Arnett *et al.* 1988) and were included on the basis of the following demographic and disease characteristics: age, sex, RA disease duration, morning stiffness, professional status, education level, RA global functional ACR Index, DAS28, VAS, functional status (HAQ), quality of life (NHP and AIMS2-SF), dexterity and aerobic fitness (DHI, SODA, home bicycle) and concurrent treatment (Biotherapy (anti-THF-alpha or IL-1 beta), glucocorticoid and SENS)). They were also required to be on DMARD therapy and have ECGs completed, with at risk patients being screened by a cardiologist. Excluded patients were those who had >10mg glucocorticoid therapy per day, no or unstable DMARD therapy, disease activity (DAS28 score >1.2) in the three months prior to the study, who were aged under 18 or over 70 and / or had a global functional status in RA class 3 or 4 (Hochberg *et al*. 1992). Patients unable to implement an educational programme or complete questionnaires (e.g., due to cognitive impairment) were also excluded. The final exclusion factor was the inability to perform aerobic exercise or complete the one-year follow-up because of health problems or socio-professional status. Even with the extensive considerations, such a recruitment strategy offers a limited patient description, as several factors were not considered (such as serology) (England *et al.* 2019; Myasoedova *et al.* 2020) . But to the strength of the study exclusion of participants unable to perform necessary components of the design protects against missing data to some degree but simultaneously limits the generalisability (Kendall 2003).<br><br>Randomisation of the participants entering the study was completed by an independent physician. Participants were randomly assigned to their groups, DEP or conventional joint rehab, by sealed and opaque envelopes. All other study observers were unaware of study group allocation. These processes had the potential to limit bias, thus improving the rigour of the study. The fact that biological measures of baseline (FBC, ESR, CRP, RF, serum glutamic oxaloacetic transaminase, serum glutamate pyruvate transaminase, creatine, serum calcium and ALP) were suggested to be similar between the groups further enhanced the rigour of the study. According to power analysis a sample size of 38 patients in the DEP group and 76 patients in the control group was required to detect an HAQ difference of 0.25 between groups (90% power; significance level of 0.05). This implies the study was underpowered from the start as the authors reveal that they could not attain the required number or account for patient dropout (Kendall 2003; Christley 2010) . The outcomes of the study can thus only suggest trends and not define causality effectively, which mitigates against the limitation of bias. Even though the study was well designed in terms of patient entry it is therefore compromised in terms of the outcomes and statistical analysis.<br><br>The intervention for the DEP (Pollock *et al.* 1998) involved a five-hour session per day for a four-week period. This intervention strategy placed considerable variance on the individual interventions received by the participants, with the initial focus on knowledge and physical capacity followed in turn by an occupational therapy intervention for RA, then occupational therapy focussing on skill exercise and daily activities with increasing intensity by way of endurance and exercises against resistance and lastly occupational therapy exercises directed towards specific office tasks. The above ran concurrently with a training programme designed to improve muscle strength, endurance, flexibility, and balance, which for the most part was based on the individual evaluation of each patient and modified according to pain and fatigue. By contrast, the conventional joint rehabilitation was formulated through |

| | |
|---|---|
| | a multidisciplinary programme with about 20 hours of intervention over three days with a focus on patient education followed by a specific intervention for each patient. Given these very different protocols, there some control was achieved by educating participants and requiring them to keep an exercise diary. Nevertheless, the study showed only limited ability as regards being able to record and control for variables linked to the different intervention strategies. Interventions were detailed sufficiently for the programme/s to be understoods but replicated, thus limiting their generalisability into the clinical setting. However, it should be noted that the duration of the DEP programme was not necessarily sufficient in length to manifest physiological change, which raises concerns over the outcomes reported (Heine *et al.* 2012).<br><br>All outcomes (except SENS) were normally distributed prior to analysis and the baseline characteristics of the two groups were similar except for RA Global Functional ACR Index. These differences may result from the lack of baseline information about physical activity or rehabilitation. Additionally, not all confounding factors (particularly RA disease-related) were considered and therefore equality of distribution is unknown. Finally, although good compliance is reported, its calculation may overestimate bias. These limitations to the statistical analysis are further compounded by the small sample size and the impact of the intent-to-treat analysis; both of which are known influencers of increased bias. Thus, although the variation of 0.6 points in the DAS28 effect on this outcome is considered clinically significant (Soubrier and Dougados 2005) , these outcomes are tainted by the bias (and risk of type II error) introduced in the baseline data as well as the abnormally distributed data and the impacts of sample size on the statistical analysis (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>The validity and reliability of the outcomes were variable in reporting; HAQ (Guillemin, Braincon and Pourel 1991); AIMS2-SF (Guillemin *et al.* 1997); NHP (Bucquet, Condon and Ritchie 1990); SODA (van Lankveld *et al.* 1996); DHI (Duruöz *et al.* 1996); SENS (Van der Heijde *et al.* 1999); DAS28 (Smolen *et al.* 1995); aerobic fitness (Astrand 1960). This variability introduces a risk of error in measurement and interpretation (Coster 2013; Bahreini *et al.* 2020). To the authors' credit, they also state that some of the assessment tools had limitations and may have impacted on outcomes due to the assessment tools having less sensitivity for specific outcomes. |
| Outcome | This study provided a four-week DEP intervention programme which was compared to conventional rehabilitation, using a sample of 50 RA patients randomly allocated to DEP group or a conventional rehab group. As a result of these interventions the study reported evidence of improvement of HAQ in the DEP group at all points of the study, by comparison with the conventional rehab programme group. At one-month significant difference was reported ($p$=0.04) but at six and twelve months this significant difference was lost. Significant improvement in the DE group as opposed to the conventional care group was reported in NHP ($p$=0.01), together with aerobic fitness ($p$=0.02) at the one-month assessment, although this was not maintained at following assessments. Insignificant trends of improvement were reported in DHI, DAS28, SODA, and AIMS2-SF. As a result, the study concluded that evidence had been provided supporting the DEP programmes capacity to improve QoL and impart positive impact on the participants aerobic capacity and functional status, however these improvements were not necessarily maintained (Baillet *et al.* 2009). |
| Discussion | To conclude, according to the authors, the outcomes of this study were favourable, even though the sample size was small (underpowered) and would have required the use of non-parametric statistics, thus only allowing for trends to be observed. This limitation, along with the limitations noted above (viz., the limited patient description, the tailored/individualised components of the intervention, the lack of a natural history comparison group, the limited detail regarding the intervention programme and the risk of limited exposure due to short duration of the intervention, the difference in the baseline values of some outcome measures, the limited sensitivity of some of the outcome measures and the lack of therapist blinding), provides limited protection from confounding and bias. Collectively, these limitations negatively affect the study's ability to intervention groups accurately and fairly. Additionally, failure to report aspects of the methodology required for the analysis of control and the elimination of confounding and bias is limited. Finally, the vague participant description prevents generalisability and implementation into clinical practice as the outcomes provide only limited evidence for the intervention against "usual care", let alone against natural history.<br><br>This study needs to be repeated with a larger sample (meeting the planned power-analysis) and with a clearer description of participants (to be acquired by a more inclusive set of recruitment criteria) in order to provide a basis from which to determine any true causal link present in the study outcomes and to allow application in the clinical setting to be made with a prognostic capacity. Nevertheless, through clarity in several areas of the methodology the study provides some strength to the reported findings, hence the findings offer a basis for trends and for additional research to build on. |
| Conclusion: | To conclude, the limitations noted in the table above result in a moderate degree of bias from an external validity vantage point. The review of the study by the reviewers (94% agreement) at a 9/11 ranking indicates that the study presented outcomes that were subject to a low level of bias as regards methodological rigor / internal validity.<br><br>Therefore, the study provided moderate evidence for support of the tested interventions (primarily owing to compromised external validity).<br><br>Consequently, the study needs to be repeated with a larger sample (meeting the planned power-analysis) and with a clearer description of participants (imparted by a more inclusive set of recruitment criteria) in order to provide a basis from which to determine any true causal link present in the study outcomes so that application in the clinical setting can be made with a prognostic capacity. Despite all this, the study provides moderate strength of evidence, though its findings should be implemented with caution due to the external validity limitations. |

Table 4.3a: Tabulated feedback of reviewer data

| Author/s | Bell, M. J., Lineker, S.C., Wilkins, A.L., Goldsmith, C.H. and Badley E.M. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1998 | | | | | |
| Title | A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes, | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | Yes | Yes | No | Yes | 67% |
| 6 | There was blinding of all therapists who administered the therapy | Yes | Yes | Yes | Yes | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 10/11 | 11/11 | 10/11 | 11/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.3b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Bell, M. J., Lineker, S.C., Wilkins, A.L., Goldsmith, C.H. and Badley E.M. |
| Year | 1998 |
| Title | A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Self-Efficacy Scale (SES) (primary outcome).<br>ACREU Rheumatoid Arthritis Knowledge Questionnaire (KQ)<br>Pain VAS<br>Disease activity (morning stiffness, grip strength, and joint count). | Baseline, week-six, and week-12. | 12-weeks: part-one (baseline to week 6), and part-two (week 7 to week 12). | Total sample 150 (experimental group of 76 and control group of 74). Completion 127 (experimental group of 69 and control group of 58). | Yes | Yes, no-treatment waiting list (for first arm of study – week six to 12) | Yes | 11/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | Rheumatologist referral to a "Consultation and Therapy Service" (CTS) for physical therapy, in Ontario (Canada), resulted in the acceptance of participants onto this study. This centre provides therapy to arthritis patients. The specialist nature of this centre, raises the concern that the sample may have more severe RA compared to the general RA population (Wells *et al.* 2012). Additionally, there is concern about treatment novelty and patient naivety to treatment options that were included in this study, thus affecting their outcomes.<br><br>However, it may be considered that the carefully detailed and selected set of inclusion controlled for bias in this regard. For inclusion the RA of participants had to meet the ACR 1987 criteria (Arnett *et al.* 1988), they had to have been diagnosed after 18 years of age, they were required to present with (> three) identified areas for improvement, (> six) painful or tender joints, (> 45) minutes of morning stiffness and comply with a full the functional classification of class II or III (Steinbrocker, Traeger and Batterman 1949). Additionally, the participant had to have been referred to the CTS for the first time, be literate (English) [capable of understanding the informed consent] and be present for (> four) consults or (> three) hours of physical therapy at CTS. Exclusion criteria included; if the patient had previously participated the feasibility study (prior to this study), required urgent care, and / or reported involvement in similar programmes. This may have provided protection for patient naivety and potentially treatment novelty (Mouton 2006). However, the limited patient description and exclusion of RA dynamic characteristics (e.g., serology and medication specifications), allowed for variation in the type of participant RA. This implies a possible lack of homogeneity between the groups. An example of this is that although medication use was detailed at the outset for DMARDs and methotrexate, it overlooked steroids and NSAIDs (potentially affecting study outcomes (Zhao *et al.* 2017; Heidari *et al.* 2019), particularly as the majority of the outcomes are subjective in nature and report on pain or ADLs (Katz 2002). Additionally, medication changes were allowed during the course of the study. Finally, in terms of homogeneity, stratification was only limited by to: (1) stable or unstable medication, (2) disease duration of (≤ two) years versus (> two) years and Steinbrocker's functional classification II or III. These stratifications do not take care of the majority of RA nuances that may account for baseline differences in the groups. This lack of homogeneity can lead to increased risk of type II error / conclusions (de Boer *et al.* 2015) and the effects of incorrectly chosen stratification parameters may actually perpetuate flaws in the conclusions (Zhao *et al.* 2017; Heidari *et al.* 2019). This lack of homogeneity and ability to identify the type of RA patient under care in this study limits generalisability (by affecting the external validity required for application in clinical practice).<br><br>On a positive note, however, concealed allocation was reported, thus limiting patient assessor and researcher bias (Kendall 2003; Moustgaard *et al.* 2014). And it was also documented that there were no significant differences between groups in terms of age, RA disease duration, sex, significant partner status, education, income class, functional class, stable DMARD therapy and medication for arthritis taken daily. Similarly, in the study's favour, was that the sample size was determined by an *a priori* analysis based on the need for a 10% change in the SES (Bell *et al.* 1994) and equalled or exceeded the minimum 58 per group. Additional participants were intentionally incorporated as a means of drop-out control and intention-to-treat analysis was included. The sample size enabled parametric statistics analysis, allowing for stronger conclusions and the negation of the placebo effect (McGowan 1997). Analysis included the stratification variables that were controlled with the ANOVA analysis, which is a downside of the study as the reason for the choice of these variables has not been justified and may result in statistical errors and effect outcomes (de Boer *et al.* 2015).<br><br>With regard to comparing the results between the group receiving intervention and the no-intervention group / control (wait-listed for the first six weeks), which means that at face value during the first portion of the study (baseline to six-weeks), allowed for the determination of efficacy of intervention programmes verses no intervention. This is mitigated against in that the discussion revealed that the control group participants received community rehabilitation during in this time. Thus, the outcomes are a relative effectiveness outcome. Additionally, this "added" treatment in the waitlisted group compounds interaction effects: touch effects (Melzack and Wall 1965; Moayedi and Davis 2013; Senderovich *et al.* 2016), placebo effects (Kaptchuk 1998; Bialosky *et al.* 2011) and the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015). |

| | |
|---|---|
| | To add complexity, care to both the intervention group as well as the control group was performed by multiple therapists who did not seem to have protocol consistency training, where not compared for consistency and were not reported to have been trained similarly. Thus, although the intervention was standardised (and included education and exercise) (not clearly delineated for replication)) allowed therapist and patient individualisation. This thus seems more like a set of case series than comparison of two homogenous groups each with a specific intervention. This along with variable patient naivety are important as most of the outcomes are subjective (pain reporting and ADLs), which means that the outcomes can be a measure of the impact of these considerations by comparison with the actual treatment effect of the intervention.<br><br>With regard to the outcome measurements, the SES, KQ, VAS, morning stiffness duration, grip strength and joint count were all similar at baseline between the groups, which aid in determining change due to the intervention and not differences between the groups. There was some concern around the relationship between the intervention and measurement points and thus the actual outcomes that were being measured. This determine outcomes as immediate, short, intermediate or long term (Platt, Schisterman and Cole 2009). In this study, the measurement may be a better reading of the intervention immediately preceding the measurement as opposed to the six- or 12-week programme measures. Furthermore, the measures are only representative of short-term effect and not long-term outcomes. This is important, as RA is a chronic disorder, which means that the disease duration or chronicity mitigates against shorter treatment protocols being effective. The actual intervention (i.e., not the wait listing) provides a further complication as different modalities show clinical impact over different clinical durations (Fraser 2008). This can be explained in that exercise requires six to eight weeks before functional capacity change can be measured as compared to collagen and fibrous tissue, which requires 12 weeks or more (Fraser 2008). This conundrum is further compounded by the fact that the actual interventions were not itemised, so actual physiological impact and thus the transferability to clinical practice is impossible.<br><br>Given the conundrums of the inclusion criteria and the lack of an intervention description in addition to treatment variables and inappropriate duration for the interventions makes the replication of the study impossible to achieve. The strongest positive indicator for the study was the use of blinded assessors who were trained through a protocol or programme of assessment with reliable and valid RA measures (Huskisson 1974; Lorig *et al.* 1989; Bell *et al.* 1994; Lineker *et al.* 1997). The only downside is the fact that there were 18 evaluators. It would have been better to have these numbers limited to one or two trained assessors (Kendall 2003). |
| Outcome | The SES at six-weeks reported in the completers showed an improvement in the experimental group of 13.5% was 7.7% greater than the control, this between group difference was statistically significant ($p$=0.015). However, through intention-to-treat analysis a complete data set showed that the SES outcomes were reported to be affected by the control group receiving treatment; additionally, the inclusion of the intention-to-treat analysis further reduced the improvement differences between the groups for the SES (at six ($p$=0.106) and 12 weeks p=0.543). Significant change was also noted in the KQ and morning stiffness at six weeks ($p$=0.011 and $p$=0.036 respectively), the experimental group maintained these improvements at 12 weeks but did not differ significantly from the control at 12 weeks ($p$=0.687 and $p$=0.510 respectively). The VAS, grip strength and tender joint count were not significantly different at any time point. Nevertheless, the authors conclude that the four visits of education and exercise in a community based physical therapy intervention should be utilised as an adjunct to medical care to significantly improve self-efficacy, disease management, knowledge and morning stiffness (Bell 1998). |
| Discussion | The outcomes were correctly reported to be affected by the treatment received by the control group; additionally, the inclusion of the intention-to-treat computations for drop- out patients also reduced the significance of improvement (protocol violations are more likely in severe RA patients). This seems to suggest that the reporting of the outcomes seems to be reflective of patients with less severe RA and that the analysis was not reflective of all the participants in the study. This therefore also limits the ability to interpret the results for use in clinical practice. Although the use of more severe RA patients is commendable as this increases the chances of change, this however requires increased time for the ability to detect change. For example, the grip strength was only affected by four visits over six weeks, which is too little for this outcome measure to detect effect change (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012). It is also possible that the uncontrolled factors such as the duration and RA severity of the various participants (de Boer *et al.* 2015), the use of multiple therapists and multiple assessors (even though blinded) and the lack of medication control during the study may have contributed to the somewhat unremarkable outcomes. Therefore, it is not possible to conclusively state that the outcomes attained through this study are a result of the intervention programme alone. In addition, the outcomes as suggested by the authors are only applicable for the short term (six-week period), which may be as a result of the placebo, the Hawthorne effect and patient expectation based on prior experience (Kendall 2003; Moustgaard *et al.* 2014) as opposed to the intervention. |
| Conclusion: | In the light of the above, the conclusions drawn in this study provide limited evidence for the intervention and therefore need to be accepted with extreme caution and applied with circumspection in clinical practice as the study's external validity was compromised (partially acknowledged as such by the authors), even though the internal validity was strong (94% agreement), indicating the internal validity and methodology rigour of the study (viz. risk of bias (internal validity)).<br><br>Future studies need to consider with a larger sample (meeting the planned power-analysis) with a clearer description of participants (imparted by a more inclusive set of inclusion criteria), control of confounders like medication, limit additional interventions used (particularly in the control group) during the study period (or a natural history control), and increasing the duration and exposure to the intervention. Follow- up measures need to be structured to determine immediate, short, intermediate, and long-term effects. This would determine true causal links present in the study outcomes and that application in the clinical setting can be made with a prognostic capacity. |

Table 4.4a: Tabulated feedback of reviewer data

| Author/s | Breedland, I., van Scheppingen, C., Leijsma, M., Verheij-Jansen, N. P. and van Weert, E. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2011 | | | | | |
| Title | Effects of a Group-Based Exercise and Educational Program on Physical Performance and Disease Self-Management in Rheumatoid Arthritis: A Randomized Controlled Study. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | No | No | No | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 8/11 | 8/11 | 8/11 | |
| Overall percentage agreement | | | | | | 100% |

Table 4.4b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Breedland, I., van Scheppingen, C., Leijsma, M., Verheij-Jansen, N. P. and van Weert, E. |
| Year | 2011 |
| Title | Effects of a Group-Based Exercise and Educational Program on Physical Performance and Disease Self-Management in Rheumatoid Arthritis: A Randomized Controlled Study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Aerobic capacity (VO2 max)<br>Muscle strength (Dynamometer)<br>Self-reported health status (Dutch-AIMS2)<br>Self-efficacy (ASES) | Baseline (week one)<br>Measurement two (week nine) end of intervention.<br>Measurement three (week 22) 13 weeks after intervention follow-up. | Eight weeks intervention period with a follow-up 13 weeks after the intervention period. Total study time is 22 weeks. | 34 participants were randomised, with 19 in the intervention group, and 15 in the control group. Drop-out of two from the intervention group. | Yes, single blinding | Yes, a waiting list control. | Yes, an independent researcher achieved this by drawing sealed tickets. | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 100% |

| | |
|---|---|
| Limitations | Recruitment was carried out by means of a convenience sampling through referral of independent RA patients by rheumatologists. This study was conducted at the Rehabilitation Centre of at the Medical School of Groningen University, Groningen (Netherlands). This raises the perception that the sample consists of patients who have previously sought specialist care, implying that their RA disease activity and severity may be greater than in the case of the average RA patient (Wells *et al.* 2012). And in this raises the possibilities that a lack of patient naivety and treatment novelty exist in this sample (Wells *et al.* 2012). The inclusion criteria were extensive and attempted identify ambulant, independent RA patients based diagnosed by the ACR 1987 Criteria (Arnett *et al.* 1988); with difficulties in physical performance and / or issues with daily management of their condition, and participants had to be between the ages of 18 and 66. Participants were excluded if disease activity was severe (DAS28 score over 5.1) - in this controlling for disease activity (DAS28) (Prevoo *et al.* 1995; Fransen, Creemers and Van Riel 2004; Fransen and Van Riel 2006; van Riel and Fransen 2007), comorbidities and required a functional capacity based on the Steinbrocker classification III or higher (Steinbrocker, Traeger and Batterman 1949). These criteria did not consider control of RA serology phenotype, disease duration, use of physical modalities and intervention contraindications. Given these criteria and the use of appropriate randomization and concealed allocation it is not unexpected that the authors reported no significant difference in participants' characteristics in the baseline characteristics and distribution. Although stratification of patients was by age and sex, the authors did not seem to control either for medication (Zhao *et al.* 2017; Heidari *et al.* 2019) although they acknowledge its possible impact, or for the functional capacity of the patients and the chronicity of the RA, both of which would potentially have impacted the outcomes of the study as well as its generalisability (de Boer *et al.* 2015). Additionally, the placebo effect, which requires an *a priori* sample at minimum to be negated, the Hawthorne effect (Richardson 2007) in terms of intervention blinding (Kendall 2003; Moustgaard *et al.* 2014), patient naivety to the interventions (Mouton 2006) and environmental factors (Richardson 2007) affecting the outcome are not acknowledged in the study.<br><br>Based on the process reflected in the CONSORT flow diagram, the final sample size required the use of non-parametric statistics; these only have the ability to show trends regarding the effectiveness of the intervention. The power analysis was not *a priori* but rather post-hoc, there was no drop-out replacement plan (the authors report that the single referral practitioner did not refer sufficient patient numbers within a given time period, and they proceeded so as not to have the patients wait for treatment for an excessive period of time) and intention-to-treat was not implemented. Therefore, the effect of sample size was only calculated after the trial, which prevented the possibility of ensuring an adequate sample in order to draw definitive conclusions in the study's favour, however, the lack of intention-to-treat limited the statistical bias that could further have compounded the outcomes (Fogg and Gross 2000; Kendall 2003; Christley 2010). Lack of baseline comparability of some outcome measures questions the ability of the study to use these as comparable outcome measures. These should have been utilised as part of the stratification prior to statistical analysis correction after the fact (de Boer *et al.* 2015). It was also noted that some outcomes measures had fewer than 85% of the total population in the respective groups complete all measurement points. This incomplete data may also have negatively affected the outcomes of the study (Ibrahim, Chu and Chen 2012), therefore justifying the exclusion of patients with RA affecting the lower extremity and potentially compromising the cycle ergometer outcomes (e.g., $VO_2$max), should have occurred at the beginning of the study.<br><br>In favour of the study, the outcomes assessor was blinded to the group allocation and the measures utilised included both validated and reliable subjective and objective outcomes (Bell *et al.* 1992). These procedures decrease assessor bias, limit patient learning as a significant effect on subjective reporting and allow the results to exclude the patient effects not addressed in the baseline when objective and subjective outcomes are compared (Richardson 2007; Moustgaard *et al.* 2014). The relationship between intervention and outcome measurements allowed for short-term and longer-term reporting on the effectiveness of the protocol. The use of eight weeks of intervention would have been sufficient for muscular changes to be recorded and thus more accurately document the intervention effect (Fraser 2008; Heine *et al.* 2012).<br><br>The treatment interventions are clearly described and could be replicated. However, the intervention programme does not allow the outcomes to comment on specific aspects of the programme, in addition each participant could have had a different programme as the recreational aspect was customisable. Consequently, the lack of structure to the intervention allowed for variance in exposure. In other words, was the whole protocol completed in one day or more spread out and were the exercises done on specific days. |

| | |
|---|---|
| Outcome | The statistical outcome was affected by four dropouts from the intervention group and two dropouts from the waitlisted group. The authors report that there were no significant between group changes other than the $VO_2$max. Therefore, the conclusion drawn by the authors was that RA patients who participated in a multidisciplinary intervention using the FIT programme protocol compared to the waitlisted participants had significant benefit only with regard to improved aerobic capacity but not with any of the other outcome measures although there were suggestions of trends for improvement of muscle strength, perceived physical health and self-efficacy of the participants (Breedland *et al.* 2011). |
| Discussion | In the context of the limitations, as relating to the baseline similarities and differences between the two groups as well as the dropouts from each, it is difficult to exclude bias as a factor affecting the outcomes of the study, particularly when the intervention group lost four patients whose data may have limited the significance of the difference for the $VO_{2max}$ measures. Additionally, the relatively short period for muscle changes secondary to the intervention results in the negligible difference between the groups (Bell 1998; Heine *et al.* 2012). These issues collectively bring the objective outcomes into question.<br><br>As a result, the outcome of the study hinges on the subjective outcome measures that may have been influence by the placebo effect, the Hawthorne effect, patient naivety and environmental factors. Therefore, the collective outcome of the study does not make it possible to identify the actual treatment effect on RA patients. The results may also have compromised by the intervention group being younger and having less severe RA than the wait-listed group, even though they had had RA for a longer period than the waitlisted group, but this cannot be determined as the effect on these mean data for the dropouts is not included in the reported data.<br><br>Secondly, the inability of the study to recognise specific interventions (i.e., the programme versus individual modalities) as having played a role in the clinical outcomes of the study means that it is not possible to identify a specific intervention modality as the active ingredient in the attained outcomes. |
| Conclusion: | To conclude, the limitations noted in the table above result in a high level of bias from an external validity vantage point, while the review of the internal validity of the study by the reviewers (100% agreement) at 8/11 ranking indicate a low risk of bias in respect of the study's internal validity and a high level of methodological rigor. The study thus presents with a moderate degree of bias that is controlled for by internal validity but compromised in the sphere of external validity. At best, therefore, this study provides low to moderate evidence for a programme of multidisciplinary group therapy incorporating mixed exercise and thus for its clinical use on RA patients.<br><br>The study should be re-done with a larger sample (determined by priori power analysis) recruited from a larger and multi-centred source; a clearer description of participants should be acquired by means of a set of recruitment criteria addressing noted limitations in the sample's generalisability to the RA presentations included in this study. The methodology should include not only strategies to enhance adherence to the intervention, such as using an exercise diary to record compliance, but also a means by which the diagnosis can be confirmed. Such a study would provide a basis to determine the true causal link and thus the application in the clinical setting with prognostic capacity. |

Table 4.5a: Tabulated feedback of reviewer data

| Author/s | Brodin, N., Eurenius, E., Jensen, I., Nisell, R., Opava, C. H. and Group, P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2008 | | | | | |
| Title | Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | No | Yes | 67% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | Yes | No | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | No | Yes | Yes | 67% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | No | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 7/11 | 6/10 | 7/11 | |
| Overall percentage agreement | | | | | | 88% |

Table 4.5b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Brodin, N., Eurenius, E., Jensen, I., Nisell, R., Opava, C. H. and Group, P. |
| Year | 2008 |
| Title | Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| **Primary outcome**<br>EuroQol VAS (perceived health status)<br>**Secondary outcome**<br>Muscle function (Grippit (grip strength) and Timed-Stands test (lower extremity function)<br>Escola Paulista de Medicina (EPM) ROM scale<br>Walking in a figure-of-8 (balance)<br>VAS (pain)<br>HAQ disability index (activity limitation)<br>Self-reported questionnaire<br>DAS-28<br>Medication prescription | Pre/post design | 12 months | 228 total sample, with the intervention group (n=94) and control (n=134). Of these 82% and 85% completing the study respectively. | Yes | Yes, all participants in both groups had access to, but were not specifically encouraged to participate in, ordinary physical therapy treatment. | Yes,<br>Roll of dice | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 88% |

| | |
|---|---|
| Limitations | The study followed a multi-centred approach, with participants drawn from 10 specialist clinics (rheumatology), located in seven university and three country hospitals. All of which served the central and rural catchment area of Sweden. Over a period of four years patients were invited after their first anniversary on the country's RA registry. These enrolments were onsite and adjusted to the site's circumstances. The sourcing provides the study a degree of protection for greater representation of the general RA population; on the other hand, sites are solely medical which increases the risk of more severe RA (not generalizable) (Wells *et al.* 2012). However, neither recruitment method offered protection for patient naivety nor treatment novelty. The study provided little information on inclusion criteria (serology, functional class, activity of disease, age, flare-ups, pain, other physical modalities in use, contraindications, and co-morbidities). In essence, provided that the participants were older than 18 years of age, were diagnosed with RA according to the ACR 1987 criteria (Arnett *et al.* 1988), could communicate in Swedish and perform body-function testing they were included in the study. This does offer the study control for missing data in relation to inability to perform outcome measures. However, it opens the study up to confounding for uncontrolled variables (e.g., RA dynamic characteristics).<br><br>The study attempted to control for this (outcome and confounders) with assessment of disease activity through standard methods described by the EULAR handbook (Scott, Vanriel and Vandeputte 1992) – further details are not clear. In addition, physical activity (by means of participant self-reporting), body function and health status (such as muscle function, balance, and ROM). The study stated purpose behind the lack of exclusion through disease activity and comorbidities, as the limited generalisability of the outcomes and all can benefit from exercise. Following this, subjects were randomly allocated, the baseline characteristics (demographic data, disease activity and percentages of participants taking different types of medications) were reported as comparable (pre- and post-dropout), but it was noted that there was a significant difference in duration of RA since diagnosis with intervention resulting in a lesser duration. Notwithstanding this, it was noted that the collective participants compared well with the RA register in Sweden except as regards age and disease duration. Although the recruitment was by convenience sampling (which in itself risks heterogeneity of the sample (Jager, Putnick and Bornstein 2017)), allocation was randomised (roll of a dice) without stratification, and there was no mention of concealed allocation, which could unduly have biased the group allocations (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; Wells *et al.* 2012; Smith *et al.* 2013; de Boer *et al.* 2015). Nevertheless, groups comprised of 94 in the intervention group and 134 in the control. This uneven distribution of the sample creates comparability concerns. Sample size was determined by an *a priori* analysis (determined as 91 participants per group). Unfortunately, due to dropout the number of participants in the intervention group (n=77) dropped below the required group number and so the intention-to-treat strategy was implemented. As a result, although the study was able to perform parametric statistics (offering the ability to draw definitive conclusions) (Fogg and Gross 2000; Kendall 2003; Christley 2010), these outcomes may have been affected by the way the intention-to-treat was implemented (Fogg and Gross 2000; de Boer *et al.* 2015; Unbiased Research 2017). On the other hand, this study is least affected by the placebo effect (as it reached the required sample size to reduce the impact of this particular factor). However, the Hawthorne effect and environmental factors (Richardson 2007) were not controlled for.<br><br>The control group was not a true control seeing that the group was able to receive any treatment and was not specifically encouraged to participate in ordinary physical therapy treatment, while the treatment group received a complicated intervention containing multiple factors for 12-months. Despite this, the intervention for both groups were not clearly defined and it implied that the intervention group participants received individualised intervention plans that were documented at intervals in an activity log. The nature of the intervention, the lack of control in the |

| | |
|---|---|
| | control group and the convenience sampling from a hospital setting along with duration of disease and centre of treatment reduced the chance of naivety. Additionally, previous experience may have influenced the patients' response to the treatment, particularly as the majority of outcomes are subjective patient reporting (Mouton 2006). These factors made it difficult to ascertain whether the study as designed actually measured the effect of the intervention as these confounders detract significantly from this ability. Another factor to be considered with regard to the implementation of a uniform intervention is uniformity of standards among the 23 personal coaches, the majority of whom did not receive training. It was implied in the article that at least one coach from each of 10 clinics was trained, but that all coaches received training manuals and forms. This, however, is not the structure of preference for the best outcomes in a RCT where the control of extraneous variables is important (Kendall 2003).<br><br>Outcomes were measured by trained, independent and blind assessor/s, with validated and reliable outcome measures (Coster 2013; Bahreini *et al.* 2020). Additionally, the good mixture of subjective and objective measures provided a strong base for the study (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). This was supported by the baseline comparable outcomes except for VAS, grip strength and balance. However, significant difference in objective measures at baseline limits the ability of these measures to control for the influencers that may affect subjective measures such as the placebo effect, the Hawthorne effect and learned responses (Kaptchuk 1998; Bordens and Abbott 2002; Mouton 2006; Richardson 2007; Bialosky *et al.* 2011; Sedgwick and Greenwood 2015). The outcomes are further influenced by the consistency of measurement and type of measurement employed as some participants may have had telephonic and others in-person measurements. In addition, the intervention group seemed to have measurement intervals more frequently than that of the control (Richardson 2007). This makes it possible that the learning effect may have enhanced the outcomes of the intervention group while the control group could have been affected negatively through memory decay (Mouton 2006; Ricker, Vergauwe and Cowan 2016). The conclusions can only comment on short- and intermediate-term outcomes and not long-term outcomes, particularly as it is unclear as to whether the last measurement was taken directly after the last treatment (thus measuring the immediate effect of treatment as opposed to measuring the collective effect of the treatment plan) (Pater *et al.* 1998).<br><br>Despite reporting in the study that participants in each group did not significantly differ in percentage of medication types utilised, apart from the baseline data was not made available for the reader to evaluate medication control in terms of interruption, continuation, changes made during study, adherence, and dosage. The specific medication type in this study, among them pain medications (NSAIDs and analgesics), DMARDs and other biologics apart from anti-TNF-alpha are reported on, however specific details of these are lacking. It also failed to report on administration method, its pharmacokinetic effect and therapeutic concentration as well as the control of these effects and how it influences the condition and findings (see section 2.7.1) (Zhao *et al.* 2017; Heidari *et al.* 2019). |
| Outcome | The study reported a retention rate of 82% and 85% in the intervention and control groups respectively, over the one-year study period. With this completion rate, the conclusions drawn indicate that a one-year programme for improved physical activity resulted in perceived health status (EuroQol VAS $p$=0.025) and increased muscle strength (timed stands test; $p$=0.000 and Grippit; $p$=0.003). Nevertheless, the requirements for healthy physical activity remained similar ($p$>0.05), and insignificant change were reported for EPM range of motion ($p$>0.05), DAS28 ($p$>0.05), HAQ DI ($p$>0.05) and figure 8 oversteps ($p$>0.05) (Brodin *et al.* 2008). |
| Discussion | This study, which is loosely based on the CONSORT principles and requirements, appears to be fairly well structured (Begg *et al.* 1996; Kendall 2003; Moher *et al.* 2012). Its limitations are noted in the table above. It is not clear if the results of the study are affected by the floor and ceiling effects (especially in terms of the objective outcome measures) (Liang 2000; Coster 2013), therefore the significant improvement in muscle strength (grip strength) may be subject to a type II error (Bartley and Fillingim 2013; de Boer *et al.* 2015). This is further compounded by the lack of clarity of the characteristics of the RA patients included (including the effect of medications on the patients) and therefore the generalisability (external validity is compromised) of the results of the study is brought into question. Additionally, the lack of clarity regarding the intervention and its multimodal nature does not allow for comment on any one treatment intervention, but rather on the treatment programme.<br><br>In the light of the above the results need to be interpreted with caution and applied with discretion in clinical practice given that there is a lack of clarity regarding the relationship between cause and effect in terms of the intervention and the outcomes, which has been influenced by a significant individualisation of the programme to the patients, thus making the interaction a collection of case series as opposed to an actual review of the treatment protocol. |
| Conclusion: | When considering the evaluation of internal validity and the methodological rigor of this study, reviewers reported with a high degree of congruency (88%), that the study was at moderate risk of bias in terms of the internal validity while carrying moderate strength in the methodological rigour. Taken into consideration with the limitations noted in this table, the external validity is at moderate to high risk of reporting outcomes imparted by factors other than intervention. Furthermore, the lack of clarity with regard to the intervention used and the broad patient description (although intentional as representing the large range in the clinical presentation of RA in patients) result in moderate evidence at best for the use of the intervention.<br><br>Finally, the lack of intervention detailing hinders a re-testing of the intervention with greater consideration for patient description, improved randomisation and drop-out prevision, and a natural history comparison group, which would provide the necessary basis for evaluating the true effect of the intervention and its prognostic value, both needed for its confident implementation into RA management. |

Table 4.6a: Tabulated feedback of reviewer data

| Author/s | Buljina, A. I., Taljanovic, M. S., Avdic, D. M. and Hunter, T. B. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2001 | | | | | |
| Title | Physical and exercise therapy for treatment of the rheumatoid hand. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | No | Yes | Yes | 67% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 7/11 | 7/11 | 8/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.6b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Buljina, A. I., Taljanovic, M. S., Avdic, D. M. and Hunter, T. B. |
| Year | 2001 |
| Title | Physical and exercise therapy for treatment of the rheumatoid hand. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - ESR<br>- Joint size<br>- Hand articular index<br>- ROM<br>- ADL<br>- Patients self-assessment of pain in the hand<br><br>Hand strength:<br>- Grip strength<br>- Tip-to-tip pinch<br>- Palmar pinch<br>- Key pinch | Pre/Post design. | Three-weeks | 100 in total, with 50 per group. | The study utilised two assessors one of which was blinded. | Yes, the control was a waiting list control. | Yes, by means of table of random numbers. | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study's recruitment was through the University of Sarajevo (Bosnia and Herzegovina) clinical centre. Through whom 115 consecutive RA patient had been referred to the rehabilitation hospital. Given the referral from a hospital centre, the impression provided is that the patients were recruited from the subpopulation of RA that are seeking medical care and are likely to be suffer from more severe RA (Wells *et al.* 2012). As a result, risk is imparted to the generalisability of the study findings, and it is unlikely the sample is naïve to the intervention type. The study implemented an inclusion criteria for this study that covered age (20 to 70 years of age), diagnostic criteria used (ACR 1987 criteria) (Arnett *et al.* 1988), a disease duration of at least six months, disease activity factors (three or more swollen joints in both hands, five or more tender joints in both hands, raised ESR, and a hand problem defined as decreased ROM and grip strength). However, functional class, pain severity, use of other physical modalities, contraindications, comorbidities, and dominant-handedness were not considered. No differences between the groups ($p$>0.01) were shown with regard to the characteristics of sex, age, RA duration, serology and anatomy by means of the Steinbrocker criteria (Steinbrocker, Traeger and Batterman 1949). To the strength of the study, it controlled for medication well (three months prior to enrolment, with comparable use between groups in terms of patient numbers using NSAIDs, slow-acting anti-rheumatic drugs and small dose of oral systemic steroids). However, the study does not report on the following: interruption or change to the medication prescribed at entry to the study and during the study, adherence/compliance and comparability of dosage and type and the means to control effect of medication on the outcome of the study (Zhao *et al.* 2017; Heidari *et al.* 2019). In view of the above points, the groups, although homogenous in the reported findings, do not account for all variables that could introduce bias in the outcomes that are meant to test the interventions. Stratification of the variables unaccounted for within the reported randomization procedures such as a table of random numbers would have strengthened the study. The sample size (of 100 with 50 respectively per group) was determined without an *a priori* calculation, which prevents the reader from determining the power of the study. Furthermore, dropout numbers and the replacement of dropout data points were not clearly detailed, and no intention-to-treat analysis was stipulated. With the large sample size, the study had the ability to utilise parametric statistics, which would have allowed the study to draw more definitive conclusions (Kendall 2003; Christley 2010). However, the impact of dropouts on the baseline similarities of the groups at the conclusion of the study is unclear, thus the impact of the variables noted in the paragraph above becomes more problematic in presenting bias towards a particular group.<br><br>Measurement tools used in this study are a mix of subjective and objective, which is conducive to a stronger study (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). As noted in paragraph one above, the objectives measures can control for factors that may influence the subjective measures. Additionally, the objective measures may negate the effects of assessor blinding (which was partially employed) (Kendall 2003; Moustgaard *et al.* 2014), the placebo effect, the Hawthorne effect and patient naivety (Kaptchuk 1998; Mouton 2006; Richardson 2007; Bialosky *et al.* 2011; Sedgwick and Greenwood 2015). In support of a stronger study outcome, the study controlled for time of the day at which measurements were taken so as to exclude the effect of diurnal variation (Nandgude, Hasabe and Kolsure 2018). This is unlike other studies reviewed and significantly strengthens the study's outcome. Finally, the primary outcome and measurement tool were suited and validated for RA (Coster 2013; Bahreini *et al.* 2020). What is unclear and significantly affects these gains is the lack of a clear *a priori* calculation to determine whether the sample size was likely to detect differences in the outcome measures (Kendall 2003; Mouton 2006; Christley 2010). In congruence with this, it is also unclear when the post measurement was taken. In other words, if the post reading was taken on the day of the last intervention, this would imply measurement of the effect of the last intervention as opposed to the treatment programme. Thus, the impact on the outcomes of these latter two points raises the question to whether the outcomes do in fact measure the treatment programme effect (Pater *et al.* 1998). This is further compounded |

| | |
|---|---|
| | by the fact that exercise takes six to eight weeks at minimum to show physiological change in clinical practice, therefore the length of time that this study ran mitigates against the possibility that exercise-induced muscle strength measures would have been able to detect any clinical change reflected in the HAQ and ADLs (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012). Statistically the baseline comparison between outcomes measures was similar for all measured variables (ESR, joint size, Ritchie hand articular index, ROM, ADL scale, pain scale, grip strength, tip-to-tip pinch, palmer pinch, key pinch*): p*>0.01, which allowed a more accurate comparison between the groups and therefore a more accurate assessment of the effect of the intervention programme based on these measures (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>In terms of the interventions, the control was a no-treatment group and the technique used in the experimental group was clearly described, which provides better insight into the applicability of the intervention in clinical practice (external validity) (Heine *et al.* 2012). In addition, the clear no-treatment group enables the study to determine short-term effectiveness of the intervention programme more accurately. Given the "busy-ness" of the treatment group, no comment can be made on any single intervention. Furthermore, because RA is a chronic condition the shorter-term study can be affected in that there is limited short-term change, which may negate the ability of the study to report any positive clinical gains. The chronicity of RA can negate the effects of short-term study interventions and limit the effects detected and reported in the study outcomes (Fraser 2008). |
| Outcome | The authors concluded that a tailored physical therapy and exercise therapy programme that can address improvement in both subjective (pain score ($p$<0.005)) and objective (articular hand index (in terms of tenderness ($p$<0.005)), ROM ($p$<0.01), activities of daily living ($p$<0.005)) outcome measures in the short-term. While the outcomes of ESR and PIP joint size had improved with clinical significance but did not reach significance statistically. The control however, reflected in the outcome data, no significant changes for all outcome measures. In fact, slight deterioration was noted over the study period. Thus, concluding that the favourable short terms effects on hand function were noted as a result of the three-week physical therapy programme (including modalities and exercise) (Buljina *et al.* 2001). |
| Discussion | Within the context of the author-reported study outcomes, it is possible that, with the regression of the control group over time and the improvement of the intervention group, the statistically significant differences between the groups may be an artefact of the widening of the gap between the groups' respective outcomes and not specifically the improvement of the outcomes for the intervention group only. This could have been controlled for if the patients had been adequately stratified and there had been control of the variables that were not reported at the stage of stratified randomisation (de Boer *et al.* 2015). Additionally, bias was introduced through the lack of clarity with regard to assessment of the measures as well as the patient-related factors (see Section 4.4.1.1).<br><br>In view of the lack of an intervention programme that only included one intervention, it is not possible to comment on anything more than the full intervention programme with caution in terms of its applicability in clinical practice (external validity). Therefore, with the impact of several limitations listed in the table above, on the external validity and the effect this has on the ability to extrapolate to clinical practice. |
| Conclusion: | Given the limitations noted above, from the external validity vantage point this study potentially suffers from a moderate degree of bias and the reader cannot exclude the effect of multiple external influences. This to a degree contrast with the review by the reviewers, which ranks the study as an 8/11 with 94% agreement, indicating that there is high methodological rigour and low risk of bias to the internal validity of the study in terms of the PEDro scale. This highlights a degree of disjuncture between the two reviews (internal and external validity). It is however possible to have a rigorous study that has influencers outside of the research process that impact the well-structured study and interfere with the outcomes. Therefore, this study at the best has a moderate level of bias that is shown through both review processes, which indicates that there is a moderate degree of evidence and that this should be considered when applying these results in practice and with caution. |

Table 4.7a: Tabulated feedback of reviewer data

| Author/s | Cima S.R., Barone A., Porto J.M. and de Abreu D.C.C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2013 | | | | | |
| Title | Strengthening exercises to improve hand strength and functionality in rheumatoid arthritis with hand deformities: a randomized, controlled trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | Yes | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | Yes | Yes | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | No | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 6/10 | 6/10 | 7/10 | 6/10 | |
| Overall percentage agreement | | | | | | 91% |

Table 4.7b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Cima S.R., Barone A., Porto J.M. and de Abreu D.C.C. |
| Year | 2013 |
| Title | Strengthening exercises to improve hand strength and functionality in rheumatoid arthritis with hand deformities: a randomized, controlled trial. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| HAQ<br><br>Hand grip strength (bilateral)<br><br>Pinch grip strength (four types and bilateral) | Both at baseline then for treatment group (after session 10 and 20) and control group (after two months) | Two-months, as stated by inconsistent with reported treatment plan | Treatment group (n=13) Control group (n=7)<br><br>Completers Treatment group (n=10) Control group (n=7) | No | Yes | Yes | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 91% |

| | |
|---|---|
| Limitations | This study fails to explicitly state the location of the study, nor the source of recruitment. However, the principal author is based at the school of medicine university Sao Paulo (Brazil), furthermore, the treatment protocol was established at the rehabilitation centre of the same school. The failure to explicitly state the type and source of recruitment, as a result the reader is limited in their ability to determine influence this may have on the type of sample draw and the likelihood of participant naivety (Wells *et al.* 2012). Because of inclusion criteria of the study, the sample consisted of 20 female participants with RA (criteria used unreported), with associated hand deformities (classification criteria unreported) but noted deformities included swan-neck finger/s, Z-shaped thumb, and Boutonniere finger/s and ulnar deviation was necessary in both hands. Participants were excluded if any other RA-related diseases were present or if they were in a phase of exacerbation of RA. The study randomised patients by means of a computer-generation. This to the study's credit assumes degree of concealment and blinding for the allocation of participants to groups (Mouton 2006). These broad criteria increase the chances for a lack of homogeneity between the groups. This is particularly pertinent as the study's lack of demographic data does not allow homogeneity to be assessed and neither does it allow generalisability of the results. Areas that increased the potential for the risk of bias and confounding, include medication (Zhao *et al.* 2017; Heidari *et al.* 2019), occupation and occupational status (Naqvi, Hassali and Aftab 2019), surgical interventions (Naqvi, Hassali and Aftab 2019), use of other interventions (Mouton 2006), hand dominance (Lee, Gandevia and Carroll 2009; Dragert and Zehr 2013), contraindication to intervention (Durcan, Wilson and Cunnane 2014), lifestyle factors (Pedersen *et al.* 2006a; McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019), endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018), and other general demographic factors (Bartley and Fillingim 2013).<br><br>The study failed to provide a power-analysis, implying that the study is underpowered (Mouton 2006; Christley 2010). The small sample size was further compromised by drop-outs (Fogg and Gross 2000) and lack of utilising intention-to-treat analysis (de Boer *et al.* 2015). The lack of demographic data for the baseline and for those that dropped-out, bias cannot be assessed due to differences between completers and drop-outs (viz the drop out may have aggravated during the study and thus the results only reflect the relatively better RA patients (de Boer *et al.* 2015).<br><br>The study implemented a programme of hand exercises aimed at strengthening the hand (treatment group), the control received no intervention. On a positive note, both groups were asked to maintain habitual activities (e.g., exercise) to eliminate the effect on the outcomes. If the participants had changed anything, this would confound outcome measures (Robins, Hernán and Siebert 2004). The treatment group received 20 supervised sessions, these sessions were conducted twice a week (thus 10 weeks) and home-based exercises (completed three times a week). This limits the blinding of the multiple therapists, and increases risk of bias (Moustgaard *et al.* 2014). The intervention group completed. These exercises were carried out twice a week for two months (totally 20 sessions). Load progression was determined individually for the participant and those with orthotic aids were encouraged to utilise them when exercising, which typically makes this a case series compared to a no intervention control. However even in this context, the baseline values for the outcome measures, handgrip strength for both hands, pinch strength for both hands and HAQ values were not significantly different between groups (de Boer *et al.* 2015).<br><br>Outcome measures utilised in this study include a balance of subjective and objective tools (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). The HAQ was administered to all participants in the same day, which controlled time-sensitive confounding (Platt, Schisterman and Cole 2009), but the oversight in time of day (the diurnal RA effects) is not reported (Nandgude, Hasabe and Kolsure 2018). It was noted that the HAQ is not specific for RA of the hands and the authors note this limitation. The procedures for both grip strength measures were well-detailed, collectively controlling for measurement/protocol consistency (Kendall 2003). Unfortunately, all assessments were conducted |

| | |
|---|---|
| | by a single unblinded assessor (evaluator) providing control over inter-assessor variability despite the bias associated with lack of blinding (Kendall 2003; Moustgaard *et al.* 2014). Assessments appear to occur at baseline and post-intervention for both groups (pre/post design) with an additional assessment after 10 sessions for the intervention group only.<br><br>Given that the study was underpowered, the use of nonparametric statistics, the analysis of the time and intergroup effects at best show preliminary trends (Hoskin 2012; Bahreini *et al.* 2020). Additionally, conclusions need to be read with caution as the probably of type II errors are high considering that many baseline demographic impact factors could not be controlled for in the analysis. |
| Outcome | This study found in RA patients with hand deformities (ulnar deviation with one or more of swan-neck/Boutonniere finger/z-shaped thumb), who received 20 sessions of hand exercises performed at home and in supervised sessions, twice a week; that the intervention group showed significant improvement over time of hand grip strength and pinch grip strength (reporting a $p < 0.05$). Furthermore, the study reported improvements in functionality ($p = 0.016$). For the control group, no significant differences were reported over time. Thus, the study concluded that implementation of this strengthening exercise programme in RA participants with hand deformities allows gain in handgrip strength, pinch grip strength, and functional improvement. The study notes that significant functional and strength improvement was apparent in the hand after 10 sessions, but 20 sessions provided further benefit (Cima *et al.* 2013). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though the sample size was small (study underpowered) and would have required the use of non-parametric statistics, thus only allowing for trends to be observed. This limitation, along with the limitations noted above (lack of assessor and therapist blinding; limited control imposed by the inclusion criteria both on group homogeneity and statistical analysis, suggests that the outcomes may have been affected by factors such as natural history of the condition / chance / time and be the result of a regression to the mean given the reporting of outcomes. These limitations collectively negatively affect the study's ability to accurately and fairly compare intervention group to the control without risking bias in outcomes and conclusions. Additionally, failure to report aspects of methodology required for analysis of control and elimination of confounding and bias variables limits the reader's capacity to evaluate the degree of risk in the study. Finally, the vague participant description prevents generalisability and implementation into clinical practice as the outcomes provide limited evidence of efficacy and therefore need to be interpreted with caution. This study needs to be repeated with a larger sample (determined by power-analysis) with a clearer description of participants, in order to provide a basis from which to determine any true causal link present in the study outcomes and that application in the clinical setting can be made with a prognostic capacity. |
| Conclusion: | As a result of the limitations noted in the table above, resulting in high bias from an external validity vantage point as well as the review of the study by the reviewers (91% agreement) at a 6/11 ranking; indicates that the study presented outcomes that were subject to a moderate to highly compromised methodological rigour. Therefore, this study does not provide any evidence for support of the tested interventions.<br><br>The study does fail to provide sufficient and clear patient description so as to facilitate the ability of field practitioners to identify which RA patients would benefit most, limiting the clarity of the finding's generalisability. Additionally, when reviewing the outcomes, which reported statistically significant change over time in the intervention group, in the context of being underpowered, indicates that only trends may be reported not causation. Thus, the suggestion that a hand exercise programme is better than a control is limited in providing evidence for the intervention. At best therefore, it is concluded that this study provides limited to moderate evidence for hand exercise programme/s and thus for its clinical use in RA patients.<br><br>The study should be re-done with a larger sample (determined by power-analysis) with a clearer description of included RA presentations. This would provide a basis to determine the true causal link and application in the clinical setting with prognostic capacity. |

Table 4.8a: Tabulated feedback of reviewer data

| Author/s | de Jong, Z., Munneke, M., Lems, W. F., Zwinderman, A. H., Kroon, H. M., Pauwels, E. K., Jansen, A., Ronday, K. H., Dijkmans, B. A. and Breedveld, F. C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2004 | | | | | |
| Title | Slowing of bone loss in patients with rheumatoid arthritis by long-term high-intensity exercise: results of a randomized, controlled trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 9/11 | 9/11 | 9/11 | 9/11 | |
| Overall percentage agreement | | | | | | 100% |

Table 4.8b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | de Jong, Z., Munneke, M., Lems, W. F., Zwinderman, A. H., Kroon, H. M., Pauwels, E. K., Jansen, A., Ronday, K. H., Dijkmans, B. A. and Breedveld, F. C. |
| Year | 2004 |
| Title | Slowing of bone loss in patients with rheumatoid arthritis by long-term high-intensity exercise: results of a randomized, controlled trial. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| **Disease activity:** DAS28 (joint count for swollen and tender joints, ESR, VAS [patients' global assessment of disease activity measures])<br>**Physical capacity:** Aerobic fitness (ergometer), Muscle strength (isokinetic dynamometer), HAQ<br>**Radiological:** Bone mineral density (BMD), and Joint damage | Disease activity and physical capacity measures were conducted at baseline and months 0, 3, 6, 9, 12, 15, 18, 21, 24.<br>HAQ measures were conducted at months 0, 6, 12, 18, 24.<br>Radiological measures were conducted at months 0, 12, 24. | 24-months | 309 in total started the study (151 RAPIT and 158 usual care).<br>On completion 281 competed the study (136 RAPIT group and 145 usual care group).<br>With 15 dropouts from the RAPIT group and 13 dropouts from the usual care group. | Yes | Yes, a usual care group treated by a physical therapist only if deemed necessary by their attending physician. | Yes, by means of random digit generator | 9/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 100% |

| | |
|---|---|
| Limitations | The study draws its sample via cross-referencing the inclusion criteria with the disease related and demographic details drawn from the patient files in the register of four rheumatology outpatient clinics in the Netherlands. Those found eligible were contacted and invited to join the study. Of those that showed interest screening by trained investigator was undertaken, according to the set of recruitment criteria. The recruitment method carries with it the risk of excluding sections of the population and in doing so provide a limited generalisability of the outcomes to a subpopulation of the true RA population (Wells *et al.* 2012). Furthermore, the likelihood of a naïve patient or novel intervention is in question as they all have a history specialist care. The study outlined inclusion criteria that detailed the diagnostic criteria ACR 1987 (Arnett *et al.* 1988) and ACR functional class (I to III) (Hochberg *et al.* 1992), age range (20 - 70 years), stable use of DMARDs (three months prior to the study) and ability to cycle on an exercise bicycle. The study excluded comorbidities that contraindicate high-intensity exercise (Cardiac / pulmonary disease), those with weight bearing joint prostheses, and any comorbidity likely to reduce life expectancy. In much the same manner as previous study (de Jong *et al.* 2003), the criteria lack consideration for RA serology, duration and activity of disease, pain and physical modality use. Given the limited inclusion criteria, randomization (random digit generator) and stratification for study by centre, age and sex, the study does achieve similar baseline characteristics between the groups, with no significant differences. The exceptions were that the disease duration in the RA Patient In Training (RAPIT) programme group was shorter than that of participants in usual care, and the RAPIT participants had less damage of the small joints. What is not clear due to lack of demographic data for the groups at baseline (as a standalone publication) from the study is whether the participants in the study are representative of the general RA population as would be seen in clinical practice. Thus, the external validity is unclear, particularly with regards to disease activity and other comorbidities that may have excluded participants, thus making the sample population different from the general RA population. Unlike the several studies in the reviewed (e.g., Bell (1998), Häkkinen *et al.* (2001), and Piva *et al.* (2019)), this study monitored the method of administration and /or changes in medication in view of the need to control for changes allowed during the treatment period of the study as well as for patient adherence, dosage, and type of DMARDs. The limitation, however, was that there was no use of a medication diary, instead the patients were asked to report medication changes every three months for the preceding three-months. This approach is subject to memory decay (Mouton 2006). This study determines the sample size by an *a priori* analysis, which produced the minimum sample size (n=96) [while the study utilises a sample of 150 each compensating for dropout and a completing sample of intervention (n=145); control (n=136)] (Fogg and Gross 2000; Kendall 2003; Christley 2010). In addition, the study also controlled for dropouts and applied intention-to-treat analysis (de Boer *et al.* 2015; Unbiased Research 2017). Consequently, the sample size allowed for the use of parametric statistics, enabling the study to show causal links between intervention and outcome (Bordens and Abbott 2002; Hoskin 2012; Bahreini *et al.* 2020). From a statistical vantage point, this study utilised effect correction for baseline differences in the statistical analysis prior to comparison computations as reported by de Boer *et al.* (2015), this imparts bias, which detracts from the study's ability to draw conclusions about the intervention.<br>The intervention technique used was clearly described for the intervention group. However, during the study the intervention group, much the same as the control group, was also able to receive changes to medication (e.g., bisphosphates, vitamin D and calcium supplements) and other interventions with the exception of a high-intensity |

| | |
|---|---|
| | weight-bearing exercise programme. By contrast, the control group had what is defined in the study as usual care, applied with a physical therapist treating only if it was deemed necessary. Thus, the study could only determine the relative effectiveness of the RAPIT programme as a (usual care supplemented with therapy programme) to the usual care alone. It appears the principal outcome was BMD changes over the two years. The fact that this principal outcome is objective in nature implies that the outcomes are potentially reflected with limited bias having influenced them. The secondary outcomes offer a balance of subjective and objective outcomes reported in this study (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), collectively well-established outcome tools in RA research for validity and reliability (Coster 2013; Bahreini *et al.* 2020). All of which to the credit of the authors, are evaluated and recorded by four trained blinded assessors, (Kendall 2003; Moustgaard *et al.* 2014). Given the "usual care" approach in both groups, the relationship between treatment interventions and measurements is not entirely clear although measures were taken at regular defined intervals (as noted in the study properties section of this table). The intervention group seemed more likely to have a structured intervention, but this would not have been the case in the control. The question therefore arises as to whether bone mineral density changes due to the intervention are being tested or whether the effect of frequency of care influences BMD. This question is compounded by the study's participant attendance / participation rate; where the median percentage of sessions attended was 74% (IQR 55%, 75%), averaged over two-years. Thirty percent of all participants had a sufficient attendance rate (defined as 50 – 75% attendance) and 49% has a high attendance rate (defined as 75 – 100% attendance). In support of the study and perhaps negating the limitations of participation rates, it is noted that exercise as an intervention requires six to eight weeks to show impacts on functional capacity and at least three months for change in fascial and collagen structures it is potentially plausible that the duration of two-years exercise intervention would manifest effect on BMD, to a level that would allow for the detection of effect (Fraser 2008). |
| Outcome | Conclusions indicate "long-term high-intensity weight-bearing exercise programme" in addition to the studies defined usual care as opposed to this usual care alone for RA is effective in slowing the loss of BMD. This was measured principally by hip and lumbar spine BMD, imparted by the programmes emphasising muscle strength and aerobic fitness. With regarding to the hip, the initial changes were a negative (decrease) for the usual care group, with the RAPIT group reporting stable BMD at one-year, followed by a decrease for both groups at two-years. However, comparing these groups showed that the intervention group measures had a reflected significantly slower rate of demineralisation over two-years ($p$=0.026). The lumbar spine had an increase in BMD at the one-year assessment with stability maintained at two-years and no difference between the groups ($p$=0.738). The change in hip BMD was significantly and independently associated with changes in both muscle strength (multivariate odds ratio [OR = 1.75, 95% CI; 1.07 – 2.86) and aerobic fitness (OR = 1.79, 95% CI; 1.10 – 2.90), but not with the attendance rate (OR = 1.00, 95% CI; 0.99 – 1.00) (de Jong *et al.* 2004). |
| Discussion | As has been noted in the table above, the study falls prey to several limitations that must be considered when reflecting on the conclusions the authors have drawn from the study. These limitations risk misattributing the findings solely to the intervention programme. The RAPIT programme participants had access to usual care, as a result the paradigm of the study is shifted from determining the efficacy of the RAPIT programme to comparing usual care to the intervention programme in which usual care is supplemented by the RAPIT programme. This means that the outcomes can only talk of a particular approach as being more favourable, and not specific or individual interventions within each of the approaches. This comparison is, however, also limited by the effect of a lack of controlling for several demographic and baseline factors, in addition to controlling for this statistically (de Boer *et al.* 2015); but also by the limited external validity that does not allow for effective identification of when these approaches would be best applied in clinical practice. The applicability to clinical practice may also be limited by the lack of information on how the clinical measures were chosen in order to run analysis to determine the OR relationships between these measures and the changes in BMD (de Boer *et al.* 2015). |
| Conclusion: | As noted above, the external validity of the study is compromised by several limitations, deriving from baseline difference between groups (in disease duration and articular damage), failure to assess comorbidity distribution (among other RA dynamics) and their effect on the outcomes. These collectively raise concerns about confounding and bias effects on the outcomes reported. The study's external validity suffers a moderate degree of bias. However, the outcome of the review by the external reviewers with total congruency / agreement reflects a PEDro score of 9/11, indicating that from the internal validity vantage point the study's methodological rigour is of high quality and the validity is subject to low risk of bias. Like other studies reviewed this study presents conclusions in the two aspects of the analysis that reflect disjuncture between the external and internal validity risk analysis, indicating that the bias this study is subject to may derive from an external source to that examined within the internal validity PEDro criteria. The collective would indicate that evidence provided by this study should be viewed as subject to moderate to low bias risk. Additionally, the study's choice of control renders the evidence in this study as evidence of relative effectiveness between two treatment arms. Fundamentally, the interventions share a baseline in the form of the usual care, which is supplemented in the experimental group with the RAPIT intervention. Given the comparison of these two interventions and the study's failure to cite evidence of efficacy in the study for the usual care, the review has no clear basis from which the intervention can be rated. Therefore, the study at best provides moderate evidence for a relative effect, while the efficacy of either intervention remains unclear. The evidence base for clinical application is thus low to moderate, if not low, and the findings of this study should be utilised in RA management with caution. Furthermore, the identification of the RA patient/s most suited to such an intervention is restricted by the study's limited patient description. Given the amendable short-falls in the study, it is recommended that the study be re-done with more inclusive set of recruitment criteria and a more detailed patient description, as well as the employment of a no-treatment control (no physical therapy) for a shorter period against both treatment arms to provide some form of comparative basis from which greater context can be drawn for this study. |

Table 4.9a: Tabulated feedback of reviewer data

| Author/s | de Jong, Z., Munneke, M., Zwinderman, A. H., Kroon, H. M., Jansen, A., Ronday, K. H., van Schaardenburg, D., Dijkmans, B. A., Van den Ende, C. H. and Breedveld, F. C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2003 | | | | | |
| Title | Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis?: Results of a randomized controlled trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | Yes | Yes | Yes | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 9/11 | 9/11 | 9/11 | |
| Overall percentage agreement | | | | | | 97% |

Table 4.9b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | de Jong, Z., Munneke, M., Zwinderman, A. H., Kroon, H. M., Jansen, A., Ronday, K. H., van Schaardenburg, D., Dijkmans, B. A., Van den Ende, C. H. and Breedveld, F. C. |
| Year | 2003 |
| Title | Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis?: Results of a randomized controlled trial. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - McMaster Toronto Arthritis (MACTAR) Patient Preference Disability Questionnaire<br>- HAQ<br>- Aerobic fitness<br>- Muscle strength<br>- Hospital Anxiety and Depression Scale (HADS)<br>- Disease Activity Score with 4 variables (DAS4) | Measured at month -zero, -six, -12 -18 -24). | Two-years | 309 total, of these nine withdraw before baseline assessment, thus groups of 150 each. Drop out in intervention group (14) and control group (five). Therefore, completers included intervention group (136) and control group (145) | Yes | Yes, a usual care group treated by a physical therapist only if deemed necessary by their attending physician. | Yes, permutated-blocked randomization with stratification (4 blocks) | 9/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 97% |
| - Radiological damage (Larsen method) | At month -zero, -12, -24) | | | | | | | |

| | |
|---|---|
| Limitations | The study draws its sample via cross-referencing the inclusion criteria with the disease related and demographic details drawn from the patient files in the register of four rheumatology outpatient clinics in the Netherlands. Those found eligible were contacted and invited to join the study. Interested candidates were screened by a trained investigator, according to the set of inclusion criteria. This inclusion method carries with it the risk of limited generalisability of the outcomes to all RA patients (Wells *et al.* 2012). Furthermore, the likelihood of a naïve patient or novel intervention is unlikely to be avoided as all candidates have a history specialist care.<br><br>The study outlined inclusion criteria that detailed the diagnostic criteria ACR 1987 (Arnett *et al.* 1988) and ACR functional class (I to III) (Hochberg *et al.* 1992), age range (20 - 70 years), stable use of DMARDs (three months prior to the study) and ability to cycle on an exercise bicycle. In much the same manner as other studies in this review, the criteria lack consideration for RA serology, duration and activity of disease, severity of pain and use of other physical modality. Given the limited stringency of the inclusion criteria, randomization (random digit generator) and stratification for study (by centre, age, and sex), does achieve similar baseline characteristics between the groups with no significant differences. The exceptions were that the disease duration in the RA Patient in Training (Rheumatoid Arthritis Patients In Training (RAPIT)) programme group was shorter compared with that among the participants in the usual care group, and the RAPIT participants had less damage of the small joints. What is not clear from the study is whether the participants in the study are representative of the general RA population as would be seen in clinical practice. The external validity is therefore unclear, particularly regarding disease activity and a number of comorbidities that may have excluded the participants, thus making the sample population different from the general RA population.<br><br>A strength of this study rarely noted in other studies, is that this study monitored the method of administration and changes in medication (given the need to control for changes allowed during the treatment period of the study), as well as patient adherence, dosage, and type of NSAIDs and DMARDs. Control of these factors allows for a deceased impact on the outcome measures, thus allowing the study to measure the tested intervention. Another strength of this study is that the sample size was determined by an *a priori* analysis, which produced the minimum sample size (n=96). Furthermore, the study makes provision for drop-out in recruitment of 150 per group, with a completing sample of intervention (n=145); control (n=136). In addition, intention-to-treat analysis was applied. In this way, the sample size allowed for the use of parametric statistics, offering the study the ability to show causal links between intervention and outcome (Kendall 2003; Mouton 2006; Christley 2010). From a statistical vantage point, this study utilised effect correction for baseline differences in the statistical analysis prior to comparison computations. As reported by de Boer *et al.* (2015), this does impart bias, which detracts from the study's ability to draw conclusions about the intervention. It should be noted that the study reported that the completers were comparable with those of the dropouts/attendance failures in terms of the outcomes implying that the intention-to-treat analysis did not create bias in the study.<br><br>The intervention technique used was clearly described for the intervention group. However, the intervention group was also able to receive changes to medication and other interventions (apart from the prescribed high-intensity weight-bearing exercise programme) during the course of the study in much the same way as the control group. By |

| | |
|---|---|
| | contrast, the control group had usual care applied where a physical therapist treated only if it was deemed necessary. Thus, the study can only determine the relative effectiveness of the RAPIT programme of the Intervention group (RAPIT + usual care) against the usual care programme. This has implications for subjective bias as participants are unlikely to be naïve to the intervention/s and can easily identify if they are in the usual care group, in addition to which there are the variable differences in treatment exposure and its relationship to the Hawthorne effect, touch therapy and placebo effects in the control group (Melzack and Wall 1965; Kendall 2003; Moustgaard *et al.* 2014). These potentially affect the manner of reporting the subjective outcomes (Mouton 2006), though this may have been minimised by the good balance between subjective and objective outcomes reported in this study (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). However, this may be limited by the fact that the objective outcomes seemed to have limited change whereas the subjective outcomes had some significant changes. Thus, the question arises as to whether the results reflect the intervention changes, or the patient factors present in the trial.<br><br>To the credit of the authors, the outcome measures were recorded by a blinded assessor, by means of validated and reliable measures for RA (Moustgaard *et al.* 2014). Given the usual care approach in the control group, the period relationships between treatment interventions and measurements are not entirely clear (Pater *et al.* 1998). It would seem that the intervention group was more likely to have a structured approach with interval outcome measures taken, but this would not have been the case in the control. Therefore, the question arises as to whether the study was actually testing an intervention or was it merely testing the frequency of care required over a specific period in order to induce change. This may be compounded by the study's report on participant attendance / participation rate for the full study. The median percentage of sessions attended was 74% (IQR 55%, 75%), averaged over 2 years, during which time 30% of all participants had a sufficient attendance rate (defined as 50 – 75% attendance) and 49% has a high attendance rate (defined as 75 – 100% attendance). In support of the study and perhaps negating the limitations of participation rates, exercise as an intervention requires six to eight weeks to show impacts on the functional capacity and thus ADL would only show improvement after this period. This means that the study's two-year duration would allow for the detection of effect (as will be relevant in outcome measurement tools like HAQ) (Bell 1998; Breedland *et al.* 2011; Williams *et al.* 2015). |
| Outcome | A long-term high-intensity exercise programme plus a usual care programme is more effective than usual care in improving functional ability of RA patients. Intensive exercise does not increase radiographic damage of the large joints, except possibly in patients with considerable baseline damage of the large joints. After two-years, participants in the RAPIT programme and usual care showed more improvement in functional ability than participants in usual care alone. The MACTAR (trends) ($p$=0.421), aerobic fitness ($p$<0.001), muscle strength ($p$<0.001) and HADS ($p$=0.007) scores were significantly different in favour of the RAPIT programme. The median radiographic damage of the large joints did not increase in either group. In both groups, participants with considerable baseline damage showed slightly more progression in damage (de Jong *et al.* 2003). |
| Discussion | Given the above analysis of limitations and the conclusions drawn from the study, the review must consider the influence these have on the outcomes and the risk of mis-attributing the findings in fact the intervention programme. The RAPIT programme participants having had access to usual care is a prime example of the confounding in this study. For this reason, the study in fact compared usual care on its own against usual care supplemented by the RAPIT programme. This means that the outcomes can only talk to a particular approach as being more favourable and not to specific or individual interventions within each of the approaches. The study is further impaired by not statistically controlling for several demographic and baseline factors and their effects on the outcomes reported (de Boer *et al.* 2015). Additionally, the limited external validity does not allow for the effective identification of when these approaches would be best applied in clinical practice. |
| Conclusion: | The limitations noted above, together with the study's failure to establish the baseline difference between groups in disease duration and articular damage, failure to assess comorbidity distribution (among other RA dynamics) and their effect on the outcomes, inability to exclude effect of subjective bias, and the effect of variable frequency and consistency of intervention exposure over two years combine to compromise the study's external validity which thus suffers a moderate degree of bias. By contrast, the review by the external reviewers with high congruency (97% agreement) reflects a PEDro score of 9/11, which from the internal validity vantage point indicates that the methodological rigor is of high quality and the validity is subject to low risk of bias. Thus, there is disjuncture between the external and internal validity reviews, indicating that the bias this study manifests may be derived from a source external to that of a structured research protocol and design. The collective would indicate that evidence provided by this study is moderate. Aside from the above, the study suffers a significant clinical limitation in respect of a chosen control in that the findings are representative of intervention effectiveness that is relative to usual care. The baseline (which the usual care represents) to which the test intervention (RAPIT) is compared has no cited efficacy evident in the study. As a result, findings have no clear basis from which the intervention can be rated. Therefore, this study at best has a moderate degree of evidence that reflects the relative effect, while efficacy remains unclear. The evidence for clinical application should be viewed as moderate and applied in RA management with caution. Moreover, the identification of the RA patient/s most suited to such an intervention is restricted by the limited patient description. Given the amendable shortfalls in the study, it is recommended that the study be re-done with a more inclusive set of recruitment criteria and thus a more detailed patient description as well as the employment of a "no-treatment" control (no physical therapy) for a shorter period against both treatment arms to provide some form of comparative basis from which greater context can be drawn for this study. |

Table 4.10a: Tabulated feedback of reviewer data

| Author/s | Dellhag, B., Wollersjö, I. and Bjelle, A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1992 | | | | | |
| Title | Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by intention-to-treat | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | No | Yes | Yes | 67% |
| Total score | | 7/11 | 6/11 | 6/11 | 7/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.10b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Dellhag, B., Wollersjö, I. and Bjelle, A. |
| Year | 1992 |
| Title | Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| ROM (bilateral hand)<br>Grip strength (bilateral hand)<br>Grip strength (dominant hand only)<br>Pain with resisted motion (dominant hand only)<br>Pain with non-resisted motion (bilateral hand)<br>Stiffness (bilateral hand) | Measurements at pre/post two to five days before and after the intervention period. | Intervention lasted four weeks with measurements at pre/post two to five days before and after the intervention period. | Total of 65 participants.<br>13 dropped out prior to start of study.<br>52 participants took part and completed the study. Group 1 = 13; Group 2 = 11; Group 3 = 15; and Group 4 (control) = 13 | Not reported | Yes, a no intervention group, although this is not clearly stated | Yes | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study recruited participants via surveying the RA patients of the Sahlgren University Hospital (Sweden), that were seropositive. All of which had been treated via the in- or out-patient setting through the rheumatology department. The survey gathers information pertaining to age; disease duration; functional capacity (relating to ADL); and interest in physical- and / or occupational- therapy. From the returned questionnaires screening and eligibility was conducted. The context of this sampling method does raise the question, confounding variable introduction unknown to the study. As well as the risk of unintentionally limiting the generalisability of the study, through subpopulation representation (such as, limitations to adjunctive therapy accessibility, severity of disease influencing willingness to participate but as well as the convenience of currently admitted or attending patients) (Wells *et al.* 2012). Furthermore, the interest in such therapy may be influenced by the past experiences (positive or negative), or fear of the intervention. Opening the study up to assessor bias (cherry picking the positive responses) and the samples naïve, novelty, and perception bias (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013; Moustgaard *et al.* 2014). Without clarity on serological variant of RA (RF/ACAP/etc), the study's inclusion criteria outlined seropositive patients without concomitant diagnoses who were younger than 70 years of age, had a RA duration of six to 10 years, and a functional class of I or II and hand problems (decreased ROM and / or grip strength) (Steinbrocker, Traeger and Batterman 1949). These criteria did not include diagnostic criteria (given the date of publication, the reader is left to assume RA diagnostic criteria used was the ACR 1987 (Arnett *et al.* 1988)), control for activity of RA, flare-ups, use of other physical modalities, contraindications to interventions and / or comorbidities. The above factors make it difficult to identify the type of patient that was included in the study, compounded by the aforementioned source of recruitment (Wells *et al.* 2012; England *et al.* 2019; Myasoedova *et al.* 2020). It also negatively influences the ability of the study to test the intervention as the inherent differences between patients on any one or combination of factors not controlled for which can affect the outcomes for the group/s and thus negate the ability of the study to test the intervention (de Boer *et al.* 2015). Randomization was achieved by sequential allocation into four groups according to sex, age, duration of the disease and previous hand surgery. Concealed allocation was thus not applied (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Mouton 2006; Nair 2019). Stratification was not conducted in this study for any of the dynamic factors relevant to the study (inclusion criteria, baseline characteristics, and dynamic RA characteristics such as disease activity (DAS28), or comorbidities). This is important, particularly when there is a small sample size per group and the condition is multifactorial; resulting in the possibility of an outlier in any one group affecting the outcomes being tested (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). Sample size was reported to have been calculated by an *a priori* analysis, but the specifics of the calculation are not report (Kendall 2003; Christley 2010). However, the authors state that a larger sample would not have produced different results. Based on this assertion it would seem to support the use of this particularly small sample. On the other hand, it is not made clear if the power of the study was maintained after suffering dropouts (Mouton 2006; Christley 2010). The completing sample size was 52 (from the initial 65 – arguably the concern raised here is mitigated by the dropout occurring prior to the start of the intervention) participants randomised into four groups (13;11;15;13 post dropouts). Dropout control was not utilised, being hampered by the limitation of patients exclusively to those who had been at the Sahlgren University Hospital and who were seropositive and met the other inclusion criteria. Intention-to-treat was not possible as the dropout occurred before the study commenced (Gupta 2011). The impact of the small sample size prevented the use of parametric statistics and thus the outcomes need to be considered as trends and not accepted categorically (Hoskin 2012; Bahreini *et al.* 2020). The above, together with the lack of stratification, are areas of concern and could have included bias in the outcomes.<br><br>The group interventions were clearly delineated, with the control group being a "no treatment" group. The treatments options were provided three times a week for four weeks in all groups except the control group and thus the study acknowledges that a Hawthorne / observer or attention effect cannot be excluded as having influenced the outcome of the study (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015). This is not dissimilar to the placebo effect that also becomes a concern in underpowered studies (Mouton 2006; Christley 2010) and the |

| | |
|---|---|
| | effect of touch therapy that would have favoured the two wax bath groups and significantly detracted from the control group (Melzack and Wall 1965; Moayedi and Davis 2013; Senderovich *et al.* 2016). The one positive is that the study looked at controlling for environmental factors insofar as measurements were taken at the same time in the day (Pater *et al.* 1998). A concern in terms of the subjective outcome measures is that readings were taken (not that these were the majority) before and after each of the 12 interactions and then also between two and five days after the last intervention; the effect of repeated measures is to bring about a learned response (Bordens and Abbott 2002). This latter concern would therefore have been affected by a lack of patient naivety (Mouton 2006) and the possibility that lack of improvement in the control group was related to patient disgruntlement.<br><br>Regarding the statistical analysis, no statistically significant differences were reported in baseline characteristics (age, sex, functional class, or disease duration, medication, and intra-articular injections) and distribution between the four groups. The outcome measures (ROM, Grip function test, pinch function, grip strength, pain, stiffness) were also reported as not being statistically significantly different at the baseline. It was also reported that medications were not changed, and no intra-articular injections were given during the two months before or during the study, which is a consideration that other reviewed studies did not take into account (Zhao *et al.* 2017; Heidari *et al.* 2019) and therefore works in this study's favour. The small sample size and very small scores possible (the range possible for change) also limited the study's ability to detect change in these particular measurement tools (the floor and ceiling effects cannot be determined) (Liang 2000; Coster 2013). Systematic blinding of the patients, patient assessors and outcomes assessors are not clearly defined and the impact of this is not clear on the outcomes of the study. If not stated as such, it is nonetheless assumed that there is an impact on the outcomes due to bias on the side of the patient, the group allocator and / or the outcomes assessor/s (Kendall 2003; Moustgaard *et al.* 2014). This would have undermined the use of the subjective and objective measures that were reliable and valid outcomes chosen for this study (Coster 2013; Bahreini *et al.* 2020), with the exception of the Grip function test which the authors suggest is only reliable as a post-surgical measure.<br><br>In terms of outcomes, literature (Fraser 2008) indicates that exercise requires six to eight weeks to impact on functional capacity and ADL measures. This study only employed a four-week period, which would decrease the study's ability to detect change in the outcome measures related to strength measures. This could explain the lack significant change or difference attained in the majority of outcome measures (including ROM, Grip function test, pinch function, grip strength, pain and stiffness). Another possible consideration is the region of interest in the study – the hand. In some outcome measures, only the dominant hand is tested, while others tested bilaterally. The reason for this is unclear and it is not possible to determine whether the authors considered the effects of the intervention as having only unilateral or bilateral effects. Literature suggests that unilateral intervention is known to have bilateral effects (Dragert and Zehr 2013; Villafañe, Cleland and Fernandez-de-Las-Peñas 2013), which may also be the reason that the outcomes aggregated to the mean, and this did not achieve a significant difference between the groups. Thus, the lack of significant differences particularly in the primary outcome may result in an incorrect assumption that the interventions are not significantly different from each other and from no treatment at all (as reflected in the control group). The last factor that needs to be considered is that there seems to be a mismatch between the study duration of four weeks and RA chronicity (six to 10 years), which, quite aside from the small sample size, may have negated the possibility of attaining significant changes in the outcomes. |
| Outcome | After four-weeks of therapy the study reported on significant changes noted in the three intervention groups and control. The intra-group differences reflected that group one noted mobility in the nondominant hand and grip strength had significantly improved ($p$<0.05); group two noted flexion deficit improvement in the dominant hand ($p$<0.05) and in reduction of pain on non-resisted motion ($p$<0.05) as well as in stiffness ($p$<0.05). However, groups -three and -four reported no significant change over the study. The participants of groups -one and -two, reported immediate and significant improvement post-treatment in pain relief (groups -one ($p$<0.01) and -three ($p$<0.05)). Additionally, participants of all intervention groups, reported immediate and significant improvement post-treatment in stiffness ($p$<0.01). Inter-group comparison revealed that significant improvement in flexion deficit of the dominant hand ($p$<0.05) in groups one and two when compared with control, and reduction in pain with non-resisted motion ($p$<0.05) in group two when compared with groups three and control. Hand function improved, in terms of total grip function and pinch test ($p$<0.05) for group one when compared with other groups. The inter-group comparison between all groups reported that exercise (as is delivered in this study) reduce pain significantly with non-resisted motion (p<0.01). improved the participants stiffness (p<0.05), and improved flexion deficit in both hands (non-dominant ($p$<0.01) and dominant (p< 0.05). while wax bath therapy showed trends toward reduce pain with resisted motion, but this did not reach significance (Dellhag, Wollersjö and Bjelle 1992). |
| Discussion | Although the outcomes provided by the authors make logical sense in terms of the physiological effect of the interventions, it is unclear whether the bias that this study was exposed to negate the effects of attaining more accurate outcomes. This may also be the reason why the outcomes of this study agree with some previous studies and conflict with others. Therefore, given the weaknesses that this study manifests, it is difficult to accept on face value the outcomes that were obtained. This implies that the use of these interventions singly or in combination needs to be applied with caution in clinical practice and that in terms of obtaining informed consent the practitioner needs to be cautious in suggesting that these interventions will attain particular outcomes. |
| Conclusion: | Given the analysis conducted above and the caution flagged in the clinical application of the interventions based on supplied evidence, from the external validity vantage point this study suffers from a moderate degree of bias and, excluding the effect of an external influence/s, cannot be achieved by the reader. This lines up with the findings of the PEDro review by the reviewers, ranking the study at 7/11 with 94% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study. Future studies need to consider an increased sample size, improved control of blinding and due consideration of the intervention effects (unilateral versus bilateral effects) along with consideration of patient effects on the outcomes to allow for accurate reporting of the actual effects of interventions. It should also be considered that all outcomes are applicable to the patient group (i.e., post-surgical outcome measures should not be used for non-surgical patients). Thus, as a final note, this study at best provides moderate evidence and until further research (as noted prior) can be carried out, the clinical practitioner should lean on the side of caution when considering implementing these interventions into management of RA. |

Table 4.11a: Tabulated feedback of reviewer data

| Author/s | Dhondt, T. W., L. A Verbruggen, RA B Oostendorp, W Duquet, W. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1999 | | | | | |
| Title | Pain threshold in patients with rheumatoid arthritis and effect of manual oscillations. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | No | No | No | 67% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | No | No | No | 67% |
| 5 | There was blinding of all subjects | Yes | Yes | Yes | Yes | 100% |
| 6 | There was blinding of all therapists who administered the therapy | Yes | Yes | No | Yes | 67% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by intention-to-treat | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 10/11 | 8/11 | 7/11 | 8/11 | |
| Overall percentage agreement | | | | | | 91% |

Table 4.11b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Dhondt, T. W., L. A Verbruggen, RA B Oostendorp, W Duquet, W. |
| Year | 1999 |
| Title | Pain threshold in patients with rheumatoid arthritis and effect of manual oscillations. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Pain threshold (PTT) (measured by pressure algometer) at several designated positions. | Part one: a single measure compared between two groups.<br><br>Part two: the measures of part one of the RA group represented the Pre-intervention measures. Post-intervention conducted at the end of the same interaction, following the prescribed intervention. | Study was conducted over two months. However, the period between measurements was conducted in a single consult/interaction. | Part one: n = 60 (30 control, 30 RA group) with no fall out reported.<br><br>Part two: n = 30 (15 RA control, 15 RA experimental group) no fall out reported. | Yes | Yes,<br>Part one: 30 healthy controls.<br>Part two: 15 participants from part one RA group. | Yes, method not defined. | 8/10, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 91% |

| | |
|---|---|
| Limitations | The study made use of two sampling methods. The RA participants were recruits participants from the Free University of Brussels, Academic Hospital, Rheumatology Clinic (Belgium). While the healthy control was drawn by convenience from the participants family and staff of the Academic Hospital. Because of the specialist medical clinic functioning as the source of RA participant, the study is at higher risk of recruiting a sample that is poorly controlled for confounding introduced through this, such as those suffering a greater degree of severity in disease activity (Wells *et al.* 2012). As a result, risks unintentionally limiting the generalisability of the study, and introducing confounding factors and bias that the study design is not controlling for. In such a scenario the outcomes reflect external influence rather than intention. Furthermore, such a sample is unlikely to be naïve to or consider the intervention novel (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013).<br><br>This discussion only centres around part two of the study, which is the actual intervention portion of the study as it was designed.<br><br>All RA patients were required to meet the revised ARA criteria (Arnett *et al.* 1988). The collective group of RA patients had baseline signs and symptoms reported, but it is unclear as to whether the two RA groups in part two of the study have similar noted characteristics. However, the study changes the comparative analysis of part two, by comparing the two RA groups against the control as appose to each other. This was due to the randomly allocated groups (of RA in part two) not carrying comparability between them in all areas at baseline. Furthermore, the descriptors of the RA patients fail to consider reporting on RA serology, RA functional class, RA duration and activity, flare-ups, contraindications and / or comorbidities (England *et al.* 2019; Myasoedova *et al.* 2020). Nor do they consider whether these would be comparable between the groups (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). This limits the clarity of description for both the sample and patient (to which outcomes would be applicable to), with this study's capacity to be repeated and the generalisability of its findings are both limited (Wells *et al.* 2012; Smith *et al.* 2013). There was reported randomisation, but not stratification or clarity on concealed allocation of the participants (Mouton, 2006). There was no reported *a priori* analysis determining the size of the two groups (n=15 for both) (Kendall 2003; Christley 2010), requiring use of non-parametric statistics, thus allowing only trends analysis of intervention effect (Hoskin 2012; Bahreini *et al.* 2020). The study at the very least is therefore underpowered to detect difference that is greater than chance alone (Mouton 2006; Christley 2010). To the study's credit, the intervention technique used, and control (no intervention) were clearly described, and from an intervention protocol vantage point, is easily replicated (Kendall 2003; Heine *et al.* 2012).<br><br>In this pre- and post- pain pressure threshold (PPT) measures study, the outcome of back tenderness could only be reported as an immediate outcome. Nine RA participants reported never having had any back pain, 12 had intermittent back pain and nine had back pain at the time of the study, but it is not clear whether these were distributed equally between the groups. Since PPT is affected by the patients' pain tolerance (Morris, Cruwys and Kidd 1997; Melia *et al.* 2019), this would be confounded by the absence or presence of overt pain (i.e., the patient is more sensitive). The window between pain sensitivity and tolerance is thus decreased, making the possibility of tenderness being reported more likely (Melia *et al.* 2019). This would obscure the ability to test the intervention when there is variance among the patients in terms of the tolerance / sensitivity window (Dhondt 1999; Melia *et al.* 2019). These concerns are reflected in the study as the authors acknowledge that the two randomly chosen groups were not equivalent and thus the study was at risk of a type II error (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). This is further compromised by the fact that the RA duration and severity, if different between the groups, could influence the response of the patient to the PPT readings particularly as this is a once off intervention (pre- and post- study) that may actually have no effect or only a negligible one which may not be visible, given all the confounders not accounted for and acknowledged in the study. All RA participants were receiving medication as treatment during the study period, although the effect of medication cannot be determined per group owing to limited detail regarding |

| | |
|---|---|
| | the prescribed medications, dosage, and the duration of the therapy. Control of the short-term use of NSAIDs or steroids may have affected pain sensitivity, thus altering the range between pain tolerance and pain sensitivity which, if different between the groups, may affect the outcomes (Bartley and Fillingim 2013; de Boer *et al.* 2015; Zhao *et al.* 2017; Heidari *et al.* 2019).<br><br>Given the effects of diurnal factors on RA, it is unclear whether the PPT measures were taken at the same time or in the same time period daily (Pater *et al.* 1998). Additionally, the touch effect between the manual oscillation group and the no treatment control also needs to be considered as a confounding factor (Melzack and Wall 1965; Moayedi and Davis 2013; Senderovich *et al.* 2016), along with the placebo effect which likewise requires an appropriately powered study to eliminate the placebo effect which needs to have been controlled for (Kaptchuk 1998; Bialosky *et al.* 2011). These factors may have been negated if the patients were naïve regarding the intervention (Mouton 2006).<br><br>The primary outcome and measurement tool of the study was objective, suitable and validated (Reeves, Jaeger and Graff-Radford 1986; Fischer 1987; Hogeweg *et al.* 1992; MacGregor *et al.* 1997; Ohrbach, Crow and Kamer 1998). Nevertheless, as the algometer tests soft tissue tenderness (Park *et al.* 2011; Więckiewicz *et al.* 2015) and RA typically affects joints (Lewis and Battaglia 2019), the algometer is typically seen as an indirect measure of RA pathogenesis. The application of the pain pressure threshold applied through an algometer may also have been affected by the positioning of the patient when the measures were taken (Melia *et al.* 2019). With the patients in the prone position, it is likely the joint orientation (Mutlu and Ozdincler 2015), normal or abnormal spinal curvatures (Grob, Frauenfelder and Mannion 2007) and / or RA joint involvement at variable levels may have affected the readings as directly opposed to the intervention, as can be seen in part one of the study. This may be partially negated due to the pre/post design and the consistent measurement environment. Given the manner in which the report is written, it appears the participants were blinded, as were the assessor/s. It is however unclear if the participants knew of group allocation considering that informed consent from participants would have been required for the study to be approved. |
| Outcome | Part two of the study, dealing with pre/post measurement of the pain pressure threshold in RA patients, showed an increase in sensitivity, with a decreased pain pressure threshold (or an increase in tenderness) in all subjects that was slightly less pronounced after a single manual oscillation session than after rest (Dhondt 1999). |
| Discussion | This particular study is conceptually well-structured in terms of the pre-post design and complies with the required outcomes as outlined in the PEDro scale (PEDro 1999) Nevertheless, there are multiple extraneous factors relating to the individual patients, as well as to the manner in which patients were allocated to the groups and to controlling for these differences. There are also implications regarding the manner in which data was collected with the algometer that provide points of bias in the study. It is therefore impossible to state categorically that the changes are only due to the intervention. As a result, the outcomes of this study need to be interpreted with caution and applied with discretion in clinical practice. |
| Conclusion: | Given the limitations noted above, from the external validity vantage point this study suffers from a moderate degree of bias and the reader cannot exclude the effect of an external influence/s. This is mildly out of agreement with the review by the reviewers, which ranks the study as an 8/11 with 91% agreement, indicating that there is high methodological rigour and low risk of bias to the internal validity of the study in terms of the PEDro scale. The findings of this study are, therefore, moderately difficult to apply to the clinical setting in that the parameters of the sample that the intervention worked on lacks clarity and thus the clinical practitioner will have difficulty identifying RA patients best suited to and most likely to benefit from such an intervention. Furthermore, the long-term effect is unknown, while causal effect cannot be defined. Therefore, this study at the best has a moderate evidence and until further research can be carried out the intervention should be considered viable but implemented cautiously for management of RA.<br><br>This study should be repeated in the future with due diligence to considering the sample size as well as the confounding factors in order to be able to arrive at more acceptable conclusions that are clinically applicable and valid. |

Table 4.12a: Tabulated feedback of reviewer data

| Author/s | Durcan, L., Wilson, F. and Cunnane, G. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2014 | | | | | |
| Title | The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | No | No | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least one key outcome | No | No | No | No | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | Yes | Yes | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| | Total score | 6/11 | 7/11 | 7/11 | 7/11 | |
| | | | | | Overall percentage agreement | 91% |

Table 4.12b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Durcan, L., Wilson, F. and Cunnane, G. |
| Year | 2014 |
| Title | The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Health Assessment Questionnaire (HAQ) Disability Index.<br>Visual Analog Scale (VAS) for pain and stiffness<br>Fatigue Severity Scale (FSS)<br>Pittsburgh Sleep Quality Index (PSQI)<br>Exercise Benefits and Barriers Scale (EBBS) | Baseline and week-12 for the control group<br><br>Baseline, week -three, -six, -nine, -12 for the intervention group. | 12-weeks | Total sample at baseline (n=80): Intervention (n=42) and control (n=38). From the intervention group, two were excluded after baseline assessment due to lake of suitability to the study. Thus, leaving a total of 78. | No | Yes, the control group received advice only on the benefits of exercise in RA. | Yes, randomised by excel random numbers | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 91% |

| | |
|---|---|
| Limitations | The sample was recruited from a teaching hospital outpatient rheumatology clinic (Ireland). This specialist centre is associated with the tertiary referral hospital for the area infers and leads the reader to assume that the patients recruited likely to be suffering a greater degree of severity in disease activity, in doing so this sample risks representing a sub-population of RA rather than the true RA population (Wells *et al.* 2012). As a result, risk is imparted to the generalisability of the study findings, and it is unlikely the sample is naïve to the intervention type. The eligible participants had to meet the following set of recruitment criteria, established RA in compliance with ACR (vision and citation not provided), capable of independent ambulation, were not a risk of fall, English literate and capable of providing informed consent, and within an hour commute of the hospital. The study controlled for comorbidities ("severe medical conditions") via exclusion criteria. By means of this the criterion, offered protection (to a degree) over contradiction to therapy and enhancing adherence. The participants were then screened for age, sex, serology (RF and ACCP), erosive disease (on x-rays of hands or feet), disease duration, medication (biologic therapy use) and smoking to confirm compliance with inclusion criteria. Given this thorough assessment, the study on the other hand failed to consider functional class, disease activity and use of other physical modalities.<br><br>A pre-assigned protocol was utilised to randomise the patients into one of two groups, but no form of stratification was utilized to enable mitigating the various factors related to the RA that could influence the outcomes. Whether the allocation to groups was per the concealed method is unclear. Reviewing the baseline comparison between the groups indicates that there was a difference (reported as significant) in the VAS readings for pain, which is not seen in the other baseline characteristics (HAQ, VAS (Stiffness), PSQI, Barriers, EBBS, and FSS). Given the differences at baseline, the authors note that they did not control for this difference/s statistically (which is in line with de Boer *et al.* (2015)), as it was stated by the author as not being the primary outcome and subject to fluctuation, however they went on to acknowledge this could lead to bias and a type II error in their conclusions. The effect of this may be secondary to floor or ceiling effects related to the differences between the groups that temper the change possible and cause artificial outcomes (Liang 2000; Coster 2013).<br><br>Using the MCID of the HAQ, the authors completed an *a priori* power analysis which required a sample size of 38 per group. The study met these requirements with n=42/n=40 in the intervention and n=38/n=38 in the control for entry into the study and final analysis [which is incongruent with primary outcome measure measuring sleep]. This sample size and completer numbers per group enabled the use of parametric statistics, thus allowing the study to produce causal links between intervention and outcome that are better than chance alone (Hoskin 2012; Bahreini *et al.* 2020). The outcome measures consisted of a mix of subjective and objective tools (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), which were suited, reliable and validated for use in the measuring change in RA patient clinical status (Coster 2013; Bahreini *et al.* 2020). However, in terms of the patient effects, the study does not control for the potential patient learning processes in response to subjective outcomes, the placebo effect, the Hawthorne, patient naivety and environmental or diurnal changes related to RA (Kaptchuk 1998; Mouton 2006; Richardson 2007; Bialosky *et al.* 2011; Moustgaard *et al.* 2014; Sedgwick and Greenwood 2015; Nandgude, Hasabe and Kolsure 2018). Furthermore, subjective outcomes are threatened by Memory recall / decay (Mouton 2006; Ricker, Vergauwe and Cowan 2016), particularly with outcomes like PSQI self-assessment evaluates for the preceding month at the time of assessment. Another shortcoming relates to the control of medication, where only baseline analysis of biologic therapy is reported. In addition, the biologics used are not detailed in terms of type, duration, and dosage and, additionally, no data is reported in relation to antidepressant medication, NSAIDs, DMARDs and steroids. Furthermore, the interruption and continuation of and the changes and adherence to medication are not detailed (Zhao *et al.* 2017; Heidari *et al.* 2019). These shortcomings are acknowledged by the authors in that the anti-inflammatory effects of exercise (intervention) (Baillet *et al.* 2012; Palmefors *et al.* 2014) may be masked by the medication. |

| | |
|---|---|
| | The impact of the medication is further compounded by the fact that the control was not a true control. This is because the control group was educated on the benefits of exercise. As an exercise diary was not included in the study it is not clear whether the control group participated in their own schedule of exercise. Thus, with an intervention possibly implemented in the control the study becomes one of determining the relative effectiveness of the interventions. This is unfortunate in the light of the fact that the intervention technique was clearly described and is reproducible (Heine *et al.* 2012). This 12-week home-based tailored prescription for the intervention group, the tailoring of the exercise prescription was targeted to functional limitations noted in examination at the start of the study.<br><br>The measurement environment is described with procedure and protocol, and the assessments occurred at baseline and post-intervention for all participants. However, the intervention group had additional assessments dispersed through the study timeline at three-week intervals (i.e., at baseline and week -three, -six, -nine, and -12). This structure is, however, at odds given that the data is presented only in the form of the pre/post design (i.e., data of the intervention groups intermediate assessment not published). The study goes further to state that the intervention group all attended the assessments, aside from three reportedly missing one (missing data). The lack of this data does not allow identification of the periods in which patients attained the best outcomes and only provides an overall outcome of the study. It is unclear whether patients were blinded to group allocation, but it seems that outcomes assessors were not blinded. These limitations imply that there may be bias as related to the patient effects as well as the presence of assessor bias (Kendall 2003; Moustgaard *et al.* 2014). Furthermore, the outcomes measurements relating to the week -12 assessment fails to report the period between the last intervention and the measurement. This suggests that the outcomes may reflect either the immediate effect of the last intervention (if taken to soon after the last intervention session) or reflect the effect of the 12-week programme as a whole. The authors' statement that a "challenge for the future includes … re-evaluating those who took part in our 12-week programme to see whether it has had a lasting effect" implies that measurements past the twelfth visit are necessary to measure the outcomes more accurately and evaluate longevity to the effects (Pater *et al.* 1998). But in favour of the study, the 12-week period of exercise intervention was sufficiently long to be able to measure the effects of the intervention. Exercise as a modality requires six to eight weeks at minimum to observe impacts on the functional capacity and thus ADL (Fraser 2008), therefore 12-weeks would show improvement if yielded. |
| Outcome | The authors concluded that their study reflected evidence that RA participants undertaking a 12-week home-based exercise programme, imparted significant improvement in sleep quality (PSQI), fatigue levels (FSS), functional disability (HAQ), pain and stiffness (VAS), barriers to exercise and opinion regarding it benefit (EBBS) [$p<0.001$]. While in the control imparted significant improvement in perceived benefit of exercise (EBBS) [$p<0.001$] only. When comparing the change noted between the groups over the study, it was evident sleep quality (PSQI) and fatigue (FSS) had significantly more improvement in the exercise group [$p=0.040$]; as well as functional disability (HAQ) in favour of the exercise group [$p<0.001$], pain and stiffness (VAS) [$p=0.050$]. From this the authors claimed these findings suggest exercise may aid in addressing important negative impact on patients with RA's life and provides reasoning for the incorporating of prescribe exercise in managing this population (Durcan, Wilson and Cunnane 2014). |
| Discussion | In view of the limitations regarding the patient factors, the lack of control of the baseline characteristics related to RA and the effects of medication on the outcome of a study. Despite the study being sufficiently powered to attaining an outcome greater than chance alone, it is difficult to contextualise the outcomes (although tempting) to accept them, due to comparison with a control gaining educational intervention. For these outcomes to be applicable to patients in a clinical context, they would not only be required to meet the requirements of the participating RA patients as outlined in this study, but the study would need to be repeated with consideration of the shortcomings presented here in order to validate the outcomes. Therefore, the external validity of this study is limited, and caution should be applied in clinical practice. |
| Conclusion | Given the limitations noted in this analysis, from the external validity vantage point this study suffers from a moderate degree of bias and the reader cannot exclude the effect of an external influence/s. This is consistent with the review by the reviewers, which ranks the study as an 7/11 with 91% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study in terms of the PEDro scale. Therefore, this study at best offers moderate evidence and until further research can be carried out this evidence should be taken in the context of effectiveness rather than efficacy when considered for management of RA. |

Table 4.13a: Tabulated feedback of reviewer data

| Author/s | Eversden, L., Maggs, F., Nightingale, P. and Jobanputra, P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2007 | | | | | |
| Title | A pragmatic randomised controlled trial of hydrotherapy and land exercises on overall wellbeing and quality of life in rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | No | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by intention-to-treat | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | No | Yes | Yes | 67% |
| Total score | | 9/11 | 7/11 | 8/11 | 8/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.13b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Eversden, L., Maggs, F., Nightingale, P. and Jobanputra, P. |
| Year | 2007 |
| Title | A pragmatic randomised controlled trial of hydrotherapy and land exercises on overall wellbeing and quality of life in rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Primary outcome:<br>Self-rated overall effect of treatment<br><br>Secondary outcomes:<br>Pain VAS<br>HAQ<br>EQ-SD (EQ-VAS + EQ-5D index)<br>10-meter walking speed. | Primary outcome: week 6<br><br>Secondary outcomes:<br>Week 0, 6, 18 | Intervention of six weeks with a follow-up at three months<br><br>4.5 months | Total start (n = 115):<br>Land group (n = 58) and Hydrotherapy (n = 57).<br><br>Drop-out:<br>Land group (n = 17) and Hydrotherapy (n = 13).<br><br>Completers:<br>Land group (n = 41) and hydrotherapy (n = 44). | Yes | No control. Two interventions groups. | Yes, by means of sealed opaque envelopes (indicating treatment allocation). Randomisation of envelopes were prepared at study inception and random number sequence by flipping a virtual coin. | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study recruited participants attending specialist clinics (rheumatology) at the Selly Oak Hospital (UK). The participants were invited by means of physiotherapy referral from the clinics, invitation during clinic attendance, or via mail. Given the nature of the source/s, the impression is that the sample is more likely to consist of participants suffering a more active disease status (Wells *et al.* 2012). As a result, the study unintentionally risks limiting the generalisability of results and introducing confounding factors (bias) that the study design is not controlling for. In such a scenario the outcomes could reflect external influence rather than the tested intention. Additionally, such a sample is unlikely to be naïve to or consider the intervention novel (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013). The participant eligibility was assessed by a single physiotherapist according to the following set of inclusion criteria, through which efforts were made to offer patient description and thus generalisability of the outcomes. The included participants were older than 18 years of age, diagnosed with RA (unknown ACR criteria) with a functional class of between I-III (Hochberg *et al.* 1992), to be English literate and to be on a stable medication dosage (DMARD therapy: six weeks; NSAIDs: four weeks and corticosteroid injections: four weeks). Once enrolled in the study, participants were randomly allocated (sealed opaque envelopes), which supported concealed allocation (Mouton 2006). Stratification was not done even though there was a possibility of incompatibility between the two RA patient groups, given the large number of dynamic RA characteristics such as fatigue, RA activity (DAS28) and serology (England *et al.* 2019; Myasoedova *et al.* 2020). Notwithstanding this, the authors report baseline characteristics and distribution to be comparable between the groups (age, sex, RA duration and medication (DMARDs, corticosteroids, and NSAIDs)), thus the sample groups seem homogenous. Participants were excluded if they had surgery (either planned for or within three months preceding their enrolment), participation in hydrotherapy or physiotherapy within six months of the enrolment, pregnancy, known carriers of MRSA infection, and a collection of contraindications of hydrotherapy/safety (e.g., chlorine sensitivity, and weight >102 kgs) of the intervention equipment and comorbidities. Through these, the sample is further defined, and the study is offered protection of missing data (due to contraindication to intervention and / or outcome measure), protection against carry over effect of other physical intervention/s, and safety measures (Bordens and Abbott 2002; Kendall 2003; Durcan, Wilson and Cunnane 2014).<br><br>The study utilised an *a priori* analysis to determine the minimum sample size, which was then supplemented by dropout provision of 10% dropout rate (Fogg and Gross 2000; Kendall 2003; Christley 2010). This collectively indicated a need for 60 participants in each group. To further limit the impact of dropout, the principle of intention-to-treat was applied (de Boer *et al.* 2015; Unbiased Research 2017). Even with these fail-safes, the study recruitment was terminated early when the closure of the hydrotherapy facility. This resulted in the intervention groups having less than the intended sample size (however, still meeting power requirements at baseline), this was followed by significant dropout rates. Accordingly, the sample did not meet the requirements for detecting statistically relevant change in the primary outcome. This indicates that the study was underpowered (Mouton 2006; Christley 2010), which does not allow for patient effects to be negated. Nonetheless, the sample size still allowed for the use of parametric statistics, potentially offering outcomes better than chance (Hoskin 2012; Bahreini *et al.* 2020).<br><br>In addition, the authors controlled for attendance by recording the attendance (not significantly different between the groups), although the patient number reduction did have an effect on sample size. Through post hoc analysis three assumptions were tested: First, if both group non-completers were counted as non-responders, the response rate would be hydrotherapy (70%) and land exercise (33%). This *post hoc* assumption is in favour of hydrotherapy ($p$<0.001); Second assumption, non-completers of both groups were analysed as responding as the completers did this was the same for both groups, the response rate is then hydrotherapy (82%) and land exercise (55%), This post hoc assumption is in favour of hydrotherapy ($p$=0.002); Finally, the third assumption. Non-completers of the hydrotherapy group were analysed as non-responders, and land exercise non-completers analysed as responders, in this case response |

| | |
|---|---|
| | rates for hydrotherapy (70%) and land exercise (64%), This post hoc assumption reflects insignificant difference between interventions (*p*=0.59). To the strength of the study, effort to present objectively the intention-to-treat analysis in more than one manner. However, post hoc analysis can report data that is misleading and overlooks confounding and bias factors pertinent to the objectives of the study and to the extraction of findings into clinical practice (de Boer *et al.* 2015). Secondly, the assumptive scenarios stipulated above, are statistical forecasting/predictions and carry the risk of misrepresenting clinical outcomes (Gupta 2011).<br><br>With the exception of participant perception (Kaptchuk 1998; Bialosky *et al.* 2011), the reliable and valid and appropriate (Coster 2013; Bahreini *et al.* 2020) baseline measurement outcomes (participant perception, EQ-5D VAS, HAQ, Pain VAS, 10-meter walk time), were all comparable between the groups at the beginning of the study (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). Data was collected by an independent blinded assessor, but timing and location of the assessments is unclear (Pater *et al.* 1998; Kendall 2003; Richardson 2007; Moustgaard *et al.* 2014). Additionally, the relationship between intervention and measurement is unclear and allows questioning of what the outcomes actually measured (the protocol versus only the prior treatment). This has an impact on immediate, short-term, and long-term evaluations of the outcomes (Pater *et al.* 1998). In terms of medication, there was control at the baseline as participants entered the study. However, the study then allowed drug changes and injections to reflect the pragmatic nature of the study and RA in clinical practice. This means that the medication control (interruption / continuation / discontinuation / compliance / dosage changes and type) was lacking. The only exception was that medication data collection was again taken at the final assessment. Common changes in medication included 10 hydrotherapy and nine land exercise group participants seeing an increase, while four hydrotherapy and two land exercise group participants reported a decrease or cessation in DMARDs therapy, with four in both groups receiving steroid injections. These changes in medication were identified by the authors' limitation.<br><br>In this study, the well-managed, customised, and consistent programme of six-weeks of exercise as an intervention, with one land or hydrotherapy intervention per week complies with the minimum period required to observe effect on the functional capacity and thus ADL (Fraser 2008). For this reason, the measurements around this could have a lack of sensitivity to observable effect (Coster 2013; Bahreini *et al.* 2020). Given that each group's programme contained exercise, the outcomes of the study are related to the relative effectiveness and not the efficacy of the interventions, this is contrary to what is purported by the title, which implies that there is a control group. Further complications are that the intervention technique used in either group was vaguely described thus lacking the capacity to be repeated specifically, and also that in both groups' participants had the exercises tailored to the participants ability (with tailored instructions). However, in support of the study, the Hawthorne effect would be minimised as each participant would have had the same levels of exposure to the researchers and participants would have been encouraged to participate through a peer encouragement group since exercises were done in groups. In contrast, to this the participant preference for land or hydrotherapy interventions may have played a role, not only in the number of dropouts but also in how they responded to the outcome measures (Mouton 2006; Moustgaard *et al.* 2014). |
| Outcome | The authors reported that on completion of the programme the RA patients treated with hydrotherapy are more likely to report feeling much better or very much better than those treated with land exercise. This perceived benefit was not reflected by differences between groups in 10-meter walk times and functional scores (quality of life measures (EQ-5D utility; HAQ) and pain (EQ-5D VAS) scores) (Eversden *et al.* 2007). |
| Discussion | Given that this study utilised a customised exercise intervention programme for both groups in the study, additionally the study suffered the impact of the small sample size, the patient effects as well as the structural components of the study would require this study to be repeated in more controlled circumstances to validate the outcomes. Furthermore, the fact that all the participants were recruited from one hospital-based site does not necessarily allow for the outcomes of this study to be generalised, as it is unclear whether the participants truly reflect a more general RA population. In the light of these outcomes, the limited conclusion that can be drawn is that the exercise may benefit RA patients and that this exercise should be tailored to patient preference. |
| Conclusion: | Given the significant limitations noted above, from the external validity vantage point this study suffers from a high degree of bias and prevents the reader from excluding the effect of a set of external influences. This is in direct contrast to the outcome of the review by reviewers that with high degree of congruency (94% agreement) reflects a PEDro score of 8/11, indicating that from the internal validity vantage point the methodological rigor is of high quality and the validity is subject to low risk of bias. This highlights a disjuncture between the two reviews. It is however possible to have strength in the rigorous study, but the bias and confounding in this study is from external influencers outside of this research process. Therefore, this study at best provides limited to no evidence. Until further research can be carried out in a RCT design with true control comparison and control measures introduced to address the limitations noted above, this review cannot draw evidence as a basis for either intervention in the management of RA. |

Table 4.14a: Tabulated feedback of reviewer data

| Author/s | Garner, S., Fenton, T., Martin, L., Creaser, C., Johns, C. and Barnabe, C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2018 | | | | | |
| Title | Personalized diet and exercise recommendations in early rheumatoid arthritis: A feasibility trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | Yes | Yes | Yes | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | No | No | No | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | No | No | No | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by intention-to-treat | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 6/11 | 6/11 | 5/11 | 6/11 | |
| Overall percentage agreement | | | | | | 91% |

Table 4.14b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Garner, S., Fenton, T., Martin, L., Creaser, C., Johns, C. and Barnabe, C. |
| Year | 2018 |
| Title | Personalized diet and exercise recommendations in early rheumatoid arthritis: A feasibility trial. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Clinical outcomes:<br>Tender and swollen joint count; DAS28; BMI; Waist-to-hip ratio; Blood pressure (systolic, diastolic, mean arterial pressure); duration of morning stiffness.<br>Subjective outcomes:<br>Global well-being VAS (physician and patient)<br>Laboratory outcomes:<br>Cholesterol (LDL), ESR, CRP | At baseline and at six months (post-intervention) | Six-months | At baseline evaluation = 30 (randomised into 15 per group).<br><br>At completion = 23 (10 in the control and 13 in the intervention groups) | No | Yes, standard care group | Yes, by means of concealed envelops. | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 91% |

| | |
|---|---|
| Limitations | The study recruits' participants via a specialised rheumatology clinic (for early inflammatory arthritis) at the university of Calgary (Canada). The impression is instilled that given the nature of the source from which participants are drawn, increases the likely of recruited of a sample that is poorly controlled for confounding introduced through this, such as the those suffering a greater degree of severity in disease activity (Wells *et al.* 2012). As a result, risks unintentionally limiting the generalisability of the study, and introducing confounding factors and bias that the study design is not controlling for. In such a scenario the outcomes reflect external influence rather than intention. Furthermore, such a sample is unlikely to be naïve to or consider the intervention novel (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013).<br><br>The study's inclusion criteria state that the participants were required to be diagnosed in the preceding year (ACR 2010 criteria (Aletaha *et al.* 2010)), be on DMARD therapy (but without considering dosage and other medication related factors) and capable of participate in all study evaluations. While controlling confounding that could be derived from comorbidities in those participants with hyperlipidaemia, hypertension and diabetes of an uncontrolled nature, or those that were pregnant were excluded (Kendall 2003). Given the fairly comprehensive inclusion and exclusion criteria, the authors did not seem to consider RA specific serology, RA functional class, RA activity, age, medication factors, prior or concurrent use of other physical modalities or additional comorbidities in these criteria or through stratification. Randomisation was noted and was by means of concealed envelopes. The randomisation without stratification led to baseline characteristics being fairly well distributed and homogenous between the groups in terms of age, sex, current smoking, disease duration and serology. The lack of consideration of some RA related factors and the stratification of these might however have introduced an unintended bias that was not accounted for in this study. The calculation of an *a priori* analysis is not reported. The sample size of 30 of which two participants withdrew prior to the study and five were lost to follow-up (dropout = 23.3%). Intention-to-treat analysis was not utilised. The small sample size prevented the use of parametric statistics, thus limiting the study to draw conclusions only on trends and not causal links between intervention and outcome (Bordens and Abbott 2002; Hoskin 2012; Bahreini *et al.* 2020). In addition, the sample size cannot rule out that the outcomes attained are due to chance alone so they must therefore be interpreted with caution.<br><br>The outcome measurements at baseline included the following: BMI, waist and hip circumference ratio, BP (systolic, diastolic, mean arterial pressure), cholesterol levels (LDL), random glucose, inflammatory markers (ESR and CRP), joint count (for both swelling and tenderness (28 joints), DAS28, HAQ, step count by pedometer, timed up-and-go test, timed stand, sit and reach test; one repetition maximum test, the six minute walk test and the number of co-morbidities and patient global evaluation score. The distribution of these outcomes was for the most part homogenous between the groups but the joint count (swollen ($p$=0.002), tender ($p$=0.001)) and disease activity score ($p$=0.001) were significantly different between the groups. As acknowledged by the authors, the distribution of these measures between study groups could potentially put the study risk of type II error being drawn in the conclusions, particularly in the context of the small sample size (Mouton 2006; Christley 2010).<br><br>Given that the study had two intervention groups, it is assumed that the control was the standard treatment group. Based on this assumption, it is therefore clear that the study set out to test the relative effectiveness and not the efficacy of the test intervention as it is unclear whether standard treatment is indeed better than placebo. Additionally, |

| | |
|---|---|
| | the comparison of interventions is that of intervention programmes and not individual treatment modalities. Therefore, the study is limited in commenting on any one modality in the intervention programmes. It is also important to consider that the higher attrition in the standard treatment group may have been due to previous exposure with regression or lack of improvement in clinical outcomes or lack of patient naivety (Mouton 2006). In order to better control for the influence on outcome of the various aspects of the intervention/s, in that clarity should be provided in terms of what constitutes both intervention programmes. This would further aid in allowing for study repetition to be achievable, the foundation of intervention is recommendation and education provided through counselling and handout material. From this foundation the intervention group adds to the intervention programme. The lack of clarity around intervention variability in this study, leads the reader to these intervention groups as a collection of participants with individualised treatment plans that loosely conform to standard care or intervention, which would make the study less of a RCT and more of a prospective case series, with this the associated problems in data analysis and reporting (Kendall 2003; Heine *et al.* 2012; Moher *et al.* 2012). Additionally, the individualised care would have introduced the Hawthorne effect, the observer effect and the attention effect which carries the potential for influence in the outcomes, particularly in the group receiving non-standard care and believing this to be different and better to previous standard care. This would have amplified the differences in outcome measures in the groups and, as acknowledged by the authors, further have led to type II errors in the conclusion.<br><br>The subjective and objective outcome measures in this study are well balanced (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), are generally accepted as well suited (Coster 2013), reliable and validated (Coster 2013; Bahreini *et al.* 2020) tools for RA in the context of the study. These outcome assessors were, however, not blinded to the group allocation which introduces researcher bias in addition to the lack of reporting on timing of the measures which influences what the outcomes truly reflect, if outcomes are assessed directly after a therapy session such that a patient may have exercised prior to attending follow-up, would reflect the immediate effect of the intervention (or component of) instead of the global effect of the intervention programme. Furthermore, the timing of the outcome assessment is not protected from bias and confounding (e.g., the diurnal changes in RA (Pater *et al.* 1998; Nandgude, Hasabe and Kolsure 2018)). Given that this study required participants engaging with the intervention such as to implement both diet change/recommendations and an exercise plan/recommendation which will lead to variable implementation/exposure and will take time for an effect to be measurable, the baseline and six-month post-intervention follow-up seem appropriate (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). A diet and exercise diary would have been helpful in considering the adherence to both groups within the six-month period and to account for compliance. A follow-up post-intervention would have been useful to consider the impact of the interventions over an additional period in terms of whether the benefits were sustained without the participant being in an active trial. |
| Outcome | The authors conclude that individualised dietary and detailed exercise prescription for patients with early inflammatory RA does have the ability to improve activity levels and dietary intakes. The statistics revealed that only 13 patients in the intervention group and 10 in the control group completed the study. The trends for improvement were non-significant for physical activity, LDL, and nutritional intake (vitamin -C and -A; iron; fibre; folate) in favour of the intervention group. The authors note that the poor enrolment and high dropout rates in this short-term study highlighted the difficulty of effecting behavioural change (Garner *et al.* 2018). |
| Discussion | The lack of control for confounding variables in the study, along with the very small sample size, suggests that the outcomes may actually be a result of chance rather than a result of the interventions. Compounding factors were the seemingly individualised care programmes for each patient, more reflective of a prospective case series as opposed to a study where the interventions are uniform in order to compare their outcome. The limitations discussed above and the acknowledgement by the authors that the study had significant limitations in patient retention together mean that the outcomes of this study are not generalizable to the RA population and therefore the recommendations stemming from this study need to be applied with caution in clinical practice. |
| Conclusion: | Given the significant limitations noted above, from the external validity vantage point this study suffers from a high degree of bias and the reader cannot exclude the effect of external influence/s. This is to a degree contrasted by the review by the reviewers, which ranks the study as a 6/11 with 91% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study in terms of the PEDro scale. This highlights a disjuncture between the two reviews. It is, however. possible to have relative strength in the rigour of the study, but the bias and confounding in this study could be induced by influencers outside of the research process. Therefore, this study at best has limited to no evidence and until further research can be carried out, this cannot be a used as a basis for management of RA. |

Table 4.15a: Tabulated feedback of reviewer data

| Author/s | Hakkinen A., Malkia E., Hakkinen K., Jappinen I., Laitinen L. and Hannonen P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1997 | | | | | |
| Title | Effects of detraining subsequent to strength training on neuromuscular function in patients with inflammatory arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | No | Yes | Yes | Yes | 67% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | No | Yes | No | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | No | No | No | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 4/11 | 5/11 | 6/11 | 5/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.15b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hakkinen A., Malkia E., Hakkinen K., Jappinen I., Laitinen L. and Hannonen P. |
| Year | 1997 |
| Title | Effects of detraining subsequent to strength training on neuromuscular function in patients with inflammatory arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - max voluntary bilateral isometric force and force-time curve<br>- max voluntary bilateral dynamic concentric strength of knee extensors<br>- max isometric force of trunk flexors and extensors<br>- max isometric grip strength<br><u>RA groups only:</u><br>- Ritchie's articular index<br>- ESR<br>- Hb<br>- physical activity questionnaire | Baseline and 42-month follow-up (for all three groups). Groups one and two had an additional assessment period at 6-months (post-intervention). | Three and a half years | At baseline: RA sample (n=43), group-one (n=23), group-two (n=20). Healthy sample (n=20) made up group-three.<br><br>Completers of group-one (n=21), group-two (n=18), and group-three (n=18). | No | Two controls, one RA control and one healthy control. | The RA groups were randomised, although no method stated. | 5/11, indicating a low to moderate methodological rigour and subject to moderate to high risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | This publication originating from University of Jyväskylä, Jyväskylä, Finland and is a continuation of the authors earlier investigations (Häkkinen, Häkkinen and Hannonen 1994; Häkkinen, Hannonen and Häkkinen 1995). One of which reported the effects of progressive strength training in RA participants on disease activity and neuromuscular function (Häkkinen, Häkkinen and Hannonen 1994), the second reported compared muscle strength between RA patients and healthy control (Häkkinen, Hannonen and Häkkinen 1995). Taking these papers into consideration, this study's primary aim was to report the impact of detraining on force production. As a result of this, this study offers a limited description of methodology and intervention, as it is based loosely on the preceding publications. It is assumed that the sample was recruited from the greater Jyväskylä but the lack of confirmed context provides no basis for understanding the severity of the RA (Wells *et al.* 2012).<br><br>The sample recruited into the study included 43 recently RA diagnosed (criteria unknown) participants. This study appears to have no further explicit inclusion criteria used. This decreases the generalisability of the findings and sources of bias from within the participants may not have been controlled for. These includes but may not be limited to medication (Zhao *et al.* 2017; Heidari *et al.* 2019), occupation and occupational status, surgical interventions (Naqvi, Hassali and Aftab 2019), use of other interventions (Mouton 2006), hand dominance (Lee, Gandevia and Carroll 2009; Dragert and Zehr 2013), contraindication to intervention (Durcan, Wilson and Cunnane 2014), lifestyle factors (Pedersen *et al.* 2006a; McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019), endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018), or other general demographic factors (Bartley and Fillingim 2013) and RA dynamic characteristics (McInnes and Schett 2011; Yang *et al.* 2017).<br><br>Randomisation was completed but there was no clarity in terms of allocation concealment (Mouton 2006). In addition, 20 health age and sex matched controls were recruited creating a second control group (group-three). The demographics of the sample were limited to the completers only, reporting on sex, age, height, weight, and BMI (no significant difference), with no comparison of RA dynamic characteristics. The study did not report a power analysis (Kendall 2003; Christley 2010). Given this, conclusions are, that the study is largely underpowered, limiting the ability to draw strong conclusions (Fogg and Gross 2000; Mouton 2006; Christley 2010). The study presents sample data according to the completers, this implies that the study did not utilise an intention-to-treat analysis (Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017), however it must be noted that the study design does not provide the needed assessment periods to incorporate this analysis. The size of the groups on completion prevents for the reporting of parametric statistics, as a result the non-parametric analysis at best reports trends (Hoskin 2012; Bahreini *et al.* 2020). |

| | |
|---|---|
| | All groups followed the same intervention design. Both controls (group-two (RA) and group-three (health)) maintained habitual physical activity. The intervention group (group-one) undertook six-months progressive supervised individualised strength training (two sessions for the first four-months, followed by two-to-three sessions for the last two-months), following this the group returned to unsupervised habitual physical activity (no strength training for 36-months). The control groups (group-one and -two) were instructed to continue with normal physical activities, supplemented by instructions to carry out stretching exercises to maintain ROM (for the duration of the study). The intervention is briefly described with reference to Häkkinen, Häkkinen and Hannonen (1994) where a full description is available. Thus, there was a limited description of the intervention variables and protocol, which prevents complete understanding of the strengths and areas for risk of bias (Heine *et al.* 2012). It was noted that there were no significant differences between the groups in the types of leisure time activities.<br><br>Another positive is that the outcome measures were balanced between subjective and objective tools (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Multiple measures of some of the outcomes suggest a regression to the mean and / or impact of memory recall (Mouton 2006; Ricker, Vergauwe and Cowan 2016), thus limiting the interpretation of the outcomes as a once off and most accurate measure of clinical ability at the time of measurement. In terms of baseline assessment, the strength of trunk muscles these were not significantly different from the healthy controls. Unfortunately, grip strength measures were significantly different ($p$<0.01); dynamic strength (knee extensor) measures were significantly different ($p$=0.01), and force-time curves (for knee extensors) were significantly different ($p$<0.05 – 0.01). These differences would indicate that there may be bias in the statistical comparison of outcomes (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). Finally, the outcomes are not defined in terms of the MCIDs (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020), which hinders clinical contextualisation. Variable habitual exercise programmes and retrospective reporting of exercise for the year prior to study further compound this. |
| Outcome | The study concluded that in RA patients the enrolment of exercise programme consisting of strength training and recreational / habitual exercise led to marked increases ($p$< 0.05 – 0.001) in all major muscle groups with regard to maximal strength (achieved without changes to the force-time curve shape). The initially reported gains in muscle strength deteriorated over the post-exercise detraining period (especially isometric contractile strength (extension ($p$<0.01) and flexion ($p$<0.01) of the trunk) and dynamic knee extension strength ($p$<0.05). However, this was not reported in grip strength. The studies RA control group reflected significant decrease of isometric strength in extension and flexion of the trunk at a consistent rate throughout the study. When RA groups were compared to an age and sex matched healthy control group, at post-test assessment, all strength values in the RA exercise and control groups were lower. According to the findings noted in this study, RA patients require continuous exercise meeting sufficient intensity, load, and frequency to minimize and / or prevent loss of muscle strength and functional capacity, frequently noted in the RA population (Häkkinen *et al.* 1997). |
| Discussion | The authors indicated that the outcome/s of this study provide evidence for deterioration of the RA patient neuromuscular function post-intervention and that there is a need to maintain exercise participation (as is seen in this studies exercise programme). The conclusions however rest on an underpowered study with the attendant statistical limitations. This, along with the limitations noted above, suggests that the outcomes may have been as a result of factors outside of and unaccounted for in this study and / or be as a result of RAs natural history / chance / time during the exercise programme and the deterioration portion of the study affected by patient adherence for example. The results are only presented as group means, not allowing the reader to interpret data variance within groups, risking misinterpretation of results and / or not being able to contextualise results, as a regression to the mean occurs and outliers cannot be accounted for. This negatively affects accurate and fair comparisons between the intervention group and control groups. With the deterioration phase being at risk of bias due to patient adherence and reporting, this may result in further misleading outcomes. These are all underpinned by limited participant factors being controlled for between the groups, making type II errors more common in accepting hypotheses, in addition to limiting clinical generalisability. |
| Conclusion: | Given the limitations noted above, particularly the unclear recruitment and patient RA dynamic characteristics being controlled for, aside from the statistical concerns, renders this study at a high risk of bias when considering external validity. In terms of the reviewers, their responses indicated a 94% agreement that the study presented with a moderate level of bias due to limitations in the methodological rigour (internal validity).<br><br>Therefore, these place the study at a moderate to high risk of bias, thus, limiting the study to providing moderate to limited evidence in support of exercise. This outcome suggests that future studies are required in order to confirm or refute the evidence provided within the confines of a tighter and more controlled methodology. |

Table 4.16a: Tabulated feedback of reviewer data

| Author/s | Hakkinen A, Sokka T, Kotaniemi A, Hannonen P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2001 | | | | | |
| Title | A randomized two-year study of the effects of dynamic strength training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | No | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 6/11 | 6/11 | 6/11 | |
| Overall percentage agreement | | | | | | 97% |

Table 4.16b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hakkinen A, Sokka T, Kotaniemi A, Hannonen P. |
| Year | 2001 |
| Title | A randomized two-year study of the effects of dynamic strength-training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Concentric strength (knee extensor)<br>Isometric grip strength<br>HAQ<br>DAS<br>BMD<br>Joint radiology | Evaluation at 0, 6, 12, 18, 24 months<br><br>Radiological evaluations at 0, 12, 24 months. | Two-years | 70 participants with 2 dropouts in the intervention group; 3 dropouts in the control group (after baseline evaluation). Then 3 diagnosis changes (but group is unclear) whose data were excluded.<br><br>intervention group drop out (n = one) (after 1st year) – data utilised in analysis. Thus 31 participants per group included in data analysis. | No | Yes, but the control was not no intervention. | Yes, by randomised clusters of four patients who had been stratified according to age and sex. | 6/11, indicating a moderate to high methodological rigour and subject to moderate to low risk of bias within the internal validity of the study. | 97% |

| | |
|---|---|
| Limitations | The sample was recruited by referral from specialists to the Jyväskylä Central Hospital Rheumatology Unit, Jyväskylä, Finland. This implies that all RA patients received specialist care on diagnosis, which limits the likelihood of novel interventions at the start of the study (Wells *et al.* 2012).<br><br>The inclusion criteria included the ACR diagnostic criteria (1987 ACR (Arnett *et al.* 1988)) and medication use (glucocorticoid or DMARDs were excluded). There seemed to be no consideration for or controlling of RA serology, RA functional class, RA duration and activity of disease, age, flare-ups, pain, use of other physical modalities, contraindications and / or comorbidities (England *et al.* 2019; Myasoedova *et al.* 2020). Of the noted baseline patient characteristics there were no significant differences between the groups in terms of sex, age, weight, height, current smoking, alcohol consumption and menopausal status. This picture of seemingly homogenous groups creates a distorted image of the sample groups in terms of their comparison with each other and with the general RA population. These limits either the testing of the intervention or the ability to generalise to the RA population seen in clinical practice (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>Stratification was utilised in the study for age (whether above or below 50 years of age) and sex. It would, however, also have been important to consider stratification for patient ability that related to the outcomes in terms of the dynamic RA characteristics such as functional ability, pain, and ADLs. This would have ensured that the outcomes were not prefaced by a significant difference in ability between the two groups of patients, with the risk of a type II error in accepting the results of the study (Bartley and Fillingim 2013; de Boer *et al.* 2015). It is reported that randomization was utilised, but the method was not described nor was it indicated whether there was concealed allocation. Additionally, the study failed to utilize an *a priori* analysis, putting itself at risk of being underpowered. This would imply that the outcomes would not be categorical but rather based on trends (Hoskin 2012; Bahreini *et al.* 2020) and that they may actually not be better than chance alone. The study also does not control for dropouts and does not utilise the intention-to-treat analysis which would only have applied to one participant (due to the available data at time of dropout). But in defence of the study, data analysis for 31 participants per group does allow the use of parametric statistics which in part negates the trend analysis from an underpowered study (Mouton 2006; Christley 2010) but does not assist in ensuring that the outcomes are better than chance alone.<br><br>The study makes reference to medication control around glucocorticoids, DMARDs and Alendronate. None of the participants were on glucocorticoids or DMARDs prior to the study starting. Later the study states that DMARDs were instituted after the initial measurements. All patients were given sulfasalazine, except for one patient in each group who was given gold sodium thiomalate because of allergies to sulphonamides. During the study, the DMARDs were changed in 15 to patients in the strength-training group and 19 in the control respectively (by reason of inefficacy and / or adverse events). Low-dose prednisone was administered in the last 12 months to two patients in the strength-training group and 10 in the control group. Additionally, if the BDM measures revealed osteopenia, alendronate was provided, as in the case of three patients in the strength-training group and nine in the control group). Notwithstanding the concerns around the baseline readings, the outcome measurement tools were a suited (Coster 2013), reliable and validated (Coster 2013; Bahreini *et al.* 2020) mix of subjective and objective outcome measures for use in measuring and detecting change for the outcome in the context of RA (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018).<br><br>Some of these outcome differences, like the medication, showed no significant difference between the groups when controlled for statistically. This detail is not offered for all medications even though RA is commonly treated with a cocktail of medications (Zhao *et al.* 2017; Heidari *et al.* 2019). This is equally applicable to medication changes, dosage adherence, type of medications and the method of administration. This discussion is equally applicable to the adherence to physical activity-based interventions. The authors acknowledge this and report exercise adherence in the experimental group (70–75%) but this is unreported in the control group. These outcomes were assessed utilizing training diaries which improve compliance (Fogg and Gross 2000). Of overall concern as regards the lack of comparability between the groups and controlling statistically is that this is not the preferred method of controlling for these impacts (de Boer *et al.* 2015). |

| | |
|---|---|
| | Additionally, the study does not control for the placebo effect, patient naivety, the Hawthorne effect, and the timing of measures (the diurnal impact) (Kaptchuk 1998; Pater *et al.* 1998; Mouton 2006; Richardson 2007; Bialosky *et al.* 2011; Sedgwick and Greenwood 2015; Nandgude, Hasabe and Kolsure 2018).<br><br>Another important area in which the study comes short is the lack of clarity as to the exact exercise protocol or whether the protocols were open for individualisation based on the RA patient's ability. This, together with the fact that the measurement assessor was not blinded introduces bias and a concern that the outcomes are indeed reflective of only their clinical impact. Compounding this is the lack of clarity around the time relationship between intervention and measurements. |
| Outcome | The authors conclude that regular exercise programme (including dynamic strength and endurance exercise) improved aspects of muscle strength and physical function, however, it fails to impart therapeutic effect on BMD in patients with early RA, while according to the authors showed no detrimental effects in terms of disease activity. The data reported in this study reflects, that participants engaging in 1.4 to 1.5 sessions per week of said programme, gained significant improvements in maximal strength of all assessed muscle groups by between 19 and 59% ($p$=0.002 – 0.025); the improvement in HAQ scores at 18- and 24-months from baseline ($p$=<0.01 and $p$<0.05 respectively), walking speed at 24-months improvement from baseline ($p$<0.001), and disease activity outcomes (ESR at six- and 12-montsh ($p$<0.05), at 18-months ($p$<0.01); pain at 18- and 24- months ($p$<0.05); morning stiffness at six- and 24-months ($p$<0.05); DAS score at six- and 18-months ($p$<0.001) and at 12-months ($p$<0.05)). Despite improvements in the same parameters seen in the control they were did not achieve the same degree of improvement. With regard to BMD the strength training group did demonstrate increase in spinal (1.17 +/- 5.34%) and femoral sites (0.51 +/- 1.64%). The control however, demonstrated loss in BMD of the study period. A sub-group of the sample (17/62 participants) at baseline suffered higher disease activity and utilised higher dosage of steroids. This is evident in the consistently lower BMD of the femur (at baseline [$p$=0.006], at 12- and 24- months [$p$=0.002]) (Häkkinen *et al.* 2001). |
| Discussion | The study's significant lack of clarity in terms of group homogeneity, the lack of medication control, protocol adherence and unblinded assessment of the participants arouses the concern as to whether the outcomes reflect the actual clinical impact of the intervention or the effect of these confounders. Accordingly, the reader is unable to identify accurately the RA patient who would benefit from a strength and exercise protocol, which limits the generalisability of the study. This is further impacted in that the vague description of the strength and exercise protocol does not allow a practitioner to implement the latter effectively for a patient. Of It is of interest, however, that the outcomes of the study contradict the findings of another exercise study in which the BMD responded favourably to an exercise protocol (Heine *et al.* 2012), thus supporting the impact of the various extraneous variables that have not been controlled for in this study. Furthermore, outcomes for BMD were derived from the sample, of which 17/62 carried higher baseline disease activity and higher therapeutic steroid use, these participants are compromised within the remaining sample as these confounding factors must be kept in the for fount when reviewing the outcomes in this outcome. |
| Conclusion: | Given the limitations discussed above, the external validity of the student ranks with a moderate-to-high risk of bias. The study suffered significant dropout and underpowering of the study outcomes. The study does make use of intention-to -treat which on the one hand recovers the sample size but has the potential to enhance outcome reporting and provide misleading information, particularly when the reader is not privy to the method by which this is achieved. Along with this, the study suffers a number of external limitations, and the reader can therefore not exclude reporting on by chance or natural history changes in the outcomes. The study's sample recruiting method risks representing a sub-population of the RA population, for which reason the reader cannot evaluate the unknown and uncontrolled for influences thus introduced into the study, additionally, the recruitment method limits the generalisability of the outcomes significantly. Finally, the randomisation method and the level of concealment achieved in the study cannot be determined. If combined with the PEDro review, the reviewers 97% agreement ranked the study as 6/11, suggesting that the study presented with moderate levels of bias and moderate levels of methodological rigor / internal validity. Therefore, when considering the study's rigour (internal and external), the study is subject to a moderate risk of bias, and at best provides a moderate-to-low level of evidence in favour of exercise for RA patients who comply with the patient description available. |

Table 4.17a: Tabulated feedback of reviewer data

| Author/s | Häkkinen, A., Sokka, T., Kautiainen, H., Kotaniemi, A. and Hannonen, P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2004 | | | | | |
| Title | Sustained maintenance of exercise-induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5-year follow-up. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | No | No | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least one key outcome | Yes | No | Yes | Yes | 67% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | Yes | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 7/11 | 7/11 | 7/11 | |
| Overall percentage agreement | | | | | | 88% |

Table 4.17b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Häkkinen, A., Sokka, T., Kautiainen, H., Kotaniemi, A. and Hannonen, P. |
| Year | 2004 |
| Title | Sustained maintenance of exercise-induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5-year follow-up. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Maximum unilateral knee extensor concentric strength<br>- Maximum trunk flexor and extensor isometric force<br>- Grip isometric strength<br>- Bone Mineral Density (BMD)<br>- Larsen method joint assessment<br>- HAQ<br>- DAS28<br>- Pain VAS<br>- Morning stiffness VAS | Baseline (of two-year intervention period), months six, 12, 18, 24.<br>Follow-up at five-years (60 months)<br>BMD and Larsen method joint assessment were measured at baseline, months 12, 24 and 60. | Five-years: two-year intervention period followed by a three-year follow-up. | Total of 70, experimental group (n=35) and control (n=35).<br>Loss during intervention period: experimental group (n=2) and control (n=3). Completers of two-year intervention 31 each. And the five-year follow-up experimental group (n=29) and control (n=30). | Yes, for the clinical status and radiological assessments.<br>For the remaining assessments it is unclear. | Yes, the control, however, received conventional training during the first two-years after which they received the experimental programme and additional instructions as did the experimental group. The control also had RA as per recruitment criteria. | Yes, "using clusters of four patients stratified according to age and sex". | 7/11, indicating a moderate to high methodological rigour and subject to moderate to low risk of bias within the internal validity of the study. | 88% |

| | |
|---|---|
| Limitations | The sample was recruited by referral from specialists to the Jyväskylä Central Hospital Rheumatology Unit, Jyväskylä, Finland. This implies that all RA patients received specialist care on diagnosis, which limits the likelihood of novel interventions at the start of the study (Wells *et al.* 2012). The inclusion criteria required confirmed RA (ACR 1987 criteria (Arnett *et al.* 1988)), with disease duration no more than 24 months and no DMARD therapy or prednisolone on entry into the study. These criteria, although controlling for confounders like medication (type) and duration of disease, do not control for the following: occupational status, surgical interventions (Naqvi, Hassali and Aftab 2019), use of other interventions (Mouton 2006), contraindication to intervention (Durcan, Wilson and Cunnane 2014), lifestyle factors (Pedersen *et al.* 2006a; McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019), medication (dosage, administration, changes and adherence) and endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018). Notwithstanding the criteria not controlled for, the baseline characteristics were not significantly different for sex (stratified), age, weight, height, disease duration and serology (RF) but were significant for active smoker status of the participants. This increases concerns for unaccounted variables that may increase the risk of type-II error (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>The 70 participants were randomised into two groups, by means of a clustered four-participant stratification utilising age (older or younger than 50) and sex (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019). This method also provides a degree of concealment (Mouton 2006) and protects against bias associated for non-comparability of groups (de Boer *et al.* 2015). As a result, the study produced an experimental and control group each with 35 participants each. During the two-year intervention period the study suffered dropouts. These group sizes were not based on a power-analysis and subjected the study to underpowering (Kendall 2003; Christley 2010), reducing the reliability of and confidence in data and making the conclusions problematic (Fogg and Gross 2000; Mouton 2006; Christley 2010), especially since intention-to-treat analysis was not employed (de Boer *et al.* 2015; Unbiased Research 2017).<br><br>The intervention was conducted over a two-year period (allowing for physiological manifestation and measurable changes due to the intervention (Fraser 2008)). The experimental group underwent a prescribed and supervised strengthening programme well defined in the study (Heine *et al.* 2012). There were, however, modifications and technique assessments (at month six, 12, 18 and 24) which introduced intra-group variability (i.e., participant- tailored). This is important for patient participation and adherence and allows continued physical loading but is problematic for the research paradigm, as the group becomes a case series within a non-case series study. The control group received a conventional training programme (ROM exercises aimed at all extremity joints and the neck, with these stretches included). The protocol was presented, and the prescribed frequency was twice a week (Heine *et al.* 2012). After the two-year comparison period, the control completers were instructed to additionally undertake the strengthening exercise programme that was carried out by the intervention group. This then changed how the two groups compared at the end of year two by comparison with the end of year three. This was further complicated by the lack of formal structural differences between the groups with the passage of years up to year five, which makes the gains achieved at different time |

| | |
|---|---|
| | points more variable for different reasons that are not fully explained in the article. Another difference between the groups is that the control group received a single instruction period, underwent no check-up assessments and the participants completed aerobic exercise (of unknown duration and intensity) two to three times a week. This introduced further variance between the groups and within the group's participants, which was compounded by lack of progression control and proper technique assessment as well as failure to assess change in outcome measures over time. Compliance to intervention (in first two-years) averaged 1.4 -1.5 times per week (experimental group), with an average 245 minutes/week (SD 115), compared to the control with an average of 195 minutes/week (SD 104). This 50-minute variance could induce change between the groups quite aside from the intervention. The five-year averages were reported as 237 minutes/week (SD 147) for the experimental group and 223 minutes/week (SD 131) for the control group, which is a reduced difference but then the control group was doing two sets of exercises which may account for the change in time. Of the participants 32 (20 experimental and 12 controls) participated in occasional "intensive aerobic exercise" (clarity of 'occasional' is a problem), and only 10 participants (all of them experimental participants) continued the strength-training programme during the follow-up period. The assumption must be made that this was based on the retrospective questionnaire covering the last 12 months that was given for completion at final follow-up.<br><br>Outcome measures and assessors: A positive outcome for the study is the balanced use of subjective and objective tools (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), including muscle strength outcomes (Häkkinen, Komi and Kauhanen 1987; Mathiowetz 1991; Rantanen, Airaksinen and Penttinen 1994), BMD, joint status (Larsen, Dale and Eek 1977; Larsen 1995), functional capacity (Finnish version) (Fries *et al.* 1980), pain and morning stiffness (VAS), and physical activity recorded in the training diaries utilised in the first two-years of the study, as well as a questionnaire (of unknown design, reliability and validity). The lack of calibration of the objective measures could introduce bias (Kimberlin and Winterstein 2008; Coster 2013), whereas the subjective outcomes could fall prey to the issues of memory recall bias (Mouton 2006; Ricker, Vergauwe and Cowan 2016) given the length of time of this study. Conversely, the use of blinded assessors for clinical and radiological outcome measures improved the study by reducing bias (Kendall 2003; Moustgaard *et al.* 2014), although it is unclear if the study used a single assessor or multiple assessors (Kendall 2003) or if assessor training took place in order to control for inter-rater variability bias. Lastly, the environment (Richardson 2007) and timing (Nandgude, Hasabe and Kolsure 2018) was not indicated. Given that the BMD and Z-scores were the only ones reported as not being different between groups, it is not possible to determine the impact of bias for the other measures, even though there were detailed assessment protocols for all measures (Kendall 2003). The study utilises a sample size large enough to accommodate parametric analysis to be conducted (for all assessment periods), even though a power analysis does not provide evidence that the power is appropriately powered. These statistics are also then confounded by the lack of baseline comparability of all but the BMD and Z-scores, which necessitated the need for the statistics to be controlled by variables that were equitable at the outset such as age, sex, disease activity (DAS), and the use of prednisolone. This can be seen in the BMD and Z-Scores that were not different at baseline, but on subsequent analysis were seen to be different, which may have been attributed to the intervention, but after adjusting with the values of age, sex, disease activity (DAS), and the use of prednisolone, this difference in the BMD and Z-Scores was eliminated. This is a good example of the type II error. Additionally, because of the change in interventions at the two-years interface, gains made in the two years are lost by the fifth-year follow-up. This may be a result of changing exercise programmes, deconditioning due to non-adherence (less likely as time spent exercising went up) or patient enthusiasm being reduced, but also not excluding the possibility that a reduced number of participants participated over time (comparing data initially larger groups to smaller groups (as a result of dropout) is comparing only those patients in the smaller group who did not get worse due to RA flare-up (impairing exercise) to a larger group that was more moderate, which is not fair and equitable). Overall, the authors acknowledge that a major limitation in this study was the requirement for more active treatment in the control group, which made it impossible to make clear causal links between outcomes and various interventions the various stages. |
| Outcome | The patients' exercise-induced muscle strength gains during a two-year training period were maintained throughout a subsequent self-monitored training period of three years. Despite substantial training effects in muscle strength, BMD values remained relatively constant and radiographic damage remained low even at the five-year follow-up. Results of the follow-up reflect that maximal muscle strength noted at the post-intervention (two-year) assessment indicated an increase from baseline(Häkkinen *et al.* 2004). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though the sample size may have resulted in under powering of the outcomes and thus the conclusions, with an increased likelihood of a type II error. This limitation, compounded by other limitations noted above, suggest that the outcomes may have been affected by factors such as exercise protocol changes, duration of exercise, RA dynamic characteristic variables not taken into account, baseline differences between outcome measures that were controlled for statistically, quite aside from the possible impacts of natural history of the condition, chance and time. These negatively affect the accurate and fair direct measures of the interventions to each other. The vague participant description further prevents generalisability and implementation into clinical practice. |
| Conclusion: | Given the limitations discussed above, the external validity of the student ranks with a moderate risk of bias. The PEDro review, the reviewers by means of an 88% agreement ranked the study 7/11suggested that the study presented with moderate levels of bias and moderate levels of methodological rigour / internal validity. Therefore, when comparing the internal and external validity outcomes, it is likely that the study presents on average with a moderate risk of bias. Thus, collectively the study provides a moderate level of evidence in favour of exercise for RA patients that comply with the inclusion criteria of this study, but outcomes are not clear. Nevertheless, it also implies that the study should be repeated with clearer delineation of the exercise protocols at different time points, which are congruent with the outcome measures so that the analysis can more accurately identify the cause of change. |

Table 4.18a: Tabulated feedback of reviewer data

| Author/s | Hamilton, D., Bywaters, E. and Please, N. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1959 | | | | | |
| Title | A controlled trial of various forms of physiotherapy in arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | No | Yes | Yes | 67% |
| 3 | Allocation was concealed | No | Yes | Yes | Yes | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | No | No | No | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | No | No | No | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 5/11 | 5/11 | 5/11 | 6/11 | |
| Overall percentage agreement | | | | | | 91% |

Table 4.18b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hamilton, D., Bywaters, E. and Please, N. |
| Year | 1959 |
| Title | A controlled trial of various forms of physiotherapy in arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - ROM (Knee and Wrist)<br>- Power (Knee extension and Hand)<br>- Time to walk up and down 4 steps<br>- Time to walk 20 yards<br>- Time to tie 3 knots. | (Weeks pre – reading)<br><br>Baseline (0), treatment 1 (1, 5), treatment 2 (6, 10), treatment 3 (11, 15), treatment 4 (16, 20), post reading after all treatments (21). | Five-months | 131 starting<br>64 drop-out<br>67 completed (five excluded). | No | Pre/post study used the participants as their own control, but then aggregated the means of the pre and compared to the means of the post for each group | Not for participant allocation<br><br>Randomised in regard to treatment sequence. For each of the three diagnostic categories (RA hand, RA knee, degenerative joint disease knee), a series of envelopes were prepared listing the order of these four treatments and sealed. | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 91% |

| | |
|---|---|
| Limitations | The sample for this study was derived through consecutive referrals of outpatients for physiotherapy in said department, of the Hammersmith medical school hospital (London, United Kingdom). Convenience sampling raises the risk of recruitment of a subpopulation easily accessible to the research team and limits the generalisability of the outcomes. Furthermore, the source of recruitment implies that the subpopulation is one seeking medical care and thus likely to be experiencing exacerbation of their RA, but when the reader considers the chronicity of the condition and the recruitment source, the participant enrolled is far more likely to lack naivety and novelty to the study interventions employed (Wells *et al.* 2012; Jager, Putnick and Bornstein 2017).<br><br>It must be noted that this study was conducted in 1959 and consequently more recent guidelines and diagnostic criteria would not have been available to the researchers. Given this, the inclusion criteria are limited to a RA diagnosis according to the rheumatism unit (Hamilton, Bywaters and Please 1959) in which RA was assumed a polyarthritis of unknown aetiology with pain, swelling, limitation of joint movement and a raised sedimentation rate. It was differentiated from degenerative joint disease (DJD) through accepted clinical and radiological evidence with a normal sedimentation rate. In view of the limitations of this classification, the criteria did not include or control for the following factors on entry or by stratification or by set of recruitment criteria: RA serology, RA functional class, RA duration and activity of disease, flare-ups, use of other physical modalities, contraindications and / or comorbidities (England *et al.* 2019; Myasoedova *et al.* 2020). The study states that baseline characteristics were similar between the groups but does not detail these characteristics in order for the reader to understand which were evaluated. This does not allow for comparisons to be drawn with the general RA population, a shortcoming that is compounded by the fact that all patients in this study were drawn from one rheumatology unit (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>The study does not report sampling by calculated *a priori* analysis, with a total sample size of 67 with 18 patients in the RA knee group, 33 in the RA hand group and 26 in the DJD knee group. Patients were allocated per involved joint and not stratified or randomly allocated (Mouton 2006; de Boer *et al.* 2015). The significant dropout rate did predispose towards the alteration of the randomly assigned treatment group's allocation in that it was possible that more patients received a particular sequence of treatments which then negated the effects that the randomisation was to attain. Convenience sampling and the significant attrition, and then negation of treatment randomisation are all significant areas of bias that occur in the study, thus decreasing the possibility that the outcomes are only as a result of the interventions (Heritier, Gebski and Keech 2003; Mouton 2006; Wells *et al.* 2012; de Boer *et al.* 2015; Jager, Putnick and Bornstein 2017).<br><br>The study focussed specifically on the knees (group 1) and hands (group 2) of RA patients and the knees (group 3) of DJD patients. If other joints of these patients were also affected, they were treated but not assessed in terms of the study outcomes. Treatment focussed on four treatment modalities (short-wave diathermy, infrared radiation, faradism for knees, wax baths for hands and cold short-wave diathermy as the control applied for a period of four-months with each treatment being applied for one-month of the four. The application sequence for each patient was randomised and concealed (Kendall 2003; Mouton 2006), so in effect the study had individualised care for the |

| | |
|---|---|
| | patients. A one-week wash-out was in effect between the four-months of treatment and in this time the post-measurement of the previous treatment and the pre-measurement of the following treatment were taken (Pater *et al.* 1998; Platt, Schisterman and Cole 2009; Evans 2010). This therefore does not negate the possibility of compounding treatment effects as a result of carry over (Bordens and Abbott 2002) as the sequence of the treatments may be synergistic or antagonistic in nature (Robins, Hernán and Siebert 2004) which in either case would invalidate the gains or losses respectively as they cannot be linked exclusively to the intervention. In addition, the monthly improvements noted by the authors with the different interventions would have led to a possible ceiling or floor effect (Liang 2000; Coster 2013), where it was less possible to attain significant outcomes in measures. This would have resulted, or potentially resulted, in the effects of the latter treatments being seen as ineffective or not significant which, if taken out of context, could predispose to a type II error in accepting the results (Bartley and Fillingim 2013; de Boer *et al.* 2015). Additionally, each group was provided active exercises (assisted or resisted), calcium aspirin or codeine and splinting as necessary. This pragmatic approach, although more reflective of clinical practice, does provide a fundamental problem as this study structure is more analogous to that of a prospective case series; the greater the number of treatment combinations, the greater the chances that this study is not reporting on the effectiveness of any one intervention or even attempting to develop an outcome reflecting the relative effectiveness of the treatments for the RA hand or knee and the DJD knee (Robins, Hernán and Siebert 2004). Furthermore, the structure of this study in no way attempts to control for patient effects (see Section 4.4.1).<br><br>The lack of assessor blinding does not help to mitigate the researcher bias and measurement bias (Kendall 2003; Moustgaard *et al.* 2014). It is unclear whether all measures are suitable, valid, and reliable for the measurement of both the RA and DJD in these patients, which is of particular importance as the patients were their own controls and therefore needed to ensure that these measures were able to detect small changes (Coster 2013; Bahreini *et al.* 2020).<br><br>The statistics utilised in this pre-post design study were not articulated and therefore, also taking into account that the group sizes varied from 18 – 33 patients, it is difficult to ascertain whether the study applied parametric or non-parametric statistics. Additionally, the tables presented are not entirely clear and the reader is left to make assumptions (Hoskin 2012; Bahreini *et al.* 2020). |
| Outcome | It would seem that within the monthly treatment regimens that only the faradism for the RA knee achieved a significant outcome (as reflected by the time to walk 20 yards), with the random combination of treatments over the five-month-long study showing improvement for the RA knee (in terms of power in knee extension), RA hand (in terms of the time taken to tie three knots) and DJD knee (in terms of the power in knee extension and stair climbing). The study claims the significance listed, while failing to provide *p*-values (Hamilton, Bywaters and Please 1959). |
| Discussion | The lack of control for confounding variables in the study (as is noted above) suggests that the outcomes may be the result of chance rather than a result of the interventions. Compounding factors were the study's analogous design characterised by with the following: an apparent prospective case series presentation and application, limited control of the patient effects, effects related to lack of concealed patient allocation, effects related to f a lack of assessor blinding and the continued impact of floor and ceiling effects as the study progressed. Collectively, these suggest that the outcomes are consequently fraught with problems in terms of their reliability and the questionable accuracy in measuring them. A further complication is the lack of an *a priori* sample size to provide adequate power to the study and therefore to support of the outcomes. The study compounds the bias by failing to report actual *p*-values calculated for each of the reported significant outcomes, thus eroding the credibility of the findings and limiting the reader's ability to contextualise the findings accordingly. Additionally, this the study is not able to comment on any one individual treatment intervention, which for this systematic review would have been exercise, given that the exercise was a supplemental adjunct to the active interventions that were applied as part of this research. |
| Conclusion: | Given the above analysis, from the external validity vantage point this study suffers from a high degree of bias and the reader cannot exclude the effect of external influence/s. This does not find itself congruent with the PEDro review by the reviewers, which ranks the study as a 6/11 with 91% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study. Disjuncture between these two reviews is an indicator of bias and confounding influencing outcomes from sources outside of the research process/methodological structure. This review, despite the disjuncture, can only conclude that at best the study provides the reader limited to no evidence. Should value be drawn for these interventions in the management of RA further research would have to be carried out addressing the reported limitations. |

Table 4.19a: Tabulated feedback of reviewer data

| Author/s | Hansen, T., Hansen, G., Langgaard, A. and Rasmussen, J. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1993 | | | | | |
| Title | Long-term physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | No | No | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | No | No | No | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by intention-to-treat | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 7/11 | 6/11 | 7/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.19b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hansen, T., Hansen, G., Langgaard, A. and Rasmussen, J. |
| Year | 1993 |
| Title | Long-term physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total % Agreement |
| - Duration of morning stiffness<br>- VAS joint pain<br>- Number swollen joints<br>- The treatment and medical expensive<br>- HAQ<br>- ESR and Hb<br>- Aerobic fitness<br>- Functional score<br>- Max isometric muscle strength (extensors of the knee and elbow, abductors of the shoulder)<br>- Larsen joint evaluation (hands, wrists, and feet)<br>- Patient questionnaire (evaluation of the training)<br>- Interviewed about physical activities during the study. | The doctor evaluations were every three months<br><br>The physiotherapist evaluated ever six months<br><br>The radiologist evaluated every 12 months | Two years | 75 patients eligible allocated randomly to five groups (15 per group). Of these the completers included self-training group (14), training in physiotherapists practice (14), group training (11), group training + pool (13), and no training (13) | Yes: the evaluating doctor, physiotherapist and radiologist are all blinded. | Yes, a no training group. | Yes, method is unclear. | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study recruited all participants from the Rheumatology Department of the King Christian X's Hospital in Denmark. Due to participants being sourced from specialist Rheumatology departments in the region (South Jutland), a concern around the perceived convenience sampling, is poor control for confounding factors introduced through this process. For example, the study population potentially reflects a subpopulation of RA, likely to suffer more severe RA than the general RA population (Wells *et al.* 2012). As a result, risks unintentionally limiting the generalisability of the study, and introducing confounding factors and bias that the study design is not controlling for. In such a scenario the outcomes reflect external influence rather than intention. Furthermore, such a sample is unlikely to be naïve to or consider the intervention novel (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013).<br><br>In this study the inclusion criteria. based on the ARA criteria for RA (Ropes *et al.* 1959) were as follows: functional stage I and II (Steinbrocker, Traeger and Batterman 1949), age between 20 and 60 years of age, no contraindications to physical training, and training less than three times per week. These broad inclusion criteria did not consider RA activity, use of other treatment modalities, comorbidities and/or contraindications. As a result, the baseline similarities in each of the groups may be significantly different as the participants were only compared in terms of the demographic characteristics of age, sex and work status as well as disease duration, serology (RF), joint erosions on x-ray, medication (slow acting drugs and glucocorticoids)). Given the table of comparison between the groups it is not clear whether there were any significant differences between the groups in terms of these baseline characteristics and whether or not homogeneity was attained. (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>Thus, although randomization was utilized in the study (the method not specified), it is unclear what bias this procedure brought into or provided control over the groups with regard to the baseline characteristics (Fogg and Gross 2000; Mouton 2006; de Boer *et al.* 2015; Mansournia *et al.* 2017). This is particularly true as stratification was not utilized to ensure homogeneity between the groups, notwithstanding the clinical complexity of RA (McInnes and Schett 2011; Yang *et al.* 2017). A further complication is the lack of *a priori* analysis to determine the minimum group sizes for the study to be adequately powered (Kendall 2003; Mouton 2006; Christley 2010). The authors claim that "it can be argued that the number of patients in each group was relatively small. The use of a two-way analysis of variance on the results from the 65 patients, however, should have detected any clinically significant effect of training during the two-year study" (Hansen *et al.* 1993). Nevertheless, this is contradicted later in the study when the authors state that "the statistical power was limited due to the small number of patients" (Hansen *et al.* 1993). The reality is that without a power-analysis the study design cannot detect either the statistically or the clinically significant changes in outcome measures (Kendall 2003; Christley 2010), which therefore places the study as a whole in risk of underpowered results (Mouton 2006; Christley 2010). The sample allowed for 15 participants per each of the five groups, thus 75 participants in total). In this context non-parametric statistics can only provide trend analysis and not causal links between intervention and outcome (Hoskin 2012; Bahreini *et al.* 2020). A compounding factor is that the study does not make provision for the high dropout rate, common in a physical-intervention long-term study (Fogg and Gross 2000), nor does it utilise the intention-to-treat analysis to protect the study (Gupta 2011). This makes it particularly difficult to determine the power of the study in terms of its capacity to detect statistical and clinical change and how this influences clinical practice. The intervention protocols used in the study (group A = self-training after instructions, group B = same as A with physiotherapy individualised care, group C the same as group A with groups training in a hospital setting) were all described |

| | |
|---|---|
| | and can be replicated except for the written instruction regarding the 15-minutes overall training programme given to all groups other than the control group. The control was a no intervention group which formed the base to determine the efficacy of the other interventions.<br><br>The subjective and objective outcome measures provided a range of measures to account for external influences (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Additionally, the assessors, a doctor, a physiotherapist, and a radiologist, were all blinded (Kendall 2003; Moustgaard *et al.* 2014), with a single evaluator for specific outcomes, thus providing consistency (Kendall 2003). The tools utilized are established, validated, and recognised in use for RA (Coster 2013; Bahreini *et al.* 2020), but the questionnaire and interview have no clear validation. The outcome measures (ESR, Hb, number of swollen joints, pain, morning stiffness, HAQ, aerobic fitness and medicine cost) were all comparable between the groups, while X-ray scores ($p<0.05$), functional score ($p<0.01$) and max isometric muscle strength of knee extensors ($p<0.01$) were statistically significantly different. These differences in outcome measures could put the study at risk of type II error) (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). The effect of the relationship between intervention and evaluation (not detailed in this study) questions what is being measured (e.g., a week post the last intervention allowed for programme outcomes to be measured while a follow-up a month later indicated short term effect). Medication use, which can have a significant impact on outcomes (Zhao *et al.* 2017; Heidari *et al.* 2019), was marginally controlled. At baseline, the distribution of slow-acting drug use and glucocorticoid use was reported (Table 1 of the publication) and as part of the outcomes medication cost over three months was monitored. Aside from this, however, other medications used (NSAIDs and DMARDs), the method of administration, changes during study, adherence, the dosage, and specific type were not commented on.<br><br>There is little control of the patient effects (reference section please), among them the touch effect (Melzack and Wall 1965; Moayedi and Davis 2013; Senderovich *et al.* 2016), the placebo effect (Kaptchuk 1998; Bialosky *et al.* 2011), the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015), particularly with regard to different levels of interaction with research staff in different groups and environmental factors such as different training sites (Richardson 2007). These factors would need to be controlled in order to minimize their impact on the outcomes of the study, particularly the subjective outcomes. Alternatively, change in objective and subjective findings could be used to determine the impact of the patient effects, but given that there are significant baseline differences in at least three of the four objective findings this is difficult to assess (de Boer *et al.* 2015). Lastly, patient naivety is not excluded in the study or controlled for, as is seen from the source of recruitment and three of the interventions were based on three types of training offered to RA patients by the Dutch health system. This therefore does not exclude participants' previous experience and its influence on the outcomes (Mouton 2006). Additionally, the study notes of a number of concurrent treatments across the groups such as synovectomy, intraarticular steroid injections, knee joint replacement, foot surgeries and other exercises in the self-treatment group, but the impact of these interventions on the outcomes is not controlled for statistically or analysed (de Boer *et al.* 2015). However, the intervention effect is accounted for in that exercise, which requires six to eight weeks to impact functional capacity and thus ADL and therefore outcomes (Fraser 2008) would have seen changes in the two-year duration of the study, though exercise adherence would have influenced the degree of change (Fraser 2008; Heine *et al.* 2012). To this, a strength in the study is incorporation of exercise diaries that were asked of the participants to log exercise compliance, this would also aid in memory recall during the questionnaire (Mouton 2006; Ricker, Vergauwe and Cowan 2016). |
| Outcome | The authors indicate that, irrespective of group, there was no effect from training (although most patients were in favour of training) on the disease activity / disease progression. This was attributed in the study to a lack of detectable effect resulting from the small and underpowered sample size, and possibly due to an insufficiency in the type or intensity of training with regard to the type of selected exercises and / or a change in the assessors' awareness of participant groups which, given the nature of the evaluation and intervention, could thus have influenced the study (Hansen *et al.* 1993). |
| Discussion | This study was a good attempt to address the possible long-term outcomes of activity and exercise on RA, even though it was compromised by underpowering and its attendant effects, along with the lack of control for patient effects, all of which influence internal validity. In addition, it is also plagued by the lack of control for baseline characteristics as well as baseline differences in most of its objective outcome measures, which limits its ability to draw conclusions. It is therefore not unexpected that the study results are inconclusive. It would be valuable to repeat the study with larger sample sizes so as to address the concerns raised with regard to improving the ability to measure the actual impact of the interventions. |
| Conclusion: | Given the significant limitations noted above, from the external validity vantage point this study suffers from a moderate to high degree of bias and the reader cannot exclude the effect of external influence(s) on the outcomes attained. This is somewhat contrasted by the review by the reviewers, which ranks the study as a 7/11 with 94% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study. This highlights a degree of disjuncture between the two reviews. It is, however, possible to have strength based in the rigour of the study, but the bias and confounding seen in this analysis could be induced by influencers outside of the research process. Therefore, this study at the best has a moderate to low evidence and until further research can be undertaken to address the above limitations, this study offers tentative foundations from which further research can be carried out; however, it falls short in offering the needed evidence to be used as a basis for the management of RA |

Table 4.20a: Tabulated feedback of reviewer data

| Author/s | Harkcom T.M., Lampman R.M., Banwell B.F. and Castor C.W. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1985 | | | | | |
| Title | Therapeutic value of graded aerobic exercise training in rheumatoid arthritis | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | Yes | Yes | Yes | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | Yes | No | No | 67% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | No | Yes | 67% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| | Total score | 7/11 | 10/11 | 7/11 | 9/11 | |
| | | | | Overall percentage agreement | | 85% |

Table 4.20b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Harkcom T.M., Lampman R.M., Banwell B.F. and Castor C.W. |
| Year | 1985 |
| Title | Therapeutic value of graded aerobic exercise training in rheumatoid arthritis |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Modified Functional Status Index (FSI)<br>- Global Assessment<br>- Grip strength<br>- 50-foot walk time<br>- Muscle strength (Isotonic extension, Isotonic flexion, Isometric extension)<br>- Joint count (pain and swelling)<br>- Exercise tolerance test (aerobic capacity ($VO_{2max}$), ECG, exercise pressure rate [heart rate + blood pressure], exercise test time) | The FSI outcome measure was recorded every 2 weeks. Global assessment was carried out only at post-intervention. The remained of the outcomes followed a pre/post design. | 12-weeks. | At baseline 20 were recruited, intervention groups were populated by; group A (n=five), B (n=four), C (n=five), with the control (n=six).<br><br>During the study three dropped out, one per intervention group. Thus completers (n=17) → group A (n=four), B (n=three), C (n=four) and control (n=six). | No | Yes, non-training control group | Yes (patient blindly chose a treatment time block and the time blocks were randomly assigned a treatment protocol). | 9/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 85% |

| | |
|---|---|
| Limitations | Recruitment from the outpatient care at the Medical School of the University of Michigan, USA, suggests more severe RA patients already attending care, thus decreasing the likelihood of participant naivety (Wells *et al.* 2012). The sample comprised 20 female participants (limiting the generalisability of the study to females) between the ages of 27 and 68 with RA (either definite or classic) according to the 1958 revised RA diagnostic criteria (Ropes *et al.* 1959); a functional classification of class II (according to the ARA (Steinbrocker, Traeger and Batterman 1949). All participants were on stable drug therapy (but a description of the medications and their variables was not provided), and no major changes in medical care were recorded during the study. The RA disease activity was stable, and no participants had acute RA flare-ups (in terms of joint symptomatology). These factors control for major RA dynamic characteristics, which usually impart significant variability in participant outcomes (Kaur, White and Bartold 2012; Walker *et al.* 2013; Bahreini *et al.* 2020). A self-administered questionnaire covered participant demographic information, including socioeconomic status, education level, and disease duration. At baseline the socioeconomic, education, age and disease duration showed no difference. These criteria allow for a clearer generalisability of the findings of the study to clinical practice (Kendall 2003; Wells *et al.* 2012), but do not exclude all variables, e.g., other interventions (Mouton 2006), surgical interventions (Naqvi, Hassali and Aftab 2019), comorbidities (Kłodziński and Wisłowska 2018; England *et al.* 2019; Levitsky *et al.* 2019; Myasoedova *et al.* 2020), concomitant diagnoses (Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011; Dagenais and Haldeman 2011), contraindication to intervention (Durcan, Wilson and Cunnane 2014), endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018) and ethnicity (Bartley and Fillingim 2013). All of these may result in study bias (Kendall 2003). Once chosen, the participants were randomly assigned to one of four groups (three intervention arms (group A-C) and a control (group D)) by means of participant selection of a time block best suiting their schedules so as to improve adherence. After this self-selected grouping, the treatment protocols were randomly assigned to the time-blocks (possibly a method of random allocation) (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019). The lack of clear description of how to achieve this limits the reader's ability to consider the quality of randomisation and concealment (Mouton 2006). After the random allocation, 'the statistical strength of the study needed to be evaluated. In this case, the authors failed to use a power analysis calculation (Kendall 2003; Christley 2010), placing the study at risk of being underpowered (Mouton 2006; Christley 2010). Additionally, no dropout prevision was made (Fogg and Gross 2000) and the study lacked intention-to-treat analysis which, although good for making up sample numbers, tends to result in outcomes regressing to the mean (Fogg and Gross 2000; Heritier, Gebski and Keech 2003; Kendall 2003; Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017). Lastly, the significantly small sample size limits the study to reporting non-parametric statistics (Hoskin 2012; Bahreini *et al.* 2020) and thus only trends of outcomes.<br><br>The intervention was a group-based, supervised bicycle ergometer programme carried out three times per week for twelve weeks. The protocol was applied to intervention groups A-C with clear explanations and reference to important exercise prescription variables (Heine *et al.* 2012). The group protocols differed in length of exercise session |

| | |
|---|---|
| | at starting point (15 minutes in group A, 25 minutes in group B, and 35 minutes in group C), as well asin progression rate and final length of session reached at six weeks and thereafter maintained to the end of the study. Exercise intensity was determined individually by heart rate (70% of maximum heart rate) and achieved by load applied through resistance settings. All interventions incorporated a warm-up and cool-down component, which included ROM exercises. The groups were unaware of interventions undertaken by the other groups. This intervention duration is appropriate as it allows the time required for physiological manifestation of exercise, and thus the detection by outcome measures (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). In the analysis of outcomes, the control of group baseline distribution (similar for all outcomes) protects the study and allows for measurement of intervention outcomes (Fogg and Gross 2000; Mansournia *et al.* 2017) and avoidance of type II error (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). It must be taken into account that the study combined the total intervention arm of the study (group A, B and C) into a single group (n=11) for statistical analysis against the control for the purpose of reflecting significant improvements but as a result of three different intervention periods. This, however, negates the different treatment periods between the groups. Overall, the study found a decrease in joint count; an increase in aerobic capacity and improved exercise test time, while the control reported no significant changes. The study failed to report MCID values for the outcomes, preventing readers from contextualising outcomes into the clinical setting (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020).<br><br>Aiding study strength was the use of a balance of subjective and objective outcome measurements (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018) that included the following: a modified Functional Status Index (FSI) (Jette 1980), the reliability and validity of which is unclear (Coster 2013; Bahreini *et al.* 2020), the FSI (subjective evaluation of morning stiffness, sleep patterns, self-care, ambulation, and ADL), grip strength (by unknown instrument and protocol), 50-meter walk time (protocol undefined), muscle strength of the knee (isotonic extension and flexion, and isometric extension of the knee recorded with the Cybex II Testing protocol on a dynamometer (Lumex Inc.)), joint count assessment (both pain and tenderness parameters) (criteria not delineated), an exercise tolerance test (on bicycle ergometry for $VO_{2max}$ and exhaled $O_2$ and $CO_2$ by gas meter analysis), cardiac rhythm (electrocardiogram),and exercise pressure rate (blood pressure and heart rate monitor). These assessments and re-assessments were carried out in exactly the same manner, thus controlling for assessment variability and testing protocol (Kendall 2003); however, many of the outcome tools had no cited reliability or validity citations, hindering the reader's clinical interpretation of outcomes (Kendall 2003). The last and most subjective outcome was the global assessment inquiry measuring participants' perceptions of change in fatigue, strength, ability to do housework, participation in social activities, overall mood, amount of joint pain, and tolerance of joint pain. This, unlike the other measures, was administered at the end of the study and is thus at risk of memory recall/decay bias (Mouton 2006; Ricker, Vergauwe and Cowan 2016). Blinding of the assessor and patient and therapist (Kendall 2003; Moustgaard *et al.* 2014) was not stated and thus assumed by the reader to not have been achieved. This imparts potential bias associated with outcome reporting, such as placebo effect (Kaptchuk 1998; Bialosky *et al.* 2011) and Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015). Finally, the environment (Richardson 2007) and timing in respect of the diurnal variability (Nandgude, Hasabe and Kolsure 2018) and the implication of the time lapse between the last intervention and the assessment of the outcomes, reflecting collective intervention effect or the immediate effect of the last intervention (Pater *et al.* 1998) of the assessments, are not detailed; these can introduce bias and confounding risk. The assessors were not blinded, exacerbating the already existing risk of bias. |
| Outcome | This study of female non-acute RA participants with functional classification class II concluded that supervised, group-based exercise programmes incorporating aerobic exercise by means of bicycle ergometry and maintaining 70% maximum heart rate with progression of exercise session length leading to 15-, 25- or 35- minute sessions conducted three times a week over 12 weeks resulted in improvements in aerobic capacity, exercise time, and joint counts. This concurred with subjectively reported improvements in ADL, reduced joint pain and improvement in fatigue. Thus, it was concluded that exercise of this nature can be therapeutic for up to 35minutes without inducing exacerbation of the disease in joint symptomatology and that the addition of a bicycle ergometric for aerobic exercise three times a week for a minimum of 15minutes is sufficient to induce the reported improvements (Harkcom *et al.* 1985). |
| Discussion | The study indicated that the outcomes were favourable, even though it was underpowered and subject to type II error reporting. In addition, collapsing the groups into one analysis negatively affects the accuracy and fair comparison of the three different time intervals between groups A-C and the control and in the current context negates the use of three groups. Additionally, as is noted in the above limitations the study has a range of variables to control for and needs to report a clear participant description with and clear criteria for several of the outcome variables, all collectively limiting the generalisability of the study and the programme into clinical practice. |
| Conclusion: | Given the impact of the above limitations on external validity, the study is at risk f of a high degree of bias and thus questionable outcomes. The reviewers (85% agreement), by contrast, indicate that the study presented with high methodological rigour, allowing for accuracy and clarity because of the utilised methodology. The latter places the study in low risk of bias within the internal validity, whereas the former places the study at a high risk of bias. If aggregated, this conflict results in moderate risk of bias with conflicting basis for evidence. As a result, the study at this point only has the capacity to provide limited evidence for the intervention even in face of the authors' suggestion that the outcomes are valid and applicable. This limits translation into practice settings; therefore, it is recommended that this study be re-done with a larger sample (determined by power-analysis) with multiple different RA presentations, more methodological and systemic strategies for addressing areas of bias so as to focus on the intervention as an outcome. |

Table 4.21a: Tabulated feedback of reviewer data

| Author/s | Hoenig, H., Groff, G., Pratt, K., Goldberg, E. and Franck, W. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1993 | | | | | |
| Title | A randomized controlled trial of home exercise on the rheumatoid hand. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | No | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | Yes | No | No | No | 67% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | Yes | Yes | Yes | 67% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | No | No | No | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | No | Yes | Yes | Yes | 67% |
| | Total score | 5/11 | 6/11 | 6/11 | 6/11 | |
| | | | | | Overall percentage agreement | 85% |

Table 4.21b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hoenig, H., Groff, G., Pratt, K., Goldberg, E. and Franck, W. |
| Year | 1993 |
| Title | A randomized controlled trial of home exercise on the rheumatoid hand. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - Grip strength<br>- ROM<br>- PIP joint circumference<br>- Degree of ulnar deviation of 3rd digit<br>- Clinical impression of hand deformities<br>- Dexterity<br>- Questionnaire | Baseline, post treatment (three-months), follow-up (six-months) | Six-months total (three-month intervention period with three-month follow-up) | 57 participants in total randomised into four groups: group-one (n=14), group-two (n=14), group-three (n=15), group-four (n=11)<br>Dropouts: group-one (n= three), group-two (n= five) group-three (n= five), group-four (n= zero). Completers: group-one (n=11), group-two (n= nine), group-three (n=10), group-four (n=11) | Yes | Yes | Yes, by means of a random number table. | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 85% |

| | |
|---|---|
| Limitations | RA patients of the Bassett Hospital in Cooperstown, New York were approached over five years, in addition to the first 50 RA patients (alphabetically) from a rheumatologist's outpatient practice at the Wilson hospital in Binghamton, New York, to participate in the study. With these sources being specialist centres, it implies that RA patients may have had RA for a longer duration of greater severity when compared to general RA patients (Wells *et al.* 2012). In comparison to prior studies reviewed (Harkcom *et al.* 1985; Hansen *et al.* 1993), where patients were not sourced from different sources, means this study may be more generalizable than the previously reviewed studies. The inclusion criteria required ARA criteria defined definite or classic RA diagnosis within an ARA functional class II or III. Patients were excluded if they had changed medication six-weeks prior to start of the study (providing some degree of control as a washout period (Evans 2010)). There was no full set of exclusion criteria, which introduces variables of bias and confounding (Kendall 2003). These may include receiving other treatment intervention/s (Mouton 2006; Naqvi, Hassali and Aftab 2019), the presence of comorbidities (Kłodziński and Wisłowska 2018; England *et al.* 2019; Levitsky *et al.* 2019; Myasoedova *et al.* 2020) and concomitant diagnoses (Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011; Dagenais and Haldeman 2011), contraindications to intervention (Durcan, Wilson and Cunnane 2014), patient endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018), and general RA disease characteristic factors (Bartley and Fillingim 2013). Along with this, medication was not fully controlled as certain changes in medication were allowed. These permitted included change to NSAID therapy and decreases in steroid or DMARD therapy. However, participants requiring increase in steroid and / or DMARD therapy, were deemed as a protocol violation and thus excluded from the study. The latter was premised on worsening of disease activity in the participants (Kaur, White and Bartold 2012; Walker *et al.* 2013; Zhao *et al.* 2017; Heidari *et al.* 2019; Bahreini *et al.* 2020), but it also stands that interventions may predispose the patient to aggravations (the authors noted that the dropouts had significantly longer disease duration ($p < 0.05$) and reported longer morning stiffness, larger PIP circumference measures, and pain at baseline). Conversely, those nine participants who reported decrease in medication use, may have done so because of positive effects of the intervention. Therefore, it must be guarded against to not only present the good outcomes as this is a form of publication bias. Another factor that complicated the monitoring of medication is the use of patient records and reporting by the participants, resulting in mis-reported data hiding confounding effects of the medications (Zhao *et al.* 2017; Heidari *et al.* 2019). Other reasons for participant exclusion were failure to complete daily exercise for more than four weeks through the study duration or failure to complete exercises for more than two-weeks prior to the second assessment period. These criteria tried to limit non-compliance, as this, risks biasing the outcomes, but may have included bias by excluding the patients. This paragraph thus raises concerns potentially leading to increased risk of type II error (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019).<br><br>The enrolled sample consisted of 57 participants, with a mean age of 57 years, an average disease duration of 11.3 years, right-handed (only one left-handed) (Lee, Gandevia and Carroll 2009; Dragert and Zehr 2013). Twenty participants had participated in exercise (occasionally), 16 maintained full time employment and 46 completed housework on a regular basis. The 57 participants were randomised by means of random number table in blocks of four, into four-groups (three intervention arms and a control). Group-one was allocated ROM exercises; Group-two was allocated resistive exercise with Group-three being allocated with ROM and resistive exercise while Group-four was the control group. The demographic baseline characteristics were not significantly different in terms of age, sex, employment status and housework; clinical factors (RA disease duration, morning stiffness, global assessment, and pain scale) and previous exposure to exercise. |

| | |
|---|---|
| | The "supervised" home-based exercise programme was carried out over 12-weeks, after which an additional 12-week follow-up period occurred. Exercises were completed twice daily for 10 to 20 mins. The level of exercise over 12 weeks was appropriate as it provides ample time for physiological effects of exercise to manifest and thus measure (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). The protocols applied were clearly (orally and verbally) explained and demonstrated by a blinded occupational therapist; with reference to important exercise prescription variables (Heine *et al.* 2012; Moustgaard *et al.* 2014). This was followed by a phone call to patients by an investigator (aware of group allocation) to ensure patient compliance and understanding. The awareness of exercises comprising the programmes increase investigator bias. This may have partially been controlled for with adherence exercise logs, kept by the participants. There are however, also known limitations / inaccuracies of the self-recording in these logs. Safety protocols detailed action for pain exacerbation due to exercise, including set and frequency reduction and when to seek medical care / withdraw from the study. With this in mind the exercises were designed to minimise torque across joints (MCP and PIP) while balancing strengthening of musculature of the hand. Following the 12-week intervention period, all three groups stopped all hand exercises. Group four patients were encouraged to maintain an active lifestyle. However, the description of what "an active lifestyle is" was not recorded.<br><br>A single blinded occupational therapist assessed outcomes limiting bias (Kendall 2003; Moustgaard *et al.* 2014). Thus, inter-assessor variability was controlled (Kendall 2003). Patients were unblinded, inferring bias from this source (Moustgaard *et al.* 2014). Assessments occurred at a set time, controlling for time-modified confounding (Platt, Schisterman and Cole 2009) and diurnal RA effects (Nandgude, Hasabe and Kolsure 2018). The outcome measures utilised a balance of subjective and objective tools (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), allowing independent verification. Outcomes included are outlined above, but only the grip strength (Lee *et al.* 1974); ROM of MCP and PIP (Ritchie *et al.* 1968); timed 9-hole peg-test (Mathiowetz *et al.* 1985), were validated for RA patients. No discussion was presented on outcome tools in terms of reliability and validity and the specificity of the tool to the RA population; this hinders the interpretation of outcomes and generalisability in clinical practice (Kendall 2003). Further, it was not stated and thus assumed that the assessor re-assessed patients in the same manner, controlling for assessment variability and testing protocol deviations (Kendall 2003), particularly as relates to the environment in which assessments were undertaken, and the effect on the outcomes (Richardson 2007). The last factor not reported related to the relationship between the last intervention and outcomes assessment, which impacts interpretation of data (Pater *et al.* 1998)). Patient's feedback on exercise programme difficulties experienced was by telephonic questionnaire. This data was at odds with reviews of the conversations, patient records, and contact with the treating physician; so, this data provided little in terms of outcomes.<br><br>At baseline, there were no significant differences in intra-group baseline variance on outcome measures, except for metacarpophalangeal extension range bilaterally and ulnar deviation in the left arm. This along with the failure to use an *a priori* analysis (Kendall 2003; Christley 2010), results in an underpowered study (Mouton 2006; Christley 2010). Additionally, the study failed to make drop-out prevision (Fogg and Gross 2000), given that the study suffered a 23% drop-out (Fogg and Gross 2000). Thus, the data presented is only reflective of and bias to completers (Heritier, Gebski and Keech 2003; Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017), skewing data by excluding those participants getting worse (Fogg and Gross 2000; Kendall 2003; de Boer *et al.* 2015). All these factors in addition to use of only non-parametric statistics (Hoskin 2012; Bahreini *et al.* 2020) impacts on only trend reporting and required post hoc analysis to control for unequally distributed variables, which introduces risk of data normalcy (Fogg and Gross 2000; Kendall 2003; de Boer *et al.* 2015). Finally, the lack of published MCID values comparisons does not provide for interpretation of clinical relevance of outcomes (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020). |
| Outcome | The study concluded that the effects of a home-based hand-exercise programmes performed twice a day for 10 to 20 minutes per session for three months in the treatment of stable/remission RA patients (functional class II and III with deformity of the hand/s), was effective in improving grip strength, joint count in the right-hand, while the administration of ROM plus balanced resistive exercises showed improved left-hand dexterity. Collectively, exercise programmes reported significant ($p$<0.05*)* improvements in left-hand grip strength. Side effects were noted as transient, moderate discomfort and exercises were tolerated well (Hoenig *et al.* 1993). |
| Discussion | The outcomes were favourable, even with outcomes skewing secondary to sample size, controlled statistical analysis and thus the possibility that change being due to chance cannot be excluded. In addition, the confounding factors related to lack of parity of baseline demographics and outcome measures in addition to the lack of control of various influencers affects the study's ability to accurately and fairly compare the groups outcomes and conclusions. This study in particular may be subject to reporting of only less severe cases of RA, based on the exclusion of more severe RA cases. Thus, the ability of the study to provide clinically relevant outcomes is limited in the form of provided statistics as well as from MCID comparison even though the study suggests that the intervention is safety and has no adverse reactions (not necessarily true if patients were excluded due to worsening symptoms). |
| Conclusion: | Given the severity of the external validity bias in this study, there is very limited methodological rigour. This contrasts the internal validity review which reflected 6/11 (85% agreement), indicating that methodological rigour was moderate. Collectively, these outcomes suggest poor to moderate methodological rigour and thus that this study does not provide more than very limited evidence in favour of the interventions. Thus, for this study to have both published accuracy and clinical relevance, it needs to be repeated with a larger sample (determined by power-analysis) and addressing all concerns noted in this review. |

Table 4.22a: Tabulated feedback of reviewer data

| Author/s | Levitsky, A., Kisten, Y., Lind, S., Nordström, P., Hultholm, H., Lyander, J., Hammelin, V., Gentline, C., Giannakou, I. and Faustini, F. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2019 | | | | | |
| Title | Joint Mobilization of the Hands of Patients With Rheumatoid Arthritis: Results From an Assessor-Blinded, Randomized Crossover Trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | Yes | Yes | Yes | 67% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 9/11 | 9/11 | 9/11 | |
| Overall percentage agreement | | | | | | 97% |

Table 4.22b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Levitsky, A., Kisten, Y., Lind, S., Nordström, P., Hultholm, H., Lyander, J., Hammelin, V., Gentline, C., Giannakou, I. and Faustini, F. |
| Year | 2019 |
| Title | Joint Mobilization of the Hands of Patients With Rheumatoid Arthritis: Results From an Assessor-Blinded, Randomized Crossover Trial. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Pain: VAS and PROMS<br>- Joint count (pain/swelling)<br>- Q-Doppler (MCP and region),<br>- Synovial fluid (MCP and region),<br>- Joint space (MCP and region)<br>- **Exploratory:** M-CDAI, M-DAS28, CRP/ESR, HAQ, SF-36 PCS, SF-36 MCS, VAS (fatigue, global health – participant and physician), Wrist (Q-Doppler and synovial fluid) | Pre-/Post study with crossover design (baseline, week 2, 3, 4) Follow-up one-month after last treatment (at week four – end of second phase of therapy). | Two-months | 20 patients: 12 RA participants, eight OA participants, no fall out in the study. (40 hands, 320 joint) | Yes | Yes, crossover design (one hand intervention, the other a control -patient was their own control) | Yes, computer generated allocation with stratification. | 9/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 97% |

| | |
|---|---|
| Limitations | The study participants (RA and OA) were recruited from the University of Karolinska (Sweden), Rheumatology clinic or Swedish Rheumatism Association. This creates the perception that the sample is likely to include a variety of RA presentations. This may make the sample more similar to a general RA population, increasing generalisability but it decreases the chances of being able to actually test the intervention (Wells *et al.* 2012). The generalisability may be limited as treatment was monocentric (i.e., not all willing RA patients could have been accommodated). The inclusion criteria required a patient of >18years, with a diagnosis of RA or OA (of the hand) (unknown criteria) and hand pain for at least six-months. A clear set of exclusion criteria removed any existing bone damage or soft tissue injury to the hand, acute disease activity (inflammation) in the joints of the hand (week preceding the start of the study), surgical intervention to the upper limb or known pregnancy within three months of study onset.<br><br>Twenty-six eligible participants were found, of which 20 were included (12 RA and eight OA-hand). Each participant (by means of a crossover design) was their own control (one hand served was treated the other as the control (vice versa)). Hand assignment was determined by computer-generated randomisation with stratification according to diagnosis. This method provides concealed allocation, and thus protection from researcher bias and interference in participant allocation (Kendall 2003), while stratification aided in protecting group distribution and against external influencers (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019). By contrast the inclusion of "seropositive polyarticular juvenile idiopathic arthritis" into the study, is perceived as a breach of the inclusion criteria and may introduce type II error into the analysis. Additionally, the study does fail to report medications variables (e.g., pharmacokinetics/dynamics and prescription variables), and the effect whether the 24-hours prior to treatment was sufficient to prevent confounding.<br><br>The reported demographic data reported (sex, hand dominance, smoker status, serology (both RF and ACPA), age, disease duration) was clear and comprehensive, offering the reader a clear patient description. Although, age ($p$=0.002) and VAS (self-reported global health) ($p$=0.047) were significantly different comparing RA to OA participants. This however would not have affected the RA intervention versus control hands and acknowledges the differences between the conditions. What does however affect the RA group is the lack of reporting on comorbidity data (only available as supplementary material (not in article)).<br><br>Pragmatic control of confounders was implemented. These included, instructing patients to continue lifestyle as "normal", which keeping medication stable during the study (unless physician instructed change). Additionally, patients were asked to not consume medication the day of or the day before mobilization (i.e., steroids or pain killers), to allow effective mobilisation to occur within pain tolerance. Finally, patients were asked to not perform mobilization on themselves or seek out other forms of care for the hands (control for this is an acknowledged limitation). As a downside, medication adherence, changes recorded, and study protocol adherence were not reported (e.g., exercise diary).<br><br>The study failed to report power-analysis, implying risk of the study being underpowered (Mouton 2006; Christley 2010). This was not adversely affected as the study did not suffer dropouts which negates risk of missing information and renders the intention-to-treat analysis unnecessary (Fogg and Gross 2000; de Boer *et al.* 2015). However, the study would be limited to non-parametric statistics, in this providing trend reporting at best (Hoskin 2012; Bahreini *et al.* 2020). The authors acknowledge the risk of bias in this study in the form of regression to the mean and the implemented post-hoc analysis by means of Dunn-Bonferroni correction, risking type II error conclusions (de Boer *et al.* 2015; Unbiased Research 2017). |

| | |
|---|---|
| | The intervention programme for hand metacarpophalangeal (MCP) joints, by means of Kaltenborn – grade I to II manual mobilization (Kaltenborn 1993; Kaltenborn FM 2011), administered low-velocity low-amplitude mobilization of non-destructed hand joints. Implementation was by five blinded (outcomes and diagnosis) manual therapists (four naprapaths and one physiotherapist) on a weekly basis. This study therefore provided a clear and reproducible treatment protocol and patient (Kendall 2003). This was necessary because each participant had two or more therapists per hand during the study. This introduced risk of inter-therapist variability, even with control measures implemented (Kendall 2003). One hand was treated and the other served as a no treatment control. Each arm of the crossover design was two-weeks in duration, with treatment occurring immediately after data collection (baseline (week zero and three)) and before post-intervention assessment (week two and four).<br><br>Outcomes were assessed at immediately before first treatment (baseline and week 3 – after crossover), and immediately after intervention last treatment (week two and four). The post intervention assessment may have been too early for measuring programme effect as opposed to immediate treatment effect of last treatment (Pater et al 1998). Thereafter a follow-up occurs at two-months (one-month after intervention ceased) allowing for intermediate term effect measures (Pater *et al.* 1998). The outcomes were assessed by six blinded (diagnosis, and treatment (hand)) physicians (each participant was assigned two physicians, raising the concern of inter-assessor variability (Kendall 2003)); and ultra-sonographer (ultrasound-based outcomes). There was a good balance between objective and subjective outcome tools, providing control of participant and assessor bias (e.g., placebo and Hawthorne) (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). The study seems to suggest differences in baseline between the RA and OA groups (e.g., VAS (participant self-reported global health) ($p$=0.047)), however the focus of the study should be between the hands in the RA and OA groups (de Boer *et al.* 2015). Additionally, some baseline values are 0, which implies a floor effect (Liang 2000; Coster 2013). The outcomes seemed appropriate, although reliability and validity were not noted for all outcomes measures, particularly in the context of RA (Coster 2013; Rezaei et al. 2014; Rezaei et al. 2017; Bahreini et al. 2020).<br><br>Authors concerns arising from the study included crossover risks such as learning effect and carry-over effect (Bordens and Abbott 2002) particularly if the washout period is not sufficient. The synergistic/antagonistic effect (Robins, Hernán and Siebert 2004) relating to neurological or biomechanical influence of ipsilateral therapy conferring effect contralaterally (Dragert and Zehr 2013; Villafañe, Cleland and Fernandez-de-Las-Peñas 2013). Lastly, the study does not report MCID values, to which clinical changes can be compared, limiting understanding of clinical improvement and value (Bahreini *et al.* 2020). |
| Outcome | By means of a crossover design the study illustrated the effect of mobilisation (manual therapy) in the treatment of participants with RA or OA experiencing daily hand pain. The participants received two-weeks of directed treatment by means of Kaltenborn mobilisation. The participants were self-controlled (one hand per group) as well as the OA participants representing a clinical control for the RA participants. The RA participants reported significant bilateral hand improvement (for pain, active inflammation, synovial fluid effusion (all $p$<0.05) and joint space of the MCP ($p$≤0.01)) both from baseline to cross over and then on to follow-up at two-months. Likewise, OA participants showed significant improvement in hand pain ($p$<0.05) and joint space. Highly significant improvements were reported by all participants for fatigue scores post-treatment ($p$≤0.001), this was sustained ($p$<0.05). Other outcomes were insignificant (Levitsky *et al.* 2019). |
| Discussion | From the description above, the study lacked control over several limitations that risk confounding the external validity of the study outcomes making the findings less generalizable to patients in clinical practice. This is indicated as there is a lack of reporting of all RA dynamic characteristics, resulting in the inability of the practitioner to know which patient type is likely to respond. Secondly, the consistent comparison of the RA to OA does not provide an effective means of comparison, as these are two different conditions with different impacts from the mobilisation employed. However, there is similar weakness in the comparison of the right to the left hands as there is published literature suggesting cross over effects of treating one side and resulting in opposite limb improvement (Dragert and Zehr 2013; Villafañe, Cleland and Fernandez-de-Las-Peñas 2013). Therefore, the actual improvements versus the comparisons in this study are flawed and the outcomes cannot be extrapolated to patients in practice. A better practice would have been to have two RA groups with treatment in one group and control in other group whilst measuring both hand outcomes, in order to detect change in treated hand as well as contralateral hand versus a control. Further complications include confounders which do not allow for clinical changes to be solely as a result of the Kaltenborn mobilisation, the sample size (under powering) increasing risk of type II error, use of controlling statistics (post hoc analysis and Bonferroni correction), the associated limitations of a crossover design, outcomes measures timing and inter-assessor and -therapist variability. |
| Conclusion: | Given the external validity assessment above, it would seem that there is a moderate bias, excluding the fact that the comparisons between OA and RA, right and left hands are flawed. Thus, although the study has a 97% agreement between reviewers for very high methodological rigour and a low risk of bias, the study does not actually enable a conclusion that allows for understanding of the clinical effectiveness of the Kaltenborn mobilisation for RA of the hand (when the study is taken in isolation). If there are other studies that indicate that Kaltenborn mobilisation in OA is better than placebo; then this study may have a basis for interpreting the outcomes of the RA to the OA in this particular context, however no mention is made of this in the discussion and conclusions. Thus, there is great tension between the internal validity and external validity in this study. Had both groups in the study been RA patients with one hand treated the outcome would have been an overwhelming positive for internal and external validity, but this is not the case. Therefore, at the very best this study provides limited to no evidence in favour of the Kaltenborn mobilisation in the treatment of hand RA. |

Table 4.23a: Tabulated feedback of reviewer data

| Author/s | Lineker, S. C., Bell, M. J., Wilkins, A. L. and Badley, E. M. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2001 | | | | | |
| Title | Improvements following short term home based physical therapy are maintained at one year in people with moderate to severe rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | No | No | No | No | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | Yes | No | No | 67% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | No | No | No | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 5/11 | 7/11 | 6/11 | 6/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.23b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Lineker, S. C., Bell, M. J., Wilkins, A. L. and Badley, E. M. |
| Year | 2001 |
| Title | Improvements following short term home based physical therapy are maintained at one year in people with moderate to severe rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Stanford Arthritis Self-Efficacy Scale (SES)<br>- Arthritis Community Research and Evaluation Unit Rheumatoid Arthritis Knowledge Questionnaire (KQ)<br>- VAS (pain)<br>- AIMS2<br>- Sickness Impact Profile (SIP)<br>- Disease activity measures (tender joints, grip strength, and morning stiffness) | Baseline and at week -12 and -52. | 12-weeks of treatment with a one-year follow-up.<br>52-weeks total duration of study. | Of the 150 randomised participants, 127 completed the study protocol, and of those 117 were available for follow-up at one year. | Yes | Yes, in initial study (waiting list control) that starts intervention (week seven to -12). Thereafter, groups are both intervention and compared to dropouts and baseline. | Yes, for initial study. | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study was based at the Arthritis Community Research and Evaluation Unit, Princess Margaret Hospital, Toronto, Ontario in Canada. Which suggests that the patients were recruited from this community. The article does however neglect to clearly define the source and method of recruitment within this context. As a result, it is not possible to evaluate the effects of these variables on patient cohort similarity or differences from the normal RA patient cohort and it is difficult to affirm whether patient novelty and nativity did influence the outcomes of the study. It is possible that the patients were drawn from within the hospital and thus would represent more severe cases with less likelihood of treatment naivety or novelty (Wells *et al.* 2012).<br><br>In regards the methodology, the study refers to Bell (1998) for a detailed methodology. Thus, the methods here are less detailed and a source of bias compromising internal and external validity. Ideally, as per systematic review methodology studies are evaluated as their own standalone item of research (Moher *et al.* 2012).<br><br>Based on an *a priori* analysis a $p \leq 0.001$ was considered significant with the sample attained. The sample recruited included 150 participants that were adults diagnosed with moderate to severe RA (unknown criteria). This limited inclusion criteria prevents a clear participant description and opens the study to bias, based on confounding factors that cannot be evaluated (Kendall 2003; Wells *et al.* 2012). However, the study did report no significant differences between baseline characteristics of the groups (viz., age, disease duration, gender, marriage status, education, economic status, functional class, broad medication usage (although, there is a lack of detail regarding medication, which can impact the outcomes of the study). These 150 participants were randomly allocated to one of two groups (intervention and control). Initially intervention is a standardised six-week physiotherapist delivered home based intervention programme [education (about RA, its management and individual goal setting), pain modalities and individualised exercise] compared to a wait listed control. Thereafter both groups received intervention, thus the control group thus became a second intervention arm after initial treatment period. There is discord between the standardised intervention and the individualised exercise programme as they are contradictory by nature. This creates concern of bias around treatment consistency. The secondary concern is one of comparability of the two groups at the longer-term follow-ups as exposure to case was different between the groups.<br><br>Of the 150 participants, 127 completed the 12-week intervention period (85%), and 117 (92% of completers, 78% starting sample) completed the one-year follow-up. For the outcome measures the study reported that the baseline characteristics showed no statistically significant differences. The study reports that the loss-to-follow-up did not differ significantly from the sample protecting against type II error in the analysis and conclusions.<br><br>The programme utilised in this study is vague in terms of its components and for these reasons (without referring back to the prior study) prohibits the repetition of the study or clinical replication in practice. The use of an adherence diary was acknowledged as a limitation as the regularity of prescribed aerobic level exercise could not confirmed, within the six-weeks for which the study was carried out and after participants continued (possibly to varying degrees. This impact on being able to compare outcomes between participants and groups that have variable exposure to exercise and its subsequent effects. |

| | |
|---|---|
| | To the study's credit, the study used both subjective and objective outcomes. A majority of these were subjective – introducing risk of bias as a result of placebo, Hawthorne, memory recall/decay, and learning effect in patients (Bordens and Abbott 2002; Mouton 2006; Ricker, Vergauwe and Cowan 2016). Of the outcomes not all were validated in terms of their use in RA patients as an outcome measure (Coster 2013; Bahreini *et al.* 2020). The measures, at baseline, week-12, and week-52; were recorded by independent blinded assessors that managed the assessment time intervals as well as the data collection protocols. Collectively these limitations rise concerns in the reader of bias, type II error, and insensitivity of the outcome evaluation. Implying risk of unreliability of the study conclusions.<br><br>The type II error is enhanced by the use of MANOVA (nonparametric statistic) whereas the paired t test (parametric) protected against this along with the $p \leq 0.001$, so as to limit the impact of multiple variables. The effects of omitted data are not reported. And the authors acknowledge that limitation/s and risk to the study is introduced by means of lack of natural history control group from week-12 to -52, that the co-interventions (in particular medication) may have influenced the reported outcomes, and that natural history may have played a role in results noted in the study. |
| Outcome | The study reported on:<br>1. a six-week (short-term) home-based intervention programme emphasizing strategies for self-management.<br>2. a follow-up period of one-year of a sample of 150 participants with moderate-to-severe RA, randomised into an intervention group or waiting list control (which received the intervention from week six to 12).<br><br>At 12-weeks 127 completers with both groups showing similar significant improvement in SES (self-efficacy and management), KQ (knowledge), and morning stiffness. At 52-weeks 117 participants were available at the follow-up assessment, reporting improvements to have been maintained ($p<0.01$). Long-term measures (AIMS2 and SIP) improved significantly from baseline to 12 weeks ($p \leq 0.001$) and at 52-weeks they were maintained or improved further. The study reported pain, joint count (tender), and grip strength were also reported improved at 12-weeks ($p<0.001$). Thus, the study showed participants participating in the six-week home-based programme improved specified outcomes and were maintained at one-year (Lineker *et al.* 2001). |
| Discussion | Given that the study seemed to lack control over several patient related confounding variables in terms of recruitment, inclusion, baseline variables prior patient exposure to the treatment and whether there was treatment novelty; there is a high risk of bias, implying that external validity is compromised and thus results may not be generalised into clinical practice. This may be secondary to a lack of reporting and reference to the prior study, but also relates to the intervention used implying that the outcomes are not solely due to the intervention.<br><br>Additionally, the lack of a long-term wait listed, or control group confounds the understanding of the 52-week outcomes as it is impossible to determine whether the outcomes are due to natural history, chance, RA pathogenesis, a combination of these or just changed habits in the patients (no adherence diary). Therefore, the study is not able to state categorically that the outcomes although favourable were as a result of the intervention. |
| Conclusion: | In the context of the provided impact on the external validity of the study, the outcomes of this study are considered to be at high risk of bias – i.e., the likelihood that the outcomes are directly as a result of the treatment are small due to uncontrolled for confounding variables. By contrast the structure of the study are provided by a 94% agreement between reviewers, showed that there was a moderate risk of bias in terms of the internal validity (scored the 6/11 PEDro ranking).<br><br>Given that the study has a moderate to high level of bias, the study is not able to provide much evidence in favour of the intervention. Thus, the study should be retested with two RA groups one of which is a true control, and addressing limitations highlighted, in order to provide a basis from which to determine any true causal link and longevity of the effect imparted, as well as application in the clinical setting can be made with a prognostic capacity. |

Table 4.24a: Tabulated feedback of reviewer data

| Author/s | Lyngberg, K. K., Harreby, M., Bentzen, H., Frost, B. and Danneskiold-Samsøe, B. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1994 | | | | | |
| Title | Elderly rheumatoid arthritis patients on steroid treatment tolerate physical training without an increase in disease activity | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | No | No | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | Yes | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 8/11 | 7/11 | 7/11 | |
| Overall percentage agreement | | | | | | 85% |

Table 4.24b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Lyngberg, K. K., Harreby, M., Bentzen, H., Frost, B. and Danneskiold-Samsøe, B. |
| Year | 1994 |
| Title | Elderly rheumatoid arthritis patients on steroid treatment tolerate physical training without an increase in disease activity |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| Clinical outcomes: Functional impairment (Steinbrocker classification), joint count (swelling, tenderness, and pain in movement),<br>Laboratory outcomes: ESR, Haemoglobin, Potassium, Sodium, Creatine, Leucocytes, Ferritin, Iron, Transferrin, Thrombocytes<br>Functional outcomes: functional ability: 30-meter walk time, grip strength, stair climbing. Muscle strength (dynamometer), peak torque, angular velocity. Aerobic capacity (Astrand test)<br>Subjective assessment of morning stiffness duration and functional assessment (Fries' index). | Randomly during two-weeks preceding and following the intervention. | Four-months (three-months of intervention with a two-week period for assessment preceding and following the intervention). | 24 participants in total at the start of the study and randomised into two-groups of 12. Two from the training group dropout. Thus, completers include 10 training group (83%) and 12 controls (100%) | Yes | Yes, a non-training group, allowed to continue other previously undertaken physical activities. | Yes, no methodology was provided. | 7/11, indicating a moderately high methodological rigour and subject to low risk of bias within the internal validity of the study. | 85% |

| | |
|---|---|
| Limitations | The article fails to clearly state the source of sample recruitment but offers some guidance as to the location of the study; researchers are cited as staff of two hospitals (Herley Hospital and Fredericksburg Hospital) in Denmark. Within the context of the lack of clarity, it is not possible to identify with the participant population and thus it is impossible to translate the results to a general RA population that may present to practice (Wells *et al.* 2012). The initial 50 RA patients that where eligible, were required to meet the following inclusion criteria, a RA diagnosis with RA (ARA 1987 criteria (Arnett *et al.* 1988)), be on continuous low-dose oral steroids, have slight or moderate disease activity (criteria unknown, but loosely based on morning stiffness, ESR, joint count of painful and swollen joints). Exclusion occurred, with the presence of comorbidities and contraindications to care (e.g., heart disease). The sample was reduced to 24 participants randomised into two-groups of 12-participants (training group and control) Groups were statistically insignificantly different at baseline (age, functional class, and disease duration). Thus, in terms of patient presentation, the groups were homogenous at baseline (de Boer *et al.* 2015). However, the inclusion criteria that could have been considered and may have influenced the outcome are RA serology, global pain severity (rather than just joint pain) and joint status (deterioration/deformity – imagining) (Kendall 2003; England *et al.* 2019; Myasoedova *et al.* 2020). Medication relatively was well recorded and controlled for, as all participants joining the study had to be on DMARD therapy (unadjusted for four-months prior to the study), the use of steroids was not permitted within six-months of the study start. The problem with the DMARD therapy is that it was unequal between the groups (six in the training group and two in the control). Furthermore, the use of NSAIDs was permitted and fluctuated along with the use of supplemental over-the-counter Aspirin. Thus, even though groups were not equal, changes in medication were recorded and noted in the analysis.<br><br>The method of randomisation was not detailed, which influences interpreting the quality of randomisation, even though we know that the baseline participant characteristics were the same, the same cannot be stated in terms of the outcome measures. However, by contrast, concealed allocation was implemented as it was noted that participants allocated to the control were blinded to intervention group and that the outcome assessors "did not participate in the allocation of the participants, nor did they observe the participants during training". These points assist in reducing assessor and patient bias (Mouton 2006; de Boer *et al.* 2015).The sample size is very small, combined with the failure to report power analysis calculations (Kendall 2003; Christley 2010) and the implementation of intention-to-treat (despite the pre/post design prohibiting this) (de Boer *et al.* 2015), implies that the study was underpowered (Mouton 2006; Christley 2010). The study is also limited to employing non-parametric statistics, in doing so limiting the outcome reporting to only trends of change rather than causal links (Hoskin 2012; Bahreini *et al.* 2020). The study suffered dropout of two participants from the training group for non-intervention for RA related reasons. Although the above contribute negatively to being able to interpret the statistical outcomes, the fact that all participants lacked prior exposure to aerobic or strenuous exercise training, protected the study from influence of participant prior perception bias (Mouton 2006; Moustgaard *et al.* 2014). |

| | |
|---|---|
| | The control was a non-training group, but all participants continued with previously undertaken physical modalities (including heat application and warm-water exercise). These latter additions are a cause for concern as the confounding effects (Mouton 2006). The exercise programme consisted of progressive interval training design: aerobic training, strengthening exercises, and stretching. A brief description is offered, and reference is given to a prior study (Lyngberg *et al.* 1994). The supervised intervention was delivered in groups of four over a three-month period, twice a week for 45 minutes per session. This provided the necessary expose for exercise changes to manifest and be measured (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). Tailoring of the programme could not be determined as protocol was unclear. An important oversight in study replication (Heine *et al.* 2012). The article seemed to suggest inclusion of multiple therapists were involved in providing the intervention increasing the risk of inter-therapist variability (Kendall 2003). The therapists were not blinded, and this raises the risk of bias (Moustgaard *et al.* 2014).<br><br>It seems that participants were allowed to continue other physical activities they usually perform (heat and water exercise). This although confounding the outcomes, may present a means to balance the control and intervention groups in terms of the "additional" interventions. But it does have the downside in that synergistic / antagonistic effects of different therapies cannot be excluded as a risk (Robins, Hernán and Siebert 2004). Adherence to the programme was reportedly on "average" at 93%, which seems high compared to other such studies, but it does pose a greater risk for an underpowered study by impacting physiological and psychological outcomes (Kaptchuk 1998; Bordens and Abbott 2002; Bialosky *et al.* 2011).<br><br>Balanced outcome measures (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), included clinical, laboratory, and functional variables supplemented by subjective participant assessment. The objective outcomes included: 1. clinical outcomes of functional impairment (functional ability: 30-meter walk time, grip strength, stair climbing) (Steinbrocker classification (Steinbrocker, Traeger and Batterman 1949)), joint count (swelling, tenderness, and pain in movement) (Lyngberg, Danneskiold-Samsøe and Halskov 1988). 2. laboratory outcomes of ESR, haemoglobin, potassium, sodium, creatine, leucocytes, ferritin, iron, transferrin, thrombocytes. 3. additional outcomes of muscle strength (dynamometer), peak torque and angular velocity (knee extensors bilaterally and the ankle plantar flexors bilaterally) (Danneskiold-Samsøe and Grimby 1986), aerobic capacity (Astrand 1960) with bicycle ergometry. Whilst subjective outcomes included morning stiffness duration and functional assessment (Fries 1983). All assessment were conducted by a blinded assessor protecting against intra-assessor variability (Kendall 2003; Moustgaard *et al.* 2014). Given the limitations of this population at assessment stage it was realised that the $VO_2$ estimates were too high for this group (Astrand PO 1986), thus the aerobic capacity was adjusted by nomogram (after which 11 filled the criteria and eight had pain induced broad heart rate variation). This implies that the outcome was not only subject to the ceiling/floor effects (Liang 2000; Coster 2013), but also inappropriate for this patient group and required post hoc analysis and correction (Fogg and Gross 2000; Kendall 2003), reducing the contribution to the study. All evaluations were conducted in random, and the study did not report control over testing environment (Richardson 2007) and time-sensitive confounding (Platt, Schisterman and Cole 2009) [such as the diurnal effect (Nandgude, Hasabe and Kolsure 2018)]. However, the authors acknowledged the possibility of seasonal effect on outcomes (study completed in autumn and winter). Therefore, although the outcomes measure in most instances were valid and reliable, there was a significant limitation associated with the accuracy of some of the outcomes, even though disease activity measures (joint count (swollen, tender) morning stiffness duration and ESR), aerobic capacity and muscle torque outcomes were reported as insignificantly different between the groups. These problems increase concerns of type II error (de Boer *et al.* 2015). Finally, the study failed to report comparison to published MCID values to allow understanding of clinical significance (Bahreini *et al.* 2020). |
| Outcome | The study reports the effect of progressive interval training, on elderly patients with RA with functional class I-III (by Steinbrocker's) treated by long-term steroids. The results showed that the work capacity was doubled ($p$=0.006), with an increase of 76% in repetition capacity. The study provides evidence that muscle torque bilaterally improved in the plantar flexion of the ankle (right ($p$=0.06) and left ($p$=0.04)) and significant improvement in bilateral step climbing and number of heel lift capability. These improvements were found while evidence indicated that disease activity remained insignificantly changed. Thus, the study concludes that an individually tailored exercise programmes could be recommended for elderly RA patients on steroid treatment (Lyngberg *et al.* 1994). |
| Discussion | Although the outcomes of this study were favourable, in the face of an underpowered study, multiple confounding variables, impacts on outcome measures that reduced accuracy of reporting and statistical adjustment for many of these "faults", suggests that there are multiple concerns with regards to accepting the outcomes at face value, as it seems that the outcomes may not be a direct effect of the stated intervention. For the limitations of this study to be addressed, it is suggested that the study be repeated with a larger sample (determined by power-analysis) where a participant description including variables of RA such as serology (among others) are considered, and appropriateness of outcome measures validated to provide a base from which to determine the true causal link and application in the clinical setting with prognostic capacity. |
| Conclusion: | Given the outcomes of the external validity assessment of the article, it is evident that although the authors tried to limit negative impact, there are some significant factors that increase the likelihood of bias within the study. Therefore, at best the study is at moderate risk of bias from an external validity vantage point. This does not differ much from the internal validity review completed by the reviewers where it is seen that a 7/11 rating was achieved, which indicates that there was moderate to high rigour and thus medium to low risk of bias. When seen in combination the overall bias assessment if low to moderate and therefore this study provides limited evidence in support of the intervention for the RA patients presented in this study. |

Table 4.25a: Tabulated feedback of reviewer data

| Author/s | Mannerkorpi K. and Bjelle A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1994 | | | | | |
| Title | Evaluation of a home training programme to improve shoulder function in rheumatoid arthritis patients. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | No | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | No | No | No | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 6/11 | 5/11 | 5/11 | 5/11 | |
| Overall percentage agreement | | | | | | 97% |

Table 4.25b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Mannerkorpi K. and Bjelle A. |
| Year | 1994 |
| Title | Evaluation of a home training programme to improve shoulder function in rheumatoid arthritis patients. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - ROM (flexion/extension)<br>- Functional movement (hand-to-neck and hand-to-scapula)<br>- Pain VAS (rest, night before, active movement)<br>- Endurance (Max time, Borg's 15-point scale)<br>- ADL<br>- RF and CRP | Pre/post | Eight-weeks per intervention period. The control run the intervention programme after the experimental group therefore 12-weeks. | Approached = 35 (Exp. group = 19, Control = 16)<br><br>Dropout = seven (Exp. group = five, Control = two)<br><br>Completers = 28 (14 per group round 1) (Exp. group = 20, Control = 14) | No | Yes | Yes | 5/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 97% |

| | |
|---|---|
| Limitations | By means of consecutive sampling participants were drawn from the Department of Rheumatology Outpatient's Clinic at the Sahlgren University Hospital in Goteborg, Sweden. As for previously reviewed studies, specialist centres as a source for RA patients, suggests decreased likelihood of naivety on the part of the participant to the intervention and it also suggests that the patients may have more chronic or severe RA impacting outcomes and clinical applicability respectively (Wells *et al.* 2012; Jager, Putnick and Bornstein 2017). The authors approached >35 consecutive eligible participants for inclusion based on them being female, with seropositive RA (RF) diagnosed [revised ARA criteria 1987 (Arnett *et al.* 1988)] and fulling into class I or II in the ARA functional class (Steinbrocker, Traeger and Batterman 1949) and having shoulder pain. Potential participants were excluded if concomitant conditions were present, medication was modified or they had steroid injection six months preceding the study, had physical therapy for their shoulder pain / dysfunction. As a result, 35 patients were recruited with the following characteristics: 20–70 years of age, 1–10-year disease duration and randomised by sequential allocation by age, functional class, disease duration, and DMARD intake (Pocock and Simon 1975). This allowed high comparability of the groups (Fogg and Gross 2000) and provided a degree of concealment on allocation (Mouton 2006). The study reports that at baseline the groups did not prove significantly different in age, disease duration, functional class, DMARD use and neck / elbow / hand pain or impairment.<br><br>The power analysis was based on Boström and Ljungquist (1984), which determined the need for 14 participants per group at a power of 80% for a significance level of 0.05 (two-tailed test) (Kendall 2003; Christley 2010) to detect statistical significance. Post randomisation this was met with 19 patients in the experimental group (home training programme) and 16 in the control, whilst also accommodating possibility for drop out (Fogg and Gross 2000). This was appropriate as seven participants (five experimental and two control) dropped out, resulting in 14 per group (completer rate of 74% and 88% respectively). The impact of the drop out on group homogeneity is not provided and does predispose the study to risks in relation to type II error (Bartley and Fillingim 2013; de Boer *et al.* 2015). It is claimed that this may have been controlled for given the dropouts did not differ significantly from the total sample in terms of shoulder status and disease activity (CRP levels), however these are not the only RA dynamic characteristics that can influence outcomes<br><br>It must be noted that the sample size, allowed non-parametric analysis (Fisher's non-parametric permutation test and liner non-parametric permutation test), which limited the authors to reporting trends only (Hoskin 2012; Bahreini *et al.* 2020). This is further compounded by the lack of consideration of confounding variables (Kendall 2003), including but not limited to comorbidities of RA (Kłodziński and Wisłowska 2018; England *et al.* 2019; Levitsky *et al.* 2019; Myasoedova *et al.* 2020), contraindications to the intervention (Durcan, Wilson and Cunnane 2014), patient endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018), and other general factors (Bartley and Fillingim 2013). Given all the above, the study must be credited for ensuring lack of difference between the groups for pain, active ROM, functional movements, endurance, and ADLs at baseline. Also given the lack of numbers, the authors allowed six control to complete the intervention programme (after completion of the control period) and added their data to the intervention group in order to enhance numbers). This although good for the group numbers had an impact on patient bias, through patient perception and satisfaction. |

| | |
|---|---|
| | The intervention included one-on-one instruction over a 30-minute session for the home-based exercise programme. This exercise programme was well described with clear exercises variables described (sets, reps and resistance (although latter is unclear)) (Heine *et al.* 2012). Modifications were possible with therapist one-week after the study commenced. The intervention was carried out over eight-weeks, which provided the necessary expose for measurable physiological change (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). The use of a single therapist provided control over the risk of inter-therapist variability (Kendall 2003) despite the participant being responsible for their own therapy administration. In addition to the exercises, patients were also offered information (poorly described) about RA, its impact and appropriate ergonomics; in order to facilitate pain relief and enhance endurance. This limits outcomes to the programme and not individual interventions, where synergistic/antagonistic effects cannot be accounted for (Robins, Hernán and Siebert 2004). Adherence to the programme was not reported and prevents impact on outcomes to be considered.<br><br>Positively, the study employed subjective and objective tools as outcome measures (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), which were recorded an assessor who was also the therapist (no blinding from intervention), which introduced assessment bias, but reduced the risk of assessment variability (Coster 2013; Bahreini *et al.* 2020). Although the authors suggest that the assessor was blinded from previous outcome readings, this does not negate the influence of measures taken (Kendall 2003; Moustgaard *et al.* 2014).<br><br>These outcomes included pain (Price *et al.* 1983), ROM of the shoulder (flexion and abduction) (American Academy of Orthopaedic Surgeons 1965), functional movements (Constant and Murley 1987). However, some authors suggest that the use of combined movements introduces a questionable sensitivity and specificity to the outcome measures (Coster 2013; Bahreini *et al.* 2020). Similarly, endurance (Borg 1982) measured on the Borg scale raises questions of reliability and validity of the tool in this study, especially with the creation of an artificial floor effect (Liang 2000; Coster 2013). ADL outcome measures were a mix of validated and unvalidated tools introducing outcomes of variable strength (as was acknowledged in the publication) (Coster 2013; Bahreini *et al.* 2020). The clarity of some outcomes provided strength to the study while others aspects of the assessment introduced risk time-modified confounding (Pater *et al.* 1998) such as the diurnal effect (Nandgude, Hasabe and Kolsure 2018), testing environment (Richardson 2007), relationship between last intervention and assessment not being clear (Pater *et al.* 1998). These limit analysis of bias control and thus attained results (Coster 2013; Bahreini *et al.* 2020). Patient perceptions will influence outcome reporting of subjective tools was a result of lack of blinding (Mouton 2006; Moustgaard *et al.* 2014). The reporting of MCIDs may have allowed some understanding of clinical applicability, but this was not provided (Bahreini *et al.* 2020). |
| Outcome | It was found in RA patients with shoulder dysfunction, that post eight-weeks of home-based exercises focused on shoulder function, there was improved endurance of shoulder musculature (bilaterally) ($p$<0.05), reduced pain in active ROM (bilaterally) ($p$<0.05) and improved functional ability in terms of abduction in the left arm ($p$<0.01), and hand-to-neck mobility in the right ($p$<0.05) (Mannerkorpi and Bjelle 1994). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though it was underpowered, thus indicating that the favourable outcomes are only trends. Given the limitations of the inclusion criteria and other confounding variables, the impact on measures by the assessor, patients and addition of controls post to the intervention after the study period; all impact on delineating a clear line between the cause and effect of the intervention. This obscurity also increases the chances that the outcomes may be as a result of the natural history of the condition or chance.<br><br>Thus, this study requires repeating with a larger sample to allow for parametric analysis. In addition, all weaknesses to the external validity of the study need to be considered and appropriately limited. |
| Conclusion: | Based on the review of the external validity of the study, it becomes apparent that there is a moderate to high risk of potential bias present in the study, which confounds the ability of the study to draw firm conclusions because of the intervention. This is also seen in the internal validity review by the reviewers, where with a 97% agreement, they indicate that there is at best moderate bias. Cumulatively then the study is at risk of moderate to high levels of bias, which indicates that there is little to no evidence available in this study to promote the tested intervention for shoulder pain in RA patients at a clinical level. Thus, this study requires repeating with a larger sample to allow for parametric analysis. In addition, all weaknesses to the external validity of the study need to be considered and appropriately limited. |

Table 4.26a: Tabulated feedback of reviewer data

| Author/s | Manning, V. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2013 | | | | | |
| Title | Exercise-based upper limb rehabilitation in rheumatoid arthritis (Chapter 5: RCT) | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | No | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| | Total score | 9/11 | 8/11 | 8/11 | 8/11 | |
| | | | | | Overall percentage agreement | 97% |

Table 4.26b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Manning, V. |
| Year | 2013 |
| Title | Exercise-based upper limb rehabilitation in rheumatoid arthritis (Chapter 5: RCT) |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Disease of the arm, shoulder, and hand questionnaire (DASH)<br>- Grip ability test (GAT)<br>- Global upper limb function test (timed two ADL – dressing and eating)<br>- RAQoL<br>- Arthritis Self-efficacy scale (ASES)<br>- DAS28<br>- VAS (pain and fatigue)<br>- Duration of morning stiffness<br>- Assessor's assessment of disease activity<br>- Max isometrics strength (upper limb and hand) | Baseline, week-12 and -36. | 36-weeks | 108 total enrolment, EXTRA (n = 52), Usual care (n = 56).<br><br>Loss to follow-up:<br>n = 14 (12-weeks)<br>n = 5 (26-weeks)<br><br>Intention-to-treat utilised. | Yes | Usual care group as a comparison/ control group. | Yes, randomised by 3rd independent party using a random number generator | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 97% |

| | |
|---|---|
| Limitations | This is a grey literature inclusion. The study has a sample drawn from multiple centres (including a secondary care clinic database and three outpatient clinics of rheumatology departments at three United Kingdom Hospitals in the NHS), located in south-east London, United Kingdom. Given the recruitment from several different sources, it is more likely that the sample represents RA patients generally. It is however potentially problematic if the sample derived from the difference sources where inequitable between the groups. This would allow bias from the vantage point of patient naivety, treatment novelty and the representativeness of the two groups of general RA populations (all of which affect generalisability of the results), quite aside from group homogeneity concerns (which detracts from intervention comparisons) (Wells *et al.* 2012).<br><br>The recruitment criteria required patients to have RA (1987 ARE criteria (Arnett *et al.* 1988)), >18 years of age and >5 years disease duration. Exclusion included was a change in biologic DMARD therapy in three months prior to the study; had received steroid intra-articular/muscular injections in four months prior to the study, had surgical intervention and / or used physical therapy modalities to the upper limb during in the six months prior to the study, and finally inability to write. One inclusion (DMARD) was changed in the last six months of the recruitment period. This potentially changed the ability to measure the intervention outcomes. Notwithstanding, the study omitted consideration RA serology and disease activity and imaging studies (Kendall 2003). Of the 316 eligible participants, 120 agreed to participate, while only 108 were available at baseline assessment. These patients were randomised by concealed allocation (Mouton 2006) by an independent party (random number generator) into two-groups, EXTRA group (n=52) and usual care control (n=56). Therefore, the study was protected against researcher bias and non-comparability between groups (de Boer *et al.* 2015), which offers protection towards reporting of intervention effects alone (Fogg and Gross 2000; Mansournia *et al.* 2017). At baseline, there was a non-significant difference between the groups in terms of demographic (age, height, weight, BMI, disease duration, ethnicity, employment status) and general health (self-reported smoking status and number of cigarettes per week, number of comorbidities) measures. The only significant differences were more males in the usual care group ($p \leq 0.05$) and higher disease activity in the EXTRA group ($p \leq 0.05$).<br><br>An appropriate sample size (N=100; n=50) (defined by power analysis), was met at baseline and assisted in increasing the chances of detecting change in primary outcome (DASH) at the 12-weeks assessment (Fogg and Gross 2000; Gummesson, Atroshi and Ekdahl 2003; Rothwell 2006; Bahreini *et al.* 2020). The researcher provided for dropouts through setting recruitment at 60 per group (20% buffer) unfortunately this was not fully achieved, the recruitment achieved (N=108; n=52 and 56). This did not protect against underpowered outcomes (Fogg and Gross 2000; Kendall 2003; Mouton 2006; Christley 2010), as final numbers were well below 50; as after baseline the study suffered dropouts (viz., 14 (six EXTRA and eight usual care) at 12-weeks, and five (three EXTRA and two usual care) at 36-weeks). As a result, the study fell below the required number to appropriate power necessary to determine the cause-effect relationship between interventions and outcomes, and thus impacts on the ability to draw definitive clinical conclusions (Mouton 2006; Christley 2010). This is compounded by a comparison between dropouts with the completers, reflecting that, completers had significantly greater BMI, PADA, weaker NDOM shoulder flexor strength, and weaker NDOM elbow extensor, and flexor strength, which contribute to obscuring the outcomes attained by the intervention. |

| | |
|---|---|
| | The usual care intervention comprised of the medical care standard protocol (pharmacology, self-management education, help-line details, "arthritis research UK information booklet", and were found necessary adjunctive therapy). The EXTRA group has the same as the usual care group in addition to the EXTRA programme. The EXTRA programme consists of a short, individually prescribed programme (per set out protocol) with upper limb education, self-management, and exercise (home based and supervised sessions). The six exercises selected from a specified list of 16, well described functional exercise, with three sets of eight to 12 repetitions, at an intensity of 50 to 80% of 1RepMax at moderate contractile speed carried out daily. Progression was implemented during the first two supervised group sessions (by variance in resistance of therapy -putty and -bands, if necessary, by adjusting the selected exercises); thereafter, by means of the Borg's scale. These exercise sessions were preceded by warm-up and completed with cool-down sessions (both 5 minutes of low-intensity aerobic, mobility and stretching). During the supervised group (four to six participants each) sessions carried out in the physiotherapy department (each participant could choose from a range of times, with attendance twice a week for two weeks), incorporating 15-minutes of educational seminars covering a range of education and self-efficacy topics, followed by the individualised exercise/s. The participants were to carry this same exercise component out daily at home for 12-weeks. Both therapist and participants received a detailed and clearly illustrated handbook of these exercises. Adherence was recorded by exercise diary (70% returned, reporting that adherence to all six exercises at least six days a week by 73%) for the home-based component and the supervised sessions by attendance log (71% attended three or more).<br><br>The use of two assessment sites improved participant accessibility and aided retention. Assessments were conducted according to a clear and standardised protocol by a blinded assessor (blinding breech were reported as 42%) (Pater *et al.* 1998)). The use of a balanced set of subjective and objective outcomes (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018) mitigated against the influence of patient factors on the subjective outcomes, such as novelty, naivety, Hawthorn, placebo and other effects (Mouton 2006). Outcomes measures were mostly valid and reliable (Lorig *et al.* 1989; Dellhag and Bjelle 1995; de Jong *et al.* 1997; Gummesson, Atroshi and Ekdahl 2003; Bearne, Coomer and Hurley 2007; Hewlett, Hehir and Kirwan 2007; Raven *et al.* 2008; Uhlig, Kvien and Pincus 2009; Primdahl, Wagner and Hørslev-Petersen 2011; Ricker, Vergauwe and Cowan 2016) but not all were specific and sensitive to clinical changes in RA. All outcomes followed a clear and standardised protocol (Fess 1992; Bohannon 1997), with instruments calibrated multiple times with no reported deviation (Kimberlin and Winterstein 2008). In general, the outcome measures are selected and reported with validity or reliability (Coster 2013; Bahreini *et al.* 2020). These measures provided consistency, control over and protection of the outcomes (Coster 2013; Bahreini *et al.* 2020). By contrast, the failure in reporting assessment timing (influenced by the diurnal effect or flare-up of the condition (Kaur, White and Bartold 2012; Walker *et al.* 2013; Nandgude, Hasabe and Kolsure 2018), participant blinding (Kaptchuk 1998; Bialosky *et al.* 2011) and Hawthorne (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015) tend to negate the gains made in the rigid study protocol. |
| Outcome | The study reflected that the EXTRA programme for the upper limb in early disease RA patients was able to improve the global upper-limb disability (as per DASH score), with the EXTRA group intragroup analysis showing significant improvement from baseline to 12-weeks ($p$=0.039) and maintained at 36-weeks. Similarly, tender joints significantly reduced at 12-weeks in the EXTRA group of -2.4 ($p$=0.007) but was not maintained at 36-weeks which was also reflected in improved pain at 12-weeks ($p$=0.007) only. By contrast, functional ability (GAT) reflected a significant improvement at both 12-weeks ($p$=0.006) and 36-weeks ($p$=0.008), the DAS28 scores reduced significantly at 12-weeks and was maintained at 36-weeks ($p$=0.004 and $p$=0.011 respectively). With the intergroup assessments, the DASH analysis showed significant improvement in favour of the EXTRA group at 12-weeks ($p$=0.022) only, which is similar to the GAT significant outcomes in favour of the EXTRA group ($p$=0.011), the DAS28 significance at 12-weeks ($p$=0.047) for the EXTRA group and the tender joint inter-group difference in favour of EXTRA at 12-weeks was noted ($p$=0.016). Each of these were not maintained at 36 weeks. Pain in the intergroup analysis was in favour of the EXTRA group at both 12- and 36- weeks ($p$=0.013 and $p$=0.049 respectively). Strength of the upper limb indicated significant strengthening in wrist flexion of the non-dominant hand at 12-weeks ($p$=0.008) and in grip strength ($p$=0.013) in the EXTRA-group, with significant inter-group improvement ($p$=0.021 and $p$=0.009 respectively). In terms of the psychosocial outcome's significant intergroup improvement in favour of the EXTRA-group at 12-weeks (self-efficacy: pain ($p$=0.021) and symptoms ($p$=0.039)), of which self-efficacy (pain) maintained significance at 36-weeks ($p$=0.047). No other disease activity outcomes or strength measures demonstrated significant outcomes. Lastly, the study found no significant adverse effects among participants (Manning 2013). |
| Discussion | The study makes great effort to control for factors influencing the methodological rigour. However, bias from recruitment and under powering of the study amplified baseline differences present between the recruited groups, as well as between dropouts and completers. This reduced the study's capacity to draw conclusive clinical outcomes, given that reported outcomes may result from chance rather than the interventions. This is particularly applicable to the subjective outcomes, which are vulnerable to patient-based influences (e.g., placebo and Hawthorne). The use of *Post hoc* analysis to correct for skewed data, salvages the capacity to draw causation, but risks type II error conclusions. Finally, the intervention design is multi-disciplinary and complex, limiting specific causation links between findings and components of the intervention applied. Additionally, the lack of natural comparison group, limits the measurement of efficacy. Therefore, this high-quality study is hampered by some limitations. |
| Conclusion: | From the analysis and discussion above, the limitations noted, influence the external validity. From this vantage point, the study suffers moderate-to-low risk of bias, and as a result the reader cannot exclude the effect of an external influence/s and associated bias. This is complimented by the reviewer's review, ranking the study as an 8/11 with 97% agreement, indicating high methodological rigour of the study, subject to low risk of bias to the internal validity of the study. Therefore, this study at the best has offers the reader moderate-to-high evidence and requires until further research to determine efficacy of the programme, to strengthen the field practitioner's confidence in implementing into management of RA. |

Table 4.27a: Tabulated feedback of reviewer data

| Author/s | McManus, K. M., Visker, J. D. and Cox, C. C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2015 | | | | | |
| Title | Effect of an Arthritis Foundation Exercise Program on Sleep Quality/Sleep Disturbance in Seniors with Rheumatoid Arthritis: A Pilot Study. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | No | No | No | 67% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | No | Yes | No | No | 67% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | Yes | No | 67% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | No | No | No | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 5/11 | 5/11 | 5/11 | 4/11 | |
| Overall percentage agreement | | | | | | 91% |

Table 4.27b: Properties of the study, outcome, and discussion

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Author/s | McManus, K. M., Visker, J. D. and Cox, C. C. | | | | | | | |
| Year | 2015 | | | | | | | |
| Title | Effect of an arthritis foundation exercise program on sleep quality/sleep disturbance in seniors with rheumatoid arthritis: a pilot study. | | | | | | | |
| Study properties: | | | | | | | | |
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - RAPS<br>- PSQI | Pre/post intervention design | Eight-weeks | 40 participants at start (20 per group). one drop-out from the experimental group. | No | Yes, uncontrolled group. | Recruitment into groups at random, but not randomisation. | 4/11, indicating a moderate-to-low methodological rigour and subject to moderate0to-high risk of bias within the internal validity of the study. | 91% |
| Limitations | Recruitment was by consecutive sampling derived from a senior centre in a Kansas City suburban area of Missouri, USA. Thus, the sample is derived from a general, but older population, resulting in the possible increased disease duration or severity as is associated with more long-standing RA in elderly (Jager, Putnick and Bornstein 2017). This may influence the generalisability to the general RA population, and it may decrease the participant naivety as they have increased likelihood to have been exposed to all RA treatment options (Wells *et al.* 2012).<br><br>The recruited sample was randomly allocated into an experimental group (n=20) and a control group (n=20). The study failed did not report a formal set of inclusion criteria and thus, this is a lost opportunity to control of confounding factors and more comprehensive participant homogeneity. Both of these factors increase the chance of testing the intervention and make the study more clinically relevant (Kendall 2003). Factors that should have been considered include RA diagnostic criteria, functional classification, serology, disease duration and disease activity, patient age, pain severity, imaging studies, medication (type, administration, and dosage) and other physical modalities used in RA treatment (England *et al.* 2019; Myasoedova *et al.* 2020). Carefully structured exclusion criteria would have protected patients against intervention contraindications and protected against omitted or incomplete data (Kendall 2003). From the reported information, 90% of both groups were Caucasian females, however the experimental group had an age range of 76 to 91 years (potentially older), and the control group had an age range from 55 to 88 (potentially much younger).<br>There was no reported power analysis (Kendall 2003; Christley 2010) under powering the study to detect statistical change (Kendall 2003; Mouton 2006; Christley 2010). Dropouts were not controlled for, but that 98% retention rate mitigates against this having an effect on outcomes and the study design obviated the use of intention-to-treat analysis (Gupta 2011). Given the small sample size, only non-parametric statistics could be used, and trends reported (Hoskin 2012; Bahreini *et al.* 2020). To counter the previous assertion of participant naivety (first paragraph), all participants were reported not having had exposure to the intervention programme and all reported infrequent exercise participation prior to the study. This mitigates against patient influence in subjective outcomes, which is significant as this study solely uses subjective reporting as an outcome measure (Mouton 2006).<br><br>The Arthritis Foundation Exercise Programme (AFEP) (Callahan *et al.* 2008; Centers for Disease Control and Prevention 2011) used as an intervention consisted of low-impact exercises (seated and standing), were focussed to improve cardiovascular endurance, muscle strength and joint flexibility (Arthritis Foundation 1992). The supervised intervention classes were an hour long, twice a week for eight-weeks. The duration of the study thus allowed for necessary exercise exposure to induce clinical and thus measurable change (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). There is limited clarity on the exercise programme, limiting study replication and also transfer into clinical practice (Heine *et al.* 2012).<br><br>The outcome measures used were self-administered Rheumatoid Arthritis Pain Scale (RAPS) and The Pittsburgh Sleep Quality Index (PSQI), but subjective response forms which may inherently bias the study given patient perception (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018); Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015); memory recall/decay (Mouton 2006; Ricker, Vergauwe and Cowan 2016) and patient expectations related to treatment novelty / naivety (Mouton 2006; Moustgaard *et al.* 2014). The study only recorded immediate responses to the programme (Pater *et al.* 1998). Both the RAPS and PSQI are reported as a valid and reliable tool/s in evaluating pain (Anderson 2001) and sleep quality / disturbance (Buysse *et al.* 1989) respectfully (Coster 2013; Bahreini *et al.* 2020). The primary outcomes, however, were stated to be insignificantly different between groups. Although seasonally no effect was anticipated on the study outcomes, the study did fail to control time-modified confounding (Platt, Schisterman and Cole 2009) [such as the diurnal effect (Nandgude, Hasabe and Kolsure 2018)] and the testing environment (Richardson 2007). These may have influenced the outcomes. Notwithstanding the above, the study had standardised protocols to reduce measurement variability (Coster | | | | | | | | |

| | |
|---|---|
| | 2013; Bahreini *et al.* 2020). With variances do not account for in patient demographics, RA dynamic characteristics and certain outcome measure influencers, their effects may place the study at risk of type II error, even though these variables in some instances were controlled for by *post hoc* analysis. |
| Outcome | The authors state that the study provides preliminary evidence for the use of the AFEP exercise as an intervention programme but indicates that this study - a pilot study – requires more data to confirm or refute their outcomes. The results after an eight-week supervised group-based intervention, reported statistically significant improvement in perceived pain intra and intergroup ($p$<0.05). Whilst, sleep quality and disturbance were improved intra-group (pre/post), but not intergroup. It was therefore concluded that the study provides cautious evidence to indicate participation in the AFEP programme is beneficial and may provide positive effects on perceived pain and some evidence effecting sleep quality (McManus, Visker and Cox 2015). |
| Discussion | From the authors, the reported outcomes were favourable, even though the study was under powered with limited ability to draw strong conclusions. This implies that the authors cannot account for effects of confounding variables, outcome measure influencers, patient factors and statistical amplification or errors, implying that chance or natural history may also have contributed. This negatively affects the study's ability to accurately and fairly compare intervention group to the control without risking bias in outcomes and conclusions. Thus, the outcomes have limited translation into practice and requires further confirmation of the outcomes. |
| Conclusion: | Given the discussion above of the external validity concerns of the publication, there is a moderate to high level of bias that stems from uncontrolled variables that were not accounted for in the study. This seems to concur with the outcomes of the internal validity of the study as ranked by the reviewers, were the overall PEDro score was ranked as 4/11. This suggests a moderate to high level of bias present in the study. Thus, given both reviews, it is evident that the study does not portray more than very limited evidence possibly even no evidence in support of the intervention for RA patients. Thus, in agreement with the authors, this study needs to be repeated with a larger sample (determined by power-analysis) with a clearer description of the participants and small concerns over the methodology and analysis to be considered more carefully, in order to provide a basis to determine the true causal link between the intervention and the outcome. This would then confirm or refute the findings of this study and allow for extrapolation to clinical practice if appropriate. |

Table 4.28a: Tabulated feedback of reviewer data

| Author/s | McMeeken, J., Stillman, B., Story, I., Kent, P. and Smith, J. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1999 | | | | | |
| Title | The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | Yes | Yes | Yes | Yes | 100% |
| 6 | There was blinding of all therapists who administered the therapy | Yes | Yes | Yes | Yes | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 10/11 | 10/11 | 9/11 | 10/11 | |
| Overall percentage agreement | | | | | | 97% |

Table 4.28b: Properties of the study, outcome, and discussion

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Author/s | McMeeken, J., Stillman, B., Story, I., Kent, P. and Smith, J. | | | | | | | |
| Year | 1999 | | | | | | | |
| Title | The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis. | | | | | | | |
| Study properties: | | | | | | | | |
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Pain<br>Timed-up-and-go test (TUG test)<br>HAQ | Baseline (at session-one pre-intervention) and post-intervention at end of week-six. | Six-weeks | Data provided: Total start = 36 Completers: Exercise (n = 17), Control (n = 18). Unaccounted for in publication (n = 1).<br>The study kept the medication violators in the study. | Yes | Yes | Yes | 10/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 97% |
| Limitations | The sample was sourced in a multi-sited recruitment approach, in Melbourne, Australia. Thus, patients came from referring rheumatologists or referring physiotherapists. They identified patients and informed them of the study. This approach provided a wide range of RA patients, more similar to the general RA population (Wells *et al.* 2012). It must however be considered that the study uses a form of convenience sampling, as both referring rheumatologists and physiotherapists seemed to be based at the Royal Melbourne Hospital, this with it a risk of poor representation of the true population (Jager, Putnick and Bornstein 2017). If the sample was truly recruited from within the hospital setting it is more likely that the patients would be less naïve to the various RA treatments available – a consideration the study needs to have controlled for (Wells *et al.* 2012). The inclusion of patients required a diagnosis with RF sero-positive/negative RA (revised 1987 ARA criteria (Arnett *et al.* 1988)), joints impacted by RA and requiring chronic medical therapy, and 10 seconds or more taken to complete the "timed-up-and-go test". Exclusion occurred if patients had joint replacement(s) or any major pathologies other impacting on the lower limb(s), ongoing acute RA joint disease exacerbations (active synovitis, or warmth of joints on palpation, or rising morning stiffness for lower limb joints), major changes to the pharmaceutical therapeutic regiment three months prior to the study, and / or fragile skin. Considering the complexity of RA, there are considerations that should be accounted for and include patient age, RA disease duration, pain severity, joint deformities, medication use, other treatment (physical) modalities in use, contra-indication exclusion (England *et al.* 2019; Myasoedova *et al.* 2020). These controls would allow contextualisation of reported outcomes and thus enable clear application into clinical practice (Kendall 2003). Thus, careful consideration of inclusion and / or exclusion criteria protects against confounding of the outcomes (Kendall 2003).<br><br>The study failed to report a power analysis (Kendall 2003; Christley 2010), which under powers the study (Mouton 2006; Christley 2010), and prevents accurate consideration of the implications of the study results. It is suggested that the patients were blinded, but this can only be to group allocation, but also infers that informed consent could not have been complete or had a generic description so that patient was unaware of group allocations. In this study, 36 participants were recruited and randomly allocated (unreported methodology), into either an experimental group (n=17) or a control (n=18). Within this sample, three participants underwent medication changes, but were retained in the study, inferring bias. There were no reported dropouts and thus no need for an intention-to-treat analysis. It was not clear whether the groups were similar in terms of sex, age, body weight and height.<br><br>As intervention, there was supervised dynamic exercise by means of training sessions with blinded therapist, which assists in reducing therapist bias (Kendall 2003). Exercises were aimed at the knee extensors and flexors. Exercise prescriptive factors was determined by knee peak angular speed at baseline. Participants commenced the 45-minute intervention with warm-up followed by the exercise protocol. The prescription and intervention variables are described clearly (Heine *et al.* 2012) and undertaken roughly at the same time daily, at three-day intervals for 14-sessions over a six-week period. This offers the study control over time-confounding effects (Platt, Schisterman and Cole 2009) like that of the diurnal variance (Nandgude, Hasabe and Kolsure 2018) and provided the necessary intervention period to ensure measurable physiological changes (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). The period between intervention and assessment was not clearly delineated but influences the outcomes (Pater *et al.* 1998; Nandgude, Hasabe and Kolsure 2018).<br><br>Outcome measures used a balance of subjective and objective tools (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Each outcome was assessed by one of two blinded assessors. In order to ensure inter-assessor reliability, training should have been provided, but was not reported (Kendall 2003; Moustgaard *et al.* 2014), however the blinding is appropriate and good for reducing bias. The outcomes included VAS and HAQ as retrospective subjective outcomes are at risk of memory recall/decay bias | | | | | | | | |

| | |
|---|---|
| | (Mouton 2006; Ricker, Vergauwe and Cowan 2016), Hawthorn effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015), treatment novelty and treatment naivety (Kaptchuk 1998; Mouton 2006; Bialosky *et al.* 2011; Moustgaard *et al.* 2014; Ricker, Vergauwe and Cowan 2016). The remaining outcomes included TUG, a functional assessment of knee musculature (peak torque) and a pedometer. The outcomes assessment procedures were detailed (Harding *et al.* 1988; Podsiadlo and Richardson 1991; Stillman and McMeeken 1995; Stillman, McMeeken and Macdonell 1998), however some outcomes – like peak torque failed to consider the effect of optimal contractile angles and patient learning for outcomes such as PEAK-2D and KinCom dynamometer. These small concerns, introduce bias, and protocol reliability. On a positive note, the study carried out weekly contact with the control participants in order to assess general health and their well-being. This was utilised by the study to minimize Hawthorne effect and control its effect on the outcomes in the exercise group (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015).<br><br>The study seems to have lost a participant, but does not report drop out, which has implications for understanding results (Fogg and Gross 2000) and exacerbates the small sample size, that may not have been similar at baseline, indicating variance can be the cause of change and not the intervention (Fogg and Gross 2000). The data presented in this study are only of the completers, as this study could not by its pre-post design use intention-to-treat analysis (Heritier, Gebski and Keech 2003; Gupta 2011). The drop out and retention of three participants that received modified medication therapy all confound the outcomes of the small sample (Fogg and Gross 2000; Kendall 2003; de Boer *et al.* 2015). As a result of the sample and group sizes (small) this limits the study to reporting non-parametric statistics (Hoskin 2012; Bahreini *et al.* 2020) and thus only trends. Thus, it would have been useful for the study to compare changes to the MCID values but this is not reported in the study preventing contextualising of outcomes into the clinical setting (Fogg and Gross 2000) (Rothwell 2006; Bahreini *et al.* 2020). |
| Outcome | The study conducted a six-week specific dynamic exercise programme directed at the knee musculature of RA patients that are non-acute in phase and disease activity. The study found that physiological improvement was noted in the isokinetic torque had significantly improved at both speeds of $60^{o}/sec^{-1}$ and $120^{o}/sec^{-1}$ ($p$= 0.02 and 0.003 respectively). The TUG reported significant improvement ($p$=0.01), as did pain ($p$=0.03) and the HAQ ($p$=0.04). Because of these findings, the study concludes that specific and targeted training of knee musculature produced functional benefit in the non-acute RA patient (McMeeken *et al.* 1999). |
| Discussion | This study indicated that the outcomes of this study were significantly favourable, even though it was underpowered and affected by limitations in patient variance; measurement of outcomes; and patient factor influences. Importantly there is also omission of one patient in the reporting, which questions the reliability and accuracy of general reporting. Thus, the outcomes of significance in favour of the exercise for the knee, these outcomes may not be better than chance. The study findings are thus limited based on questions of reliability and cannot be extrapolated to clinical practice |
| Conclusion: | Thus, even though this study was rated as 10/11 indicating a high degree of methodological rigour on the PEDro scale (high level of internal validity), there are some compromises that have been made with regards to the external validity of the study, suggesting that the study is at risk of moderate bias. Thus, when considered together the level of bias is low to moderate, indicating that this study may only be able to provide limited evidence for support of the intervention tested for RA knee patients.<br><br>This conflict in review, requires that the study be repeated with due consideration for external validity compromises that this study made before the exercise regime is provided to patients in clinical practice. |

Table 4.29a: Tabulated feedback of reviewer data

| Author/s | Meeus, M., Hermans, L., Ickmans, K., Struyf, F., Van Cauwenbergh, D., Bronckaerts, L., De Clerck, L. S., Moorken, G., Hans, G., Grosemans, S. and Nijs, J. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2015 | | | | | |
| Title | Endogenous Pain Modulation in Response to Exercise in Patients with Rheumatoid Arthritis, Patients with Chronic Fatigue Syndrome and Comorbid Fibromyalgia, and Healthy Controls: A Double-Blind Randomized Controlled Trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | No | Yes | 67% |
| 5 | There was blinding of all subjects | Yes | Yes | Yes | Yes | 100% |
| 6 | There was blinding of all therapists who administered the therapy | Yes | No | No | No | 67% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | No | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | No | Yes | Yes | 67% |
| Total score | | 10/11 | 7/11 | 7/11 | 8/11 | |
| Overall percentage agreement | | | | | | 88% |

Table 4.29b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Meeus, M., Hermans, L., Ickmans, K., Struyf, F., Van Cauwenbergh, D., Bronckaerts, L., De Clerck, L. S., Moorken, G., Hans, G., Grosemans, S. and Nijs, J. |
| Year | 2015 |
| Title | Endogenous Pain Modulation in Response to Exercise in Patients with Rheumatoid Arthritis, Patients with Chronic Fatigue Syndrome and Comorbid Fibromyalgia, and Healthy Controls: A Double-Blind Randomized Controlled Trial. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Verbal numerical rating scale (VNRS) Pain<br>Temporal summation of pain (TS)<br>Conditioned pain modulation (CPM)<br>Pain Pressure Threshold (PPT) | Pre/post per intervention period of crossover design. | Nine-days (One-week wash out period). | Total sample (n=53): Control (n=18), RA (n=16), CFS/FM (n=19).<br>Drop-out of one CFS/FM participant research unrelated. Thus completers: Control (n=18), RA (n=16), CFS/FM (n=18). | Yes | Yes, a healthy sedentary pain-free control. | Yes, randomised allocation by means of a computer programme. | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 88% |

| | |
|---|---|
| Limitations | The study was conducted at the Research Unit at the University of Antwerp Hospital in Antwerp, Belgium, the source of the sample is assumed to be from the greater Ghent community, however this was not detailed in the study. This assumption does not refute or confirm whether the participants were reflective of a specific RA population or the general RA population, which limits study generalisability and also questions the degree of participant naivety (Wells *et al.* 2012). The study design utilises populations of two RA disease profiles and a healthy control. The two chronic disease profiles included an RA patient group and a chronic fatigue syndrome and fibromyalgia (CFS/FM) group. The study included only 18-70-year-old female patients (limiting generalisation). The chronic conditions required diagnostic compliance with diagnostic criteria/s based on grouping the CMP/FM (had to meet the ACR diagnostic criteria for FM (Wolfe *et al.* 2010)) and the centre of disease control diagnostic criteria for CFS (Fukuda *et al.* 1994); whereas the RA participants had no defined diagnostic criteria. This latter introduces bias (England *et al.* 2019; Myasoedova *et al.* 2020). The healthy controls had to be pain free and sedentary (< three hours of moderate physical activity per week). This healthy control and because of the crossover design the inherent RA comparative group provide a basis for evaluating efficacy and effectiveness.<br><br>Confounders where limited by considered exclusion criteria (Kendall 2003), which included pregnancy, post-natal period (one-year), participants who were unable to stop essential medications for two-weeks and NSAIDS / paracetamol 48 hours prior to starting the study. These medications controls provide protection against their influence over outcome measures, but specific controls throughout the study are not provided (DMARD and steroid therapy) and possible changes (Zhao *et al.* 2017; Heidari *et al.* 2019). Unlike many other studies, this study attempted to control for lifestyle factors as patients were not able to consume caffeinated drinks, alcohol, and / or nicotinic substances (Pedersen *et al.* 2006a; McInnes and Schett 2011; Yang *et al.* 2017; Okada *et al.* 2019; Myasoedova *et al.* 2020). In addition, the study limited completely any physical exertion (eliminating a therapy/intervention confounder) (Mouton 2006). All of these limitations provide for very good control of RA and external influencers, but the duration of control for these and adherence is not noted, which increases risk of bias due to premature fatigue, synergistic / antagonistic effect of interventions and their respective carry-over effects (Robins, Hernán and Siebert 2004). This is particularly true as baseline comparisons regarding paracetamol and placebo were reported in another publication elsewhere (Meeus *et al.* 2013). Similarly, criteria not considered in terms of the RA population include RA serology, disease duration and disease activity, imaging studies and other physical modalities in use (England *et al.* 2019; Myasoedova *et al.* 2020). This limited clinical translation of the results due to confounding (Kendall 2003). For all groups, contraindications to the intervention did not seem to have been included (Kendall 2003). Thus, preventing clear patient description, analysis of group variance (de Boer *et al.* 2015) for external influences (Pocock *et al.* 2002) and potential of type II error (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). This is exacerbated by a sample size of less than 20 patients per group, limits the study to non-parametric statistics and trend reporting (Hoskin 2012; Bahreini *et al.* 2020). Given the four groups, it would have been important for a comprehensive demographic evaluation, but the information is limited to the reporting of significant difference in RA participant age (compared with other groups) only. In the RA group, patients were 54.25 years old (mean) (SD+/-8.36 years) compared to the control (41.06 +/-14.48 years) [$p$=0.002] and compared to the CFS/FM group (mean 44.58 +/-7.34 years) [$p$=0.028]. These differences required statistical correction by *post hoc* analyses (de Boer *et al.* 2015; Unbiased Research 2017).<br><br>The groups were randomised (viz., RA and CFS/FM and healthy participants) by means of means of an independent coordinator using a “computer programme” (Kendall 2003), providing concealment (Mouton 2006) and strengthening group comparability (de Boer *et al.* 2015) and eliminates researcher bias associated to group allocation and other external influencers (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019) in to two groups. These groups (post-randomisation) were not detailed by distribution. |

| | |
|---|---|
| | An *a priori* power analysis (effect size 0.25 and 90% power) indicated a total sample of 39, performed for ANOVA analysis of three-groups in four-outcomes (TS with/out CPM under both intervention arms (paracetamol and placebo)) (Kendall 2003; Christley 2010). As a result, a total sample of 53 participants recruited into the study, met power analysis requirements and made provision for dropout (Fogg and Gross 2000). The initial placebo group and the paracetamol groups received an identical pill (by appearance) from the study nurse (not involved in the outcome assessment), after 30 minutes they underwent baseline outcome assessment. After this assessment, the participants completed a submaximal exercise protocol and repeated the outcome assessment. This was followed by a seven-day interlude, prior to returning to repeat the process as the other group (placebo group becomes the paracetamol groups and vice versa). As befalls each cross over study, the wash-out period for the medication confounding and synergistic effects of multiple exercise sessions over a short period of time may be problematic (Robins, Hernán and Siebert 2004; Evans 2010). The actual intervention was submaximal exercise delivered by means of stipulated, prescribed exercise (e.g., heart rate <75% predicted maximum (age-related calculation) for <15 minutes avoiding fatigue) on an ergometer bicycle following the "Aerobic Power Index Test" allowing repetition (Wallman *et al.* 2004; Heine *et al.* 2012). Given this measure is not validated nor reliable for the RA population, it may influence outcomes of RA patients (Coster 2013; Bahreini *et al.* 2020). However, the set protocol provides consistency to the study and protects against the variability in outcomes measure between groups (Kendall 2003).<br><br>Each evaluation required completion of a balance of subjective and objective outcome measures (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). These included VNRS, Pain Pressure Threshold (PPT) (Fischer algometer) TS and CPM (Cathcart *et al.* 2009). A single assessor completed all protocol driven outcome measures providing control over inter-assessor variability and repeatability and thus any associated bias (Kendall 2003; Moustgaard *et al.* 2014) [it was noted that only 54% of patient group allocation was identified]. However, given that TS and CPM are calculated from the PTT, thus any patient factor – novelty, naivety, prior experience, unmet expectation, Hawthorn effects and pressure thresholds are possible influencers of PTT and thus the calculated TS and CPM (Dhondt 1999; Melia *et al.* 2019). This may create unwanted floor or ceiling effects (Robins, Hernán and Siebert 2004) or synergies / antagonisms (Liang 2000; Coster 2013). With this and the unclear environmental influence of these assessments such as time-sensitive confounding (Platt, Schisterman and Cole 2009), there is concern for groups not being equivalent and thus risking type II error. The design of the study (pre/post the two intervention periods per period of the crossover), which prevents implementation of the intention-to-treat analysis (de Boer *et al.* 2015; Unbiased Research 2017). This crossover did therefore not consider carry-over of treatment effect.<br><br>The authors acknowledged that the influence of placebo could not be excluded (Kaptchuk 1998; Bialosky *et al.* 2011) and that the exercise test, my not have been sufficient to induce its associated analgesic effect in both chronic disease profiles. In addition, the study failed to compare MCID values required to determine clinical significance (Bahreini *et al.* 2020). |
| Outcome | The authors reported exercise induced analgesia ($p$<0.05) in the RA participants in both paracetamol and placebo conditions after submaximal exercise (25W, 75% age-predicted heart rate max for 15 minutes). This was less clear in the CFS/FM group where the results were inconclusive and inconsistent (Meeus *et al.* 2015). |
| Discussion | The authors indicated that the outcomes were significantly favourable for the RA population. These findings were limited by the sample size, yet according to the power analysis the numbers were sufficient to draw conclusions on parametric statistics. The obviously unclear methodology that is sometimes seemingly contradictory in its reporting in terms of patient recruitment, RA diagnostic criteria, limited reporting on and perceived inconsistency of demographic data, limits contextual understanding and thus interpretation of the outcomes. From what is reported, there seems to be moderate bias based on the inclusion criteria used, which seems to be countered by a fairly homogenous study group. The outcomes although not all valid and reliable, were consistently applied by one assessor. The extrapolation to clinical practice is limited as the reporting of RA serology, disease duration and activity, and articular destruction/deformity, hinders generalisability of the outcomes. Although the exercise protocol used was validated for CFS, it was not for RA which questions its ability to measure effective change. This in addition to the study not controlling floor / ceiling effects, synergistic influence and placebo effect and exercise stimulus may have been insufficient to produce accurate results. Finally, the lack of MICD values in the study prevents evaluation of clinically significance. |
| Conclusion: | Given the significant number and type of limitations from t/he external validity vantage point, this study potentially suffers from significant bias / external influence. This is in contrast to the review by the reviewers, which ranks the study as an 8/11 with 88% agreement, indicating that there is high methodological rigour in terms of the PEDro scale. This implies that there is a tension between the two reviews. It is however possible to have a rigorous study that, has been well structured, which has influencers outside of the research process that (e.g., the choice of a measurement tool that is not necessarily appropriate) compromise it. Therefore, this study at the very best has a moderate level of bias that is shown through both review processes, which indicates that there is a moderate degree of evidence and that this should be considered when applying these results tenuously in practice and with caution, as no evidence for MCID comparison is provided. |

Table 4.30a: Tabulated feedback of reviewer data

| Author/s | Metin, Z. G. and Ozdemir, L. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2016 | | | | | |
| Title | The effects of aromatherapy massage and reflexology on pain and fatigue in patients with rheumatoid arthritis: a randomized controlled trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | No | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least one key outcome | No | No | No | No | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by intention-to-treat | No | Yes | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 6/11 | 6/11 | 6/11 | 6/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.30b: Properties of the study, outcome, and discussion

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Author/s | Metin, Z. G. and Ozdemir, L. | | | | | | | |
| Year | 2016 | | | | | | | |
| Title | The effects of aromatherapy massage and reflexology on pain and fatigue in patients with rheumatoid arthritis: a randomized controlled trial. | | | | | | | |
| Study properties: | | | | | | | | |
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| VAS (pain) FFS (fatigue) | Base line and once a week within 1 hour of treatment finishing. | Six-weeks | 54 participants with three dropouts. Leaving 17 per group completing study. | No | Yes, the control did not receive sham versions of the interventions. They continued usual care. | Yes, by random number table. The study was also stratified. | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 94% |
| Limitations | The study recruits' participants via a specialised rheumatology clinic at a university hospital (Turgut Ozal University) (Turkey). The impression is instilled that given the nature of the source from which participants are drawn, increases the likely of recruited of a sample that is poorly controlled for confounding introduced through this, such as the those suffering a greater degree of severity in disease activity (Wells *et al.* 2012). The method of recruitment was that of convenience sampling, the introduces risk to the study to the external validity of the study (Jager, Putnick and Bornstein 2017). As a result, risks unintentionally limiting the generalisability of the study, and introducing confounding factors and bias that the study design is not controlling for. In such a scenario the outcomes reflect external influence rather than intention. Furthermore, such a sample is unlikely to be naïve to or consider the intervention novel (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013).<br><br>This study clearly detailed the inclusion criteria, although the participant must have a RA diagnosis, confirmation (such as use of a criteria) was not included. Serology and functional class were omitted. The biggest factor missing from the inclusion criteria is the need for RA to affect the participating patients' knees in view of the fact that the aromatherapy treatment appears specifically to target knee symptomatology. In contrast to other studies in this review, there was a detailed account of the medication, but the stratification was limited to medication classes as opposed to specific medications. Randomization was applied through random number tables; however, it was not stated that concealed allocation was utilised, thus introducing allocation bias into the study (Mouton 2006). In patient terms, the study made use of stratification for the dynamic characteristics of RA, namely, duration of the disease, pain, fatigue and / or DAS28 scores, (with the latter being similar between the groups at baseline) and medication (which was only limited to DMARDs or DMARDS with steroids) but excluded serology and functional class as well as comorbidities (Kendall 2003; England *et al.* 2019; Myasoedova *et al.* 2020).<br><br>An *a priori* analysis indicated a minimum sample of 17 per group with 51 total participants to detect an effect size of 0.3 at 80% power. Not having a minimum 20 per group implies the use of non-parametric statistics (Hoskin 2012; Bahreini *et al.* 2020). Despite this, parametric statistics were calculated in the study, along with nonparametric statistics. The reason for this is unclear as the manner in which dropouts were considered is unclear, control was not in place and intention-to-treat analysis was not noted (de Boer *et al.* 2015; Unbiased Research 2017).<br><br>The reflexology technique used was clearly described, thus allowing for the intervention to be repeated. However, there was no explanation for the aromatherapy massage application only to the knee, and the usual care received by the control group is not described. Medication changes were not recorded for any of the groups, though patients were instructed to not take analgesic drugs on the days of treatment. This was not complied with in the usual care group, which provided limited control over the medication effect. Patient effects such as the placebo, Hawthorne, observer and attention effects, timing (to account for diurnal effects on RA (Nandgude, Hasabe and Kolsure 2018)), the therapeutic touch effects (groups had different treatment visit numbers) (Melzack and Wall 1965; Moayedi and Davis 2013; Senderovich *et al.* 2016). and the learning process with regard to repeated measures were not addressed (Bordens and Abbott 2002). Reporting of patient naivety in terms of treatment was described, but the exposure patients had to interventions, prior to outcome assessment (within an hour post-intervention) and the effect this has on the outcomes of this study are not considered (Pater et al 1998). The implications are that the control is potentially not naïve to usual care and may react adversely to being in this group while those not exposed to the interventions may report better benefit from the novel care given (Moustgaard *et al.* 2014). Given that the study only has subjective outcomes this results in a potential for type II error where improvement is due to perception (Kendall 2003) and not the intervention applied (Kaptchuk 1998; Bialosky *et al.* 2011).<br><br>The repeated measures were in the form of the Visual Analog Scale (VAS) (Scott and Huskisson 1979; Eti Aslan 2002) and the Fatigue Severity Scale (FFS) (Schwartz, Jandorf and Krupp 1993; Gencay-Can and Can 2012) which are reliable and valid, and suited (Coster 2013; Bahreini *et al.* 2020) to measurement of change in pain and | | | | | | | |

| | |
|---|---|
| | fatigue in RA patients. Additionally, the VAS and FFS scores were not different between the groups at the baseline. Nevertheless, the impact of the assessor being unblinded (as the therapist, interviewer and deliverer of VAS and FSS assessment are all carried out by a single researcher) to the intervention group of participants implies that there may be a measurement bias that impacts on the outcomes of the study (Kendall 2003; Moustgaard *et al.* 2014).<br><br>Finally, in terms of the interventions, reflexology has two mechanisms of action, firstly, local and, secondly, global action effects (Tiran and Chummun 2005). Aromatherapy massage is cited to have more than one means of action one would be in the form of – the massage component, and another through variable essential oil/s utilised for said treatment – acting in several ways including sensory (smell) and the selected oils effect on locally, and systemically when absorbed (Steflitsch and Steflitsch 2008; Hongratanaworakit 2011). These contrast with the usual care that has an undetermined number of actions. These actions do not all impact pain or fatigue directly but instead have indirect means of impacting these outcomes, sometimes synergistically, sometimes antagonistically (Robins, Hernán and Siebert 2004). In order to better and more comprehensively understand the impact on RA patients of the respective treatments a more comprehensive set of outcome measures would have been more appropriate (Coster 2013). |
| Outcome | The authors report that the aromatherapy massage and reflexology interventions were superior to usual care in significantly reducing pain and fatigue scores in RA patients from the second to the sixth week outcomes ($p<0.05$). Pain reduction resulting from aromatherapy massage started later than pain reduction resulting from reflexology. Reflexology was noted to have a greater effect than aromatherapy massage in reducing pain and fatigue scores in RA subjects. In the light of the last paragraph under Limitations above, the latter two outcomes are not unexpected. Thus, the conclusion was that reflexology proved more effective than aromatherapy massage with regard to reducing fatigue scores from the beginning of the intervention. Nonetheless, starting from the fourth week of the study aromatherapy massage also decreased fatigue scores (Metin and Ozdemir 2016). |
| Discussion | Given that the study compared interventions that were very different both in effect and outcome for the condition of RA, it is important that all the researchers ensure that there is a comprehensive baseline analysis of RA characteristics. This would enable them to direct the outcome more specifically at the intervention and not the lack of patient homogeneity. In this study, for example, it is not clear what pain from which region was recorded by the patients. This is of particular importance as the aromatherapy was applied locally to the knee, implying the knee to be the primary area to benefit from the therapy, but it is not clear if the patients were all symptomatic for knee pain and were reporting the same pain. This strategic concern is further compounded by the lack of concealed allocation as well as the lack of patient effects being addressed and, additionally, the lack of blinding of the assessor. |
| Conclusion: | Given the limitations noted above, from the external validity vantage point this study suffers from a moderate degree of bias and the reader cannot exclude the effect of external influence/s. This is in line with the review by the reviewers, which ranks the study as an 6/11 with 94% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study in terms of the PEDro scale. Therefore, at best this study can suggest that these interventions are potentially clinically effective as an adjunct to usual care as the comparison is usual care to reflexology versus usual care to aromatherapy. However, the extent of these positive clinical effects (when read alone) needs to be interpreted with caution, as the treatments are only applicable to patients who meet the inclusion criteria for this study. The significant effects from the factors affecting external validity must also be taken into account as these may negate the effects of the clinical improvement. The study therefore provides limited evidence in support of the use of these two interventions alongside usual care. |

Table 4.31a: Tabulated feedback of reviewer data

| Author/s | O'Brien AV, Jones P, Mullis R, Mulherin D, Dziedzic K. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2006 | | | | | |
| Title | Conservative hand therapy treatments in rheumatoid arthritis: a randomized controlled trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | No | Yes | Yes | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | Yes | No | No | No | 67% |
| 6 | There was blinding of all therapists who administered the therapy | No | Yes | No | No | 67% |
| 7 | There was blinding of all assessors who measured at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention-to-treat" | No | Yes | Yes | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 9/11 | 8/11 | 8/11 | |
| Overall percentage agreement | | | | | | 82% |

Table 4.31b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | O'Brien AV, Jones P, Mullis R, Mulherin D, Dziedzic K. |
| Year | 2006 |
| Title | Conservative hand therapy treatments in rheumatoid arthritis: a randomized controlled trial. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Primary outcome:<br>AIMS 2<br>- Upper limb<br>- Hand and fingers<br>Secondary outcomes:<br>Jebsen-Taylor hand function test (right hand)<br>Power/Gross grip<br>Key pinch<br>Dominant hand index finger flexion (right hand)<br>Disease activity<br>Swollen and tender joint scores<br>Perception of their own disease activity | Baseline, months -one, -three, -six<br><br>All had 30 min appointment at start to deliver intervention and relevant issues regarding leaflet.<br><br>Two-weeks latter 15 min follow-up to monitor concordance and requested to inform assessors of any treatment changes and to do so at each follow-up. | Six-months | 67 at start with Group 1 (n = 21), Group 2 (n = 24), and Group 3 (n = 22). A 78% follow-up was achieved at six-months, thus a total dropout of 15.<br><br>Completers: At one month's follow-up: Group 1 (n = 19), Group 2 (n = 20), and Group 3 (n = 19). At three months' follow-up: Group 1 (n = 18), Group 2 (n = 17), and Group 3 (n = 19). At 6 months' follow-up: Group 1 (n = 18), Group 2 (n = 16), and Group 3 (n = 18). | Yes | Yes | Yes, | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 82% |

| | |
|---|---|
| Limitations | The 67 participants (mean age 59.6 years) in the study were recruited from a Rheumatology Department in Mid-Staffordshire general hospitals in the United Kingdom NHI trust. With this come concerns of external validity and the nature of the recruitment source, increases the likely of recruited of a sample that is poorly controlled for confounding introduced through this, such as the those suffering a greater degree of severity in disease activity (Wells *et al.* 2012). Furthermore, the fact the participants were recruited from a rheumatology department suggests that they would not have been naïve to the intervention (Mouton 2006).<br><br>The inclusion criteria included age (>18 years) and a diagnosis of RA (1987 ACR criteria (Arnett *et al.* 1988)). Potential participants were excluded if their medications (DMARDs) were changed three months preceding the study, if oral corticosteroid (>7.5mg/day) or intra-articular / intra-muscular injection had occurred in the month prior to the study. Other exclusion criteria were a surgical history relating to wrist / hand / elbow / shoulder six months prior to the study, sensory impairment of the hand, uncontrolled wrist / hand pain and pregnancy. The 67 participants were divided among three groups as follows: Group 1 (n = 21), Group 2 (n = 24), and group 3 (n = 22). The study did not document RA dynamic characteristics (functional class and disease activity, pain intensity), the use of other physical modalities, contraindications to interventions and RA comorbidities (England *et al.* 2019; Myasoedova *et al.* 2020). Its focus was limited to the hand and wrist, although mention is made of the upper extremity, therefore the impact of handedness on clinical outcomes should have been documented. The failure to do this means that the homogeneity of the sample in terms of RA and between the groups is questioned, which may be partly countered by means of a stratified group allocation process (permuted blocking within strata stratifying) using disease duration (more or less than five years) and serology (RF) (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019). Notwithstanding the above concerns and counterstrategies within the study, the study reports the baseline characteristics as similar between groups in terms of sex, age, disease duration, hand pain and baseline AIMS II scores, although it could be argued that Group 2's participants tended to be older and to have longer disease duration and higher disease activity scores (Pocock *et al.* 2002; Bland and Altman 2011; de Boer *et al.* 2015). Further to Group 2's participants not being homogenous with Groups 1 and 3 in terms of age and RA factors, this is supported by the authors' reporting that they controlled statistically for significant differences, which, particularly in the primary outcome, would put the study at risk of Type II error (in addition to these errors occurring as a result of undocumented underlying relationships (e.g., RA functional class)) (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). This is further compounded by the small sample size (according to *a prior* calculations, the study needed 20 participants per group) |

| | |
|---|---|
| | (Kendall 2003; Christley 2010); an underpowered study easily falls prey to outliers in defining outcomes and is statistically not as powerful as an adequately powered study (especially when medication dosage, administration and adherence have not been controlled for, as is the case in this study) (Mouton 2006; Christley 2010).<br><br>In terms of the intervention/s of the study, all groups received education material in the form of joint protection leaflets (covering a number of RA related content). In intervention groups 1-3, the education offered to all was supplemented by: Group 1, received prescription of hand exercises aimed at stretching and mobility/stretching; Group 2, received mobility/stretching exercises only; Group 3: joint protection leaflets only and served as this studies control. The interventions were administered in a 30-minute consult with therapist during which the education material was delivered and chance to communicate concerns and issues was afforded to the participant. Any additional outpatient therapy received or changes to medication were to be reported to the outcome assessors. The exercise programmes (group 1 and 2) followed a progression protocol intended to increase repetitions over the study duration. The exercises were to be carried out twice a day at home for the six-month study duration. To track adherence to the protocol exercise diaries were completed by the participants and submitted. Aside for the above, the reader is not supplied with sufficient exercise prescriptive factors and instructions for this to be extracted and replicated in clinical practice (Heine *et al.* 2012). However, to the strength of the study the intervention is sufficiently long to allow for manifestation of exercise-based effects (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). As can be seen the study lacks true control comparison group, thus limiting the study to reporting relative effectiveness.<br><br>The above statistical concerns are further highlighted using measures such as the Jamar dynamometer which, despite its being the gold standard in hand-held dynamometers it suffered from the 'floor' effect in this study, limiting its usefulness in measuring change due to the intervention (Liang 2000; Coster 2013). The statistical weakness of small sample also imply that the study is likely to have been influenced by intervention novelty, patient naivety, Hawthorne effects and other patient-centred effects, thus further reducing the possibility that the outcomes were better than chance alone. (Kaptchuk 1998; Mouton 2006; Richardson 2007; Bialosky *et al.* 2011; Moustgaard *et al.* 2014; Sedgwick and Greenwood 2015). On the reduction of bias, the study noted that both assessors were blinded (Kendall 2003; Moustgaard *et al.* 2014), the assessment protocol is clearly stated (Kendall 2003). The circumstances under which these evaluations occurred (environment and timing in relation to the intervention) are not detailed. This, questions the influence of the variability of assessment on the data (Pater *et al.* 1998; Richardson 2007). |
| Outcome | Of the 67 participants recruited 78% follow-up was achieved at six months. The study reported, upper limb function (by means of AIMS-2), improvements from baseline to six-month follow-up reflected an improvement in group 1 while groups 2 and 3 detreated. The between group difference at six-months was statistically significant in the favour of group 1 ($p$ = 0.007). This finding was still significant after, ANOVA adjustment for multiple testing ($p$ = 0.012). Additional statistical analysis demonstrated that difference in upper-limb function against group 3 (control) alone proved highly significant ($p$ = 0.002), while between groups 1 and 2 alone ($p$ = 0.013). However, between groups 2 and 3 insignificant difference was reported. The only significant improvements revealed from the secondary outcomes was noted in arm function in particular the dominant-hand key-grip. Between groups statistically significant trends of change in scores from baseline to six-month follow-up ($p$ = 0.014). This according to the authors demonstrated that RA patients following home-strengthening programme of hand exercises compared with simple stretches or advice alone (O'Brien *et al.* 2006). |
| Discussion | In the study's favour is the fact that the authors detail the protocols and interventions in each of the three groups. The small details, such as instructions given to the participants, are, however, lacking, thus not allowing replication of the study (as research or in clinical practice). This is also compounded in that the control fails to be a true no-intervention group since educational material was provided to the participants in all three groups at the start of the intervention period. Although the educational material does not report any advice on exercise, the information it gives on the influence of joint protection may confound the outcome measures (e.g., regarding pain and disease activity). For this reason, the study's findings must be described as those of an efficacy study rather than of a relative effectiveness study. |
| Conclusion: | Given the significant limitations analysed above, from the external validity vantage point this study suffers from a moderate degree of bias, thus the reader cannot exclude the effect of external influence/s in the reported outcomes. This contrasts with the reviewers' review according to the PEDro scale which reflected a ranking for the study of 8/11 with 82% agreement, indicating that there is high methodological rigour and low risk of bias to the internal validity of the study. This disjuncture is concerning, but it is possible in this case that bias and confounding in this study was induced outside of the research process. Therefore, this study at best, provides the reader cautious reason to implement the intervention, on grounds of the support of moderate evidence. |

Table 4.32a: Tabulated feedback of reviewer data

| Author/s | Piva, S. R., Khoja, S. S., Toledo, F. G., Chester-Wasko, M., Fitzgerald, G. K., Goodpaster, B. H., Smith, C. N. and Delitto, A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2019 | | | | | |
| Title | Neuromuscular electrical stimulation compared to volitional exercise for improving muscle function in rheumatoid arthritis: A randomized pilot study. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | No | No | No | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | Yes | No | No | 67% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | Yes | Yes | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 9/11 | 8/11 | 8/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.32b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Piva, S. R., Khoja, S. S., Toledo, F. G., Chester-Wasko, M., Fitzgerald, G. K., Goodpaster, B. H., Smith, C. N. and Delitto, A. |
| Year | 2019 |
| Title | Neuromuscular electrical stimulation compared to volitional exercise for improving muscle function in rheumatoid arthritis: A randomized pilot study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Muscle structure and function: Quadriceps area, Attenuation, MVIC<br>Performance based physical function: Stair climb, Chair stand, Gait speed, Single leg stance<br>Patient reported: LEFS, HAQ<br>Muscle biopsy (subset of sample): Myocyte area, IMCL content (type 1 and 2 fibre proportion) | Baseline (prior to randomization) and post-intervention (at four months) | 16 weeks | Total 59 | Yes | No, two separate intervention groups. | Yes, randomization sequence generated (concealed). | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study was carried out at the University of Pittsburgh Medical Centre and recruited participants into the study by means of three sources (via rheumatologist of the pre-stated medical centre, by means the medical centres arthritis network, and through public announcement). These "multi-centred" sourcing improves the likelihood that recruitment will draw a sample (within the confines of the set of recruitment criteria) representative of the true RA population (Wells *et al.* 2012). The inclusion of two medical centre sources however risks the of impact due to lack of patient naivety, this may impart influence of the subjective outcomes as perception and previous exposure may bias the participants response (Mouton 2006).<br><br>The sample was drawn based on the inclusion of participants that complied with the following: age (>21), RA diagnosis (1987 ACR criteria (Arnett *et al.* 1988)), disease duration (>5 years) and capable of independent ambulation. Patients were excluded for cardiovascular disease, uncontrolled hypertension, neurological conditions, muscle conditions, history of quadriceps/patellar tendon rupture or adverse reaction to electrical stimulation. Additionally, having participated in resistance training two or more times per week, a change in medication in the previous month, the use of cholesterol lowing medication, malignancy, pregnancy, restricted ROM (<70º knee flexion) were also considered as exclusion criteria. However, RA dynamic characteristics (e.g., serology, functional class, disease activity), RA associated pain, medication control (adherence, type of medication and administration type) or management and some comorbidities / contraindications to therapy were not considered as part of the inclusion (and stratification of patients) or exclusion criteria as pertinent to the study. Even though the study used a randomization sequence and concealed group allocation in order to try and avoid the influence of bias, there is still uncertainty as to whether there was baseline homogeneity between the two groups (given that the study did not use stratification for all variables). This may have been the reason for the baseline characteristics showing differences between groups for age (*p* unpublished) and body weight (*p* unpublished). These differences highlight the possibility that there are other unaccounted for factors that may have been different between the groups. This implies that the outcomes for each of the respective groups may well have been determined more by the patient characteristics that composed each of the groups as opposed to only the intervention. This context does not allow for adequate comparison between the groups in addition to limiting the generalisability of the outcomes. These baseline differences will also impart on the study risk for type II error (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>To the credit of the authors, the study reported a power analysis noting a group size of 27 group (total 54) to be sufficient to report outcomes at an 80% power and CI of 95%. Furthermore, the calculated for and intended for drop out provision through a 10% attrition through which sample target was set at 60 participants. However, the final analysis only included 24 in the neuromuscular electrical stimulation (NMES) group and 26 in the exercise group. The attrition was reportedly addressed through an intention-to-treat analysis; however, its calculation is not described nor whether this was actually included. Given that the groups did not both meet the required 27 participants, means that the groups were differentially powered in order to achieve comparable outcomes. This is also further compounded given that there were differences at baseline, these were adjusted for in the analysis (including sex, height, BMI, ethnicity (White / Black / Hispanic), education, employment, smoker status, disease duration, disease activity (DAS28 and ESR) and leg pain). Adjusting for these variables with a relatively small sample size of N=59 (as noted by the authors), means that there is an increased risk of type II error, particularly for those underlying variables that are unaccounted for. These outcomes data are further compromised by the attrition rate from the study and |

| | |
|---|---|
| | between the groups (viz was three times higher in the NMES group), by limiting effect size (Fogg and Gross 2000; Pocock *et al.* 2002; Kendall 2003; Mouton 2006; Christley 2010; Bland and Altman 2011; de Boer *et al.* 2015).<br><br>The NMES group was compared with high intensity volitional exercise training (no true control group). The protocols of the interventions were detailed and would permit replication specifically for RA patients. Given that the groups had intervention programmes the study at best can only determine relative effectiveness of the programmes. The patients were not blinded to group assignment; therefore, it is good that the study did not only have subjective outcomes as measures of improvement, which limited the influence of intervention naivety, touch effect, placebo effect, Hawthorne effect and environmental conditions. But of concern is that the objective between-group differences in the muscle and physical function were small, implying that the chosen objective outcomes would potentially be at odds with the subjective outcomes, even though the outcomes where well suited and validated for the given purpose of the study. To counter this the study did provide for blinded outcomes assessors, but there was no reported protocol for these assessors training and it is therefore not certain what the effects of multiple assessors had on the study outcomes (Kendall 2003; Heine *et al.* 2012; Coster 2013; Moustgaard *et al.* 2014; Snibsøer *et al.* 2018; Bahreini *et al.* 2020). In defence of the study, the interventions were carried out in comparable conditions, this included settings, timing, and therapist (Pater *et al.* 1998; Richardson 2007).<br><br>Finally, the outcomes measures timeline was not well defined, limiting the readers understanding of the relationship between intervention and evaluation. But on the positive side, the study duration of four months allowed for the effects of exercise to be more accurately measured, as exercise as an intervention requires six to eight weeks to impact the functional capacity and thus ADLs would only show improvement after this period (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012). |
| Outcome | The study reported the effect of NMES versus “high-intensity volitional resistant training” on muscle function and structure as well as physical function in a sample of mildly functionally limited RA patients with moderate disease activity. The study reports on the effect of these interventions carried out for 16 weeks. The study found that intensity of intervention differed (31% in the NMES group and 77% in the training group according to the average intervention intensity scores assessed by MVIC in the NMES group and the percentage 1-RM in the training group), although both groups achieved similar improvements in muscle function. The study showed significant improvements in both groups for muscle structure and function ($p < 0.001 – 0.019$) in respect of attenuation and MVIC values of the NMES group bilaterally (significance at $\alpha < 0.05$). The training group reported left-sided attenuation improvements (significance at $\alpha < 0.01$) and MVIC values bilaterally (significance at $\alpha < 0.05$ (left) and $\alpha < 0.001$ (right). In terms of physical function, no significant difference was noted between groups, but the training group reflected a change in stair-climb and chair-to-stand times of -0.3 and -1.2 seconds respectively, reaching a significance of ($\alpha < 0.01$). In terms of patient-reported physical function, the NMES group reflected a change of 4.2 between assessments (significance of $\alpha < 0.01$). In terms of feasibility analysis, the study did not report events of serious adverse effects, increased pain, or disease activity. However, attrition between groups was significant with NMES reporting 29% dropout and the training group reported 7%. The balance of the outcomes remained insignificantly changed or different. In conclusion, the study reflects evidence that both groups were effective with regard to improvement in muscle function and structure within the RA patient group, thus, potentially offering an alternative intervention (given the comparable outcomes) for the contradicted patient with a poor tolerance for high-intensity exercise (with resistance) (Piva *et al.* 2019). |
| Discussion | According to the authors, the study reflected a mix of significant statistical and clinical outcomes, favourable for the RA patient. These findings were limited by the sample size and exacerbated by dropout, which resulted in differentially powered groups and an unclear intention-to-treat analysis. The study thus draws conclusions on an unclear statistical basis, and the methodological rigour of the study was weakened by several baseline differences between the groups. These require statistical correction in an already small sample size. The attrition of the groups differed substantially and was exacerbated by the lack of blinding of the therapist and participants, the use of multiple blinded assessors without controls for inter-rater reliability and the lack of natural history control. A strength of the study is found in its reporting of clinical significance for findings, which offer clarity to the clinical setting. However, the limited participant description offers a non-specific catalogue for the identification of patients who would benefit most, thus limiting the generalisability of what evidence is offered by the study. |
| Conclusion: | Given the significant number and type of limitations from the external validity vantage point, this study potentially suffers from a significant bias/external influence. This contrasts with the review offered by the reviewers, which ranks the study as an 8/11 with 94% agreement, indicating that there is high methodological rigour and a low risk of bias to the internal validity in terms of the PEDro scale, which implies that the two reviews stand in contradiction of one another. It is, however, possible to have a rigour study that has influencers outside of a research process that has been well structured. Therefore, this study at best needs to be repeated with due consideration given to the external validity limitations noted above, in order to provide a basis for clinical and prognostic consideration. Thus, to conclude, the study offers at best moderate to likely limited evidence for the application of the intervention to the RA patient. |

Table 4.33a: Tabulated feedback of reviewer data

| Author/s | Romanowski. M. W., Spiritovic. M., Romanowski. W. and Straburzynska-Lupa. A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2020 | | | | | |
| Title | Manual Therapy (postisometric relaxation and joint mobilization) in knee pain and function experienced by patients with rheumatoid arthritis: a randomized clinical pilot study. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | No | Yes | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 8/11 | 9/11 | 8/11 | 9/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.33b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Romanowski. M. W., Spiritovic. M., Romanowski. W. and Straburzynska-Lupa. A. |
| Year | 2020 |
| Title | Manual Therapy (post-isometric relaxation and joint mobilization) in knee pain and function experienced by patients with rheumatoid arthritis: a randomized clinical pilot study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Primary outcome: HAQ<br><br>Secondary outcomes: Pain VAS, the Oxford knee score, the knee society score. | Pre/post design Before the first intervention and one day after the last intervention. | 12 days in total (10 days intervention with one day off and one day after the last treatment) | 47 participants (24 intervention and 23 control) with a single drop-out from the intervention group. | Yes | Yes, the control was however, to undergo "standard exercise programme, containing five exercises". | Yes, envelope and coin toss | 9/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study provides limited detail as to the location and source of the sample. However, the sample is drawn from admitted patients (to the rheumatology ward), receiving specialist care meeting the set of recruitment criteria. Based on location of the principal researcher and the ethical approval gained from the "University of medical sciences", both in Poznan (Poland). The reader is left to assume this as that source. Given the nature of the recruitment (following a convenience sampling method) and the specialist medical nature of the ward. The readers are left with the impression that the sample draw is subject to a limited external validity in that a subpopulation of RA are available to the recruitment (Wells *et al.* 2012; Jager, Putnick and Bornstein 2017). Furthermore, this limits that likelihood of novelty and naivety of the sample, which influences subjective outcomes through previous experience and perceptions carried into the study (Mouton 2006). Participants to be eligible had to meet a well-considered set of recruitment criteria, these included a RA diagnosis meeting the 2010 ARC criteria (Aletaha *et al.* 2010), 18 years or older, experiencing knee pain unilaterally (VAS score of four or above) on a daily bases, evaluated disease activity according to DAS28 of no higher than 5.1, sonographic disease activity grading of 1 or less (Bhasin and Cheung 2015). While exclusion was based on changed to dosage of NSAIDs or steroids two weeks prior to or during the study, changes to DMARDs (either synthetic or biologics) in the preceding three-months or during the study, injection therapy (local anaesthesia or steroid) one-month prior or during the study, surgery within six-months prior, arthroplasty of the treated knee, any neurological signs, daily pain in the knees bilaterally, chronic bone damage or soft tissue injury, flare-up (acute inflammation) in the lower limb joints in the week prior, pregnancy or fever. Due to the considerations taken in recruitment, the study is protected for several significant risk's confounders and biases (Kendall 2003). However, there are variables not considered at this level serology, disease duration, co-morbidities, and other physical modalities in use (England *et al.* 2019; Myasoedova *et al.* 2020).<br><br>Given the above 47 participants entered the study. a sample size by ambition of the authors was small. Following randomisation by means of independent personal, ensuring concealed treatment allocation (Mouton 2006), produced two-groups. Group-one consisted of 24 participants receiving manual therapy, with 23 participants allocated to the control group receiving standard exercise. This is not supplemented by a power analysis, preventing the reader from commenting on the strength of the study (Kendall 2003; Christley 2010), the efforts to make provision for dropout (Fogg and Gross 2000), and leads to the premise that the study was underpowered (Mouton 2006; Christley 2010). However, it should be noted that the study reported a single dropout (left for personal reasons). To this point, and to the strength of the study, intention-to-treat is implemented according to the study, although the study was a pre/post design with raises questions as to how this was achieved.<br><br>To the strength of this study, the authors acknowledged the significant influence pharmacy has over the outcomes and made significant effort to minimize these potential effects by means of exclusion criteria relating to medication. Providing stability in the baseline and through the study. Furthermore, the medications monitored (NSAIDs and DMARDs (both synthetic and biologics)) were insignificantly different between groups at baseline ($p$ = 0.8972), with no changes made. The study does, however, fall short in reporting details of type, dosage and wash out periods for these and other medications not considered (i.e., antidepressants). With all this inconsideration, with praise to the authors for significant efforts to curtail influence, the reader (like the authors) cannot exclude its confounding of outcomes. |

| | |
|---|---|
| | In pragmatic terms, this study lacked a "no intervention" control group; therefore, it tests the relative effectiveness of the two interventions (manual therapy versus standardised exercise programme) to one another and cannot comment on the efficacy of any one of the interventions. This is further compounded by the lack of clarity of the standard exercise programmes, effect size and how it compares to a "no intervention" control group. It was also noted that the control group exercise programme was derived from studies looking into osteoarthritis and ankylosing spondylosis (Carbon *et al.* 1996; O'Reilly, Muir and Doherty 1999; Anwer and Alghadir 2014), which infers that the exercise programme does not focus on RA and its measurable impact is thus questioned. Other confounding factors such as, the effect of touch on outcome was not controlled for (Melzack and Wall 1965; Moayedi and Davis 2013; Senderovich *et al.* 2016). Both of these are confounders in manual therapy, particularly when outcomes are principally subjective in nature.<br><br>To the study's credit, the intervention techniques used were clearly detailed and could be replicated. However, the clinical impact of these techniques is compromised by a lack of blinding (both the participants and the treating physician), as this introduces bias (Moustgaard *et al.* 2014). The further lack of a power-analysis implies that the study is potentially under powered to detect the differences between the groups and thus detect intervention effect differences over and above placebo and other effects (Kaptchuk 1998; Bialosky *et al.* 2011). This may be why a lack of difference was noted in HAQ and OKS. Having stated this the sample did provide 20 participants to each group, which would have allowed parametric statistics to be utilised, potentially strengthening the outcomes (Hoskin 2012; Bahreini *et al.* 2020). This is however negated by the being too short, which questions the probability of attaining outcomes from exercise that typically takes six to eight weeks to reflect in normal patients (Fraser 2008; Heine *et al.* 2012), let alone in patients with RA disease, where increased chronicity and painful muscle inhibition mitigates against shorter treatment protocols being effective and may result in different outcomes if the groups are not adequately stratified for disease duration (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019). This is further complicated in that measures were recorded one day after the interventions, which is insufficient time for exercise effects to manifest (Fraser 2008). Follow-up measures would have been more appropriate given the intervention type and its physiological effect.<br><br>As stated previously, all the measurement tools used are subjective, limiting actual physiological change as opposed to perceived change in the patient. An example would be related to whether participants perceived the intervention as "novel", as this may have resulted in reported outcomes to be higher on subjective scales as opposed to objective scales. Given that the sample seems to have received specialist care at a rheumatologically centre, it is anticipated that participants were unlikely to be naive to the care, thus potentially resulting in more negative subjective outcomes (Mouton 2006).<br><br>Given the above limitations, the authors deserve credit for noting the following limitations and their potential effect - single blinding, placebo effect, pharmacotherapy effects and the lack of long-term follow-up of the study. |
| Outcome | As noted by the authors, all outcomes based on the use of intention-to-treat analysis for all groups improved significantly over time for both groups, but there was no significant difference between the groups by the end except for pain (as noted on the VAS) ($p$<0.001). This does, however, differ from the study's results noted in Table 3 where reference can be found to a difference between the groups, with $p$=0.004 representing the difference in outcome between the groups for KKS scores.<br><br>This difference may be based on the level of significance required, as the stated outcomes for the prior analysis was set at 0.05, whereas only 0.001 figures have been reported as significant. This however is not clear in the manner in which the authors report the results. |
| Discussion | Given that the study reveals a significant number of external validity limitations that impact the outcomes, it is reasonable to assume that the outcomes were not completely a result of the intervention (i.e., affected by confounding bias). Therefore, the outcomes are not only compromised in terms of what they measure but, equally, their applicability to specific population groups within RA is not determinable.<br><br>Additionally, the reporting seems inconsistent in terms of the baseline requirements for reporting (viz. $p$<0.001 or $p$<0.05). Also affecting this is the fact that the VAS was significantly different at the outset between the two groups (viz. the intervention group had a more severe VAS rating at the outset and thus a greater chance for improvement), implying that the control group had a floor effect that limited their ability to improve to the same extent as the intervention group or more. Thus, the actual outcomes in terms of VAS readings are skewed and questionable. |
| Conclusion: | Given the above limitations, the reader cannot exclude the potential effect of external influence/s on the reported outcomes, in addition to the significantly reported outcome being based on an unequal starting point. This factor is to a degree contrasted by the review of the reviewers, which ranks the study as 9/11 with 94% agreement, indicating that there is high methodological rigour and a low risk of bias from the internal validity vantage point. This highlights a disjuncture between the two reviews. It is, however, possible to have relative strength in the rigour of a study, with the bias and confounding bias induced by influences outside of the research process. As a result, this study provides limited evidence to support the tested intervention and therefore further research should be carried out in order to offer more substantial evidence for the application of the intervention in the management of RA. |

Table 4.34a: Tabulated feedback of reviewer data

| Author/s | Scholten C., Brodowicz T., Graninger W., Gardavsky I., Pils K., Pesau B., Eggl-Tyl E., Wanivenhaus A. and Zielinski C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1999 | | | | | |
| Title | Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | No | Yes | 67% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | No | Yes | Yes | 67% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 6/11 | 6/11 | 7/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.34b: Tabulated feedback of reviewer data

| Author/s | Scholten C., Brodowicz T., Graninger W., Gardavsky I., Pils K., Pesau B., Eggl-Tyl E., Wanivenhaus A. and Zielinski C. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1999 | | | | | |
| Title | Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 1 point (a) | 0 points (c) | 0 points (c) | 0 points (c) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 2 points (a+b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (a) | 0 points (e) | 1 point (a) | 1 point (a) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67% |
| | Total score | 9/10 | 5/10 | 5/10 | 6/10 | |
| | | | | | Overall percentage agreement | 80% |

Table 4.34c: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Scholten C., Brodowicz T., Graninger W., Gardavsky I., Pils K., Pesau B., Eggl-Tyl E., Wanivenhaus A. and Zielinski C. |
| Year | 1999 |
| Title | Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - HAQ<br>- FQCI<br>- BDI<br>- 21-point scale to evaluate cognitive-behavioural and environmental impact. | Study one: Baseline, week 2, 6 and 52<br>Study two: Baseline and 5 years. | Study one one-year<br>Study two five-years | Starting total 68<br>Study one: experimental (n=38), control (n=30), all completed.<br>Study two: (all 68 in one-group) Completers = 64 | No | Study-one, yes<br>Study-two, no. | Study-one, yes<br>Study-two, no. | PEDro:7/11 and NOS: 6/10, both offering arguably for moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | PEDro: 94%<br>NOS: 80% |

| | |
|---|---|
| Limitations | This study was carried out through the Department of Physiotherapy and the Department of Orthopaedics at the University of Vienna Medical School, Vienna, Austria. Thus, the study recruitment would have been from within the communities associated with these Departments and / or the general population of Vienna, although the protocol is not outlined. This affects the generalisability of the results, as it is impossible to determine how closely the study population matches the general RA population. Additionally, the exposure to specialist departments suggest that the sample is not likely to perceive treatment as novel and / or be naïve to possible RA interventions, given that they may already have been exposed to them. This latter highlights the role of patient factors that may influence outcomes (Wells *et al.* 2012). The sample consisted of 68 consecutive participants meeting the only requirement of definite RA (Arnett *et al.* 1988). Thus, the sample had 54 females with a mean age of 48.3 years (range 21 to 79); a disease duration of 8.9 years (mean) (range 0.4 to 30 years); Steinbrocker's functional classes I (n=14), II (n=38), III (n=17) (Steinbrocker, Traeger and Batterman 1949). All participants were noted to have had prior ongoing rheumatological care; but the study did not record or report on RA serology, RA disease activity, pain severity, imaging studies, medication (e.g., type, dosage, administration) or use of other physical modalities (England *et al.* 2019; Myasoedova *et al.* 2020). These factors act as outcomes confounders and may result in type II error conclusions in the study (Pocock *et al.* 2002; Bland and Altman 2011; de Boer *et al.* 2015), aside from limiting greater applicability to clinical practice (Kendall 2003). The exclusion criteria, protected against incomplete outcomes assessments and missing data (Kendall 2003). The above complicates the understanding of group homogeneity of the RCT, as intergroup comparisons are not provided. For the RCT, the 68 participants were randomised (unknown method and no reported concealment) into an intervention group (n=38) and a waiting-list control group (n=30). This suggests a possibility of researcher bias, which is not mitigated against with homogenous groups, as no evidence of baseline comparability if offered (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). For the observational study, the control received intervention and together with the intervention group formed a single uncontrolled group that was tracked from the 2nd to the 5th year. No power analysis was reported, potentially under powering the study (Kendall 2003; Mouton 2006; Christley 2010). Provision for dropouts (Fogg and Gross 2000) and intension-to-treat analysis were not required for the RCT as it did not suffer any dropouts (Heritier, Gebski and Keech 2003; Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017). The impact of the four participants dropouts during the five-year observational study are unknown. The sample size has an impact on the use of appropriate statistics; although the numbers allow for parametric statistics and drawing of causal links (Hoskin 2012; Bahreini *et al.* 2020), the under powering may hamper these conclusions. It may therefore have been the reason for the study to make use of Mann-Whitney U test, Friedman analysis, and Fisher's extract test (all non-parametric tests), limiting the strength of the analysis and reported outcomes.<br><br>The intervention was a multidisciplinary programme, alongside all participants receiving ongoing rheumatological care, education was over nine group sessions (≤ eight participants) sessions in two weeks. Education included interactive discussions on RA related topics (Schreiber and Schreiber 1998), goal setting, and problem solving and aimed at enhancing compliance. Psychology sessions assisted with patient's self-control and self-efficacy, coping skills and morbidity of RA. Exercise training sessions (supervised monthly sessions / exercise revision sessions) [based on Lorig's Arthritis Help-book (Lorig K 1988)], outside these being completed home, aimed to reinforce the initial two-weeks of training. The 10-minute exercise sessions were aimed at maintaining mobility through therapeutic exercises, protecting the joints from further destruction, and strengthening weakened muscles. Adherence/compliance was tracked and encouraged by means of an exercise diary. The intervention protocol was strengthened due to efforts encouraging compliance/adherence (e.g., exercise diaries), use to information packages and the integration of theoretical and practical learning paradigms. This improved and maintained participant compliance (Heritier, Gebski and Keech 2003). The detractor to this study was that it failed to offer a clear description of all aspects of the intervention (e.g., the exercises are not clearly defined), which prevents evaluation of the effect/s of treatment protocol (Heine *et al.* 2012) and study duration (Fraser 2008) on the reported outcomes. Thus, with compliance there is less risk for not measuring the intervention, but without detail, the outcomes are meaninglessly interpreted as the intervention is not understood. This is further compounded |

| | |
|---|---|
| | by effects of carry-over, synergistic and antagonistic relationships between different components of the programme (Robins, Hernán and Siebert 2004), patient learning, practice, and fatigue (Bordens and Abbott 2002). Finally, the lack of detail regarding the intervention prevents clinical replication for the appropriate patient type.<br><br>All outcomes were subjective in nature, raising concerns around the influence of patient factors (see Section 2.7.1) on the outcomes reported and whether they reflect/report actual physiological change as opposed to perceived change (Mouton 2006; Wells *et al.* 2012). They were all administered in a uniform and consistent manner in a set location, limiting external influence (Richardson 2007). The RCT design evaluating outcomes at baseline, post-intervention (two-weeks), at six-weeks, and at 52 weeks (one-year); while the cohort design evaluated outcomes between baseline and five years The later assessment periods in both studies provided outcomes on medium to long term effects of the intervention (Pater *et al.* 1998). The reliability and validity of the questionnaire/s in the RA context varied (Beck *et al.* 1961; Fries *et al.* 1980; Muthny 1989, 1996; Pater *et al.* 1998). With the outcome tools applicable and appropriate, this mitigates against the risk of exaggerated or under reported outcomes (Coster 2013; Bahreini *et al.* 2020), although the increased period of time between measure does open them for risk due to memory recall/decay (Mouton 2006; Ricker, Vergauwe and Cowan 2016), placebo (Kaptchuk 1998; Bialosky *et al.* 2011), learning effect (Bordens and Abbott 2002) and Hawthorn effects (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015).<br><br>Two fundamental problems in the design of the long-term study are the fact that the exercise exposure of the initial control was delayed by 52 weeks, making the exposure to exercise less in this group even with high rates of adherence over the five-year period. Secondly the long-term study did not actually have a control to compare with, so conclusions may be by chance or natural history. Given that the control group was no different to the Given that RA fluctuates, it is not clear how RA exacerbations were dealt with at the time of outcomes assessments. Given that the two groups at the five-year assessment, where not substantially different, does not talk to the groups being similar at baseline and that they achieved the same outcome, it actually poses a question on the inverse, in that no matter the exposure to exercise there is no change. With both groups were the same after four and five years respectively, the question arises as to why a patient should continue with exercise when there is no benefit beyond the four-year interface (viz., the four year and five-year exercise exposures had the same outcomes). Lastly, the study fails to compare outcomes to reported MCID values and as a result, the outcomes cannot be read in terms of possible clinical improvement and value (Bahreini *et al.* 2020). |
| Outcome | The RCT study with 68 definitive RA participants showing degree of functional dysfunction indicated that an exercise programme supplemented with education and psychological support had positive one-year outcomes. Significant improvements in disability ($p$<0.0001) were noted at week two and week 52 ($p$<0.01) and coping ($p$<0.0001) (FQCI); depression ($p$<0.01) and knowledge of the condition ($p$<0.0001) saw significant improvements at week six and week 52. Depression improved significantly ($p$<0.0001) at week two and maintained significance at week 52 ($p$<0.01). The use of exercise (relaxation exercise and remedial gymnastics) improved significantly and peaked at six-weeks ($p$<0.0001) and maintained at 52-weeks. Use of joint protection was significant at 52-weeks ($p$<0.0001), with social assistance and engagement significantly increased ($p$<0.0001). No change in control was reported. In contrast to the positive outcome's patients reported that their view of orthopaedic surgery as a sensible therapy increased from 5.3% to 26.3% ($p$<0.02). No explanation is provided, but it does imply the exercise is not the quick fix and may not provide the long-term outcomes the patients are seeking (as the overall long-term study suggests – see last paragraph in previous section). The long-term five-year observational (uncontrolled) crossover design found that the following significantly improved: disability ($p$<0.0001), coping ($p$<0.009); knowledge (disease/self-management) and social engagement maintained; and RA therapy compliance declined but still attained s significant improvement (p<0.0001). Depression worsened to non-significant levels ($p$>0.05). Thus, most of the one-year improvements were continued between one-year and five-year follow-ups for functional status and coping, while depression regressed to baseline values (Scholten *et al.* 1999). |
| Discussion | Reporting two studies in out publication does provide a conundrum when these are analysed, as it is often difficult to separate the two components and some fundamental flaws are seen throughout both studies. In these particular students, there was a lack of control over many patient / RA factors, intervention programme and statistical analysis (e.g., power analysis) that risk confounding of the outcome and their application outside of the study (viz., clinical practice). Another limitation is understanding the synergistic and antagonistic effects of each of the various components of the multipronged intervention. This may hide or elevate positive as well as negative outcomes that cannot be seen without the study of each component making translation into practice even more complex. Collectively, all the shortfalls in reported in the limitations of the methodology above significantly limit the evaluation, control and elimination bias impacting external validity. |
| Conclusion | Given the discussion above along with the concerns, there is a high level of bias in respect of the external validity of the study. By contrast the risk of bias to the internal validity as ranked by the reviewers (94% agreement PEDro for RCT and 80% agreement for NOS) indicated respective scores the 7/11 and 6/10. These indicate that a congruent outcome of moderate risk exists in the methodological rigour and thus moderate bias. With the context of a high degree of bias from the external validity and moderate degree of bias from the internal validity vantage points, indicates that the study provides limited to no evidence in support of the intervention programme. Thus, until further research is done addressing the concerns on this review in order to refute or confirm the outcomes of this study, there is no evidence for the use of a similar programme in clinical practice. It is acknowledged by the researcher that the article is either a translation of an original document or the authors authoring the text were not principally first language English speakers, as some parts of the text read ambiguously and may have negatively influenced the outcome of this review. |

Table 4.35a: Tabulated feedback of reviewer data

| Author/s | Shinde, S.B. and Varadharahulu, G. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2017 | | | | | |
| Title | Effect of Therapeutic Exercise Programme in Adults with Early Rheumatoid Arthritis | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | No | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 6/11 | 6/11 | 6/11 | 6/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.35b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Shinde, S.B. and Varadharahulu, G. |
| Year | 2017 |
| Title | The effects of strength and endurance training in patients with rheumatoid arthritis. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - Functional status index (FSI)<br>- Maximum grip strength<br>- Perceived health (Euro-Qol 5D) | Pre- /post-intervention | 12 weeks intervention<br><br>6-month follow-up | 30 participants included with two dropouts and 28 completing the six-month follow-up. | No | The control group was not a no-intervention group. It was a conventional therapy group | Yes, the modus operandi was "a simple random sampling method," but this was not explained. | 6/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study conducts recruitment through by means of the out-patient department, and through the department of physiotherapy, at the medical faculty of "Krishna Institute of Medical Sciences Deemed University" (India). The source, of recruitment takes the form of a specialist centre, which raises the question, are the sample representative of the RA population or that of a subpopulation. Limited information is provided to the reader to dismiss the plausible risk that participants available at such a source, would be seeking medical care for potential exacerbations in their condition. If this were true the sample is restricted in terms of generalisability as this would no-longer represent the RA population, but that of a subpopulation of RA (Wells *et al.* 2012). Further to this point, the sample are receiving therapy through the out-patients department, the likelihood of naivety or introduction to novel intervention is reduced (Mouton 2006; Moustgaard *et al.* 2014).<br><br>The study utilised a limited set of recruitment criteria, These, were limited to a diagnosis of RA with no criteria specified, a disease duration of more than three years, undergoing medical treatment (not specified), and an age of between 20 to 55 years. There were no noted exclusion criteria, meaning that there was no control of factors associated with RA serology, RA functional class, RA disease activity, RA flare-ups, pain, other physical modalities in use, contraindications, and comorbidities (England *et al.* 2019; Myasoedova *et al.* 2020). This broad outline of the inclusion criteria means that the homogeneity of the groups was not likely to be attained (Wells *et al.* 2012). In addition, baseline comparisons were not stated, which is important as any significant differences in the distribution of the baseline variables, particularly in the primary outcome, puts the study at risk of type II error (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). This is compounded by the seeming lack of control for confounding introduced by medication regarding the interruption and / or change of (dosage or type as well as the participants' adherence to regimen and the method of administration of the medication (Zhao *et al.* 2017; Heidari *et al.* 2019). As a result, the impact of medication on the outcomes of this study are unknown and have not been fully accounted for. Lastly, the fact that details regarding homogeneity / stratification for the dynamic RA characteristics were not reported on introduces bias and allows the clinical outcomes to be affected by the lack of homogeneity. The use of a simple random sampling method does not allow the reader the necessary information to determine how much the strategy has counteracted the lack of stringent patient inclusion criteria and a wide disease variance (Fogg and Gross 2000; Kendall 2003; Mouton 2006; de Boer *et al.* 2015; Mansournia *et al.* 2017).<br><br>The intervention groups comprised a therapeutic exercise programme group versus a conventional therapy group, making this a relative effectiveness study as no formal benchmark was set for the conventional therapy (or citation offered to the reader making this evident). Thus, the small sample and the lack of homogeneity together make it very difficult to determine the effect size of each intervention and the relationship between them. Additionally, the study is not able to determine the "active ingredient" in the outcomes or that part of the intervention programme that resulted in the outcomes. This occurs for two reasons. Firstly, there is an agonist / antagonist relationship (Robins, Hernán and Siebert 2004) between different interventions at a physiological level and, secondly, this particular study does not provide a detailed description of the intervention protocols. Both these factors impact on interpreting the clinical impact for patients generally.<br><br>The effects of the lack of homogeneity and the clarity surrounding the outcome measures of the study are further affected not only by the lack of a reported power analysis (Kendall 2003; Christley 2010), as a result of which the study is assumed to be underpowered (Mouton 2006; Christley 2010), but also, by a lack of clarity regarding drop-out compensation for dropouts and / or the use of an intention-to-treat analysis (Fogg and Gross 2000; de Boer *et al.* 2015; Unbiased Research 2017). The sample of 30 participants resulted at best in 15 participants per group. This requires the use of non-parametric statistics, offering the ability to show trends, not causal links, between intervention and outcome (Hoskin 2012; Bahreini *et al.* 2020). All these collective methodological concerns significantly weaken the rigour of the study.<br><br>The rigour of the methods is further compromised by the reliance on two subjective outcomes and one objective clinical outcome. The slightly higher weighting of the subjective measures means that the study should have controlled for effects of the three following factors: |

| | |
|---|---|
| | - touch effect, as touch will be different between the groups and its impact on pain perception (Melzack and Wall 1965; Senderovich *et al.* 2016);<br>- the placebo effect (in terms of the patients understanding whether they are in the therapeutic or conventional therapy group as well as whether they are naïve or not to the interventions (Kaptchuk 1998; Mouton 2006; Bialosky *et al.* 2011), which is unlikely given the source of recruited,<br>- the Hawthorne effect, especially in those instances where the assessor was not blinded to the intervention group (Kendall 2003; Moustgaard *et al.* 2014).<br>All these biases require the study to have employed a greater sample size to allow the clinical outcomes to be better than chance alone.<br><br>Notwithstanding these limitations, the outcome measures and the tools are valid and reliable in the context of RA (Coster 2013; Bahreini *et al.* 2020), which increases the likelihood of outcomes measuring the intervention effect in these patients, but it is uncertain whether this is likely to have overcome the noted limitations discussed so far. Moreover, the time frame between evaluation and last treatment was not stated, which raises concern as to whether the last outcomes application measured the last intervention or the whole programme (Pater *et al.* 1998); A further concern is that the assessor of the outcomes was not reported to be blinded to group allocation, which introduces bias and may have facilitated the Hawthorne effect as the practitioner attitude to the different interventions may be the overriding factor influencing clinical improvement (Kendall 2003; Moustgaard *et al.* 2014).<br><br>To the authors' credit, they do note that the patients' RA muscle strength compromise may result in gains being so low that even a small increase in strength may not result in significant improvement in function and independence and, conversely, small decreases in strength may impair function and independence. Therefore, the concerns around this study as mentioned above and the small gains apparent in the participants against this backdrop of "noise" make it almost impossible to determine the statistical, let alone the clinical, significance of either the interventions used or the differences in their outcomes. |
| Outcome | The study reported significant benefit in functional status, grip strength and personal care after 12-weeks of therapy. From baseline to follow-up the intervention group archives significant improvement in FSI (functional status: mobility/assistance, pain, and difficulty) ($p$<0.0001), Grip strength bilaterally ($p$<0.0001), and Euro-QoL 5D (QoL) ($p$<0.0001). While the control reported no significant change in any outcomes. Difference in outcomes between groups at post-study assessment, illustrated significant improvement in favour of the intervention group for FSI (functional status: pain, and difficulty) ($p$<0.0001), and Euro-QoL 5D (QoL) ($p$<0.0001), while personal care (reflected in the Assistance component of FSI) was insignificant ($p$=0.011) and Grip strength bilaterally ($p$>0.05). In conclusion, the study claims that the therapeutic exercise programme over 12-weeks proved effective for adult patients with early RA in the improving of functional status, perceived health, and grip strength (Shinde and Varadharajulu 2017). |
| Discussion | The significant limitations noted above, including the small sample size that was underpowered, compromised homogeneity at baseline, unbalanced reliance on subjective outcomes, and risk of subjective bias (touch, placebo, hawthorn) in the study suggests that the outcomes may actually be a result of chance rather than a result of the interventions. This is compounded by lack of blinding, and the time-confounding relating to outcome measurement in relation to the most recent intervention exposure, risks the capacity of the outcomes to truly reflect the interventions clinical effects. Finally, use of non-parametric statistics in a study lacking no-treatment comparative control significantly detracts from the clinical value offered by the study, as it can only report on trends in the context of relative effectiveness. Thus, between the limitations discussed and the acknowledged limitations made by the authors, the study is significant burdened by a compromised external validity. |
| Conclusion | The external validity is compromised by the limitations discussed above, and this leads the study to a high degree of bias and the reader cannot exclude the effect of an external influence/s, as can be seen in the analysis. The review offered by PEDro analysis via reviewers, ranks the study as a 6/11 with 94% agreement, this indicates that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study. However, this lack of congruency reveals to the reader that the study maintained a degree of strength in the internal rigour, and bias and confounding were introduced outside of this framework. At best the review acknowledges moderate to low evidence. However, the limitations have led to the poorly described patient, compromising the field clinician's capacity to identify best suited RA patients to this intervention type. With the lack of detail to the intervention itself the prevents replication, and the programme-based intervention prevents the reader defining therapeutic value of a single component to the programme. Thus, the review can only extract from the study restricted evidence for exercise (in this case) from this context and suggests further research be carried out before confident application in the management of RA can be achieved. |

Table 4.36a: Tabulated feedback of reviewer data

| Author/s | Strasser, B., Leeb, G., Strehblow, C., Schobersberger, W., Haber, P. and Cauza, E. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2011 | | | | | |
| Title | The effects of strength and endurance training in patients with rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | No | Yes | Yes | 67% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | No | Yes | No | No | 67% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 7/11 | 7/11 | 7/11 | |
| Overall percentage agreement | | | | | | 94% |

Table 4.36b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Strasser B., Leeb G., Strehblow C., Schobersberger W., Haber P., and Cauza E. |
| Year | 2011 |
| Title | The effects of strength and endurance training in patients with rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Echocardiogram (organ manifestations, valve morphology and ejection rate)<br>- BP (taken in 3 versions)<br>- Inflammatory markers (ESR and CRP)<br>- Medications<br>- DAS28<br>- VAS (general health and pain)<br>- HAQ-DI<br>- Max workload<br>- Heart rate<br>- Max muscle strength (1 rep max for bench press, bench pull, leg press)<br>- Anthropometry (body weight; height; fat mass; lean body mass) | Mostly pre/post<br><br>BP: before, after each training session and during last minute of ergometry<br><br>ESR: not clear<br><br>Heart rate (continuous during ergometry)<br><br>Max workload (before and after each training session – ergometry) | Six-months | Total staring = 40 Intervention and Control group each = 20.<br><br>Drop-out 5 from intervention group.<br><br>Completers = 35 (87.5%) Intervention = 15 (75%) and Control = 20 (100%) | Yes | Yes, stretching exercises and normal physical activities omitting study intervention exercises. | Yes, simple randomisation in a 1:1 ratio. | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | This study had a small sample of 40 patients that were drawn from the Rheumatology unit at Wilhelminen Hospital in Vienna. These patients ranged in age from 41 to 73 years and were required to have been diagnosed with RA (1987 ACR criteria (Arnett *et al.* 1988)) with a duration of greater than two years and to have been receiving stable drug therapy for three months preceding entry into the study. The participants were excluded if they had participated in another research study, or had experienced cardiac arrhythmias, a recent myocardial infarction, a stroke, cancer, and / or an untreated hypertension. These exclusion criteria were, however, not limited in that each patient underwent a clinical examination to screen for contraindications in respect of the intervention. Medication was controlled for in the study as all participants were asked to continue their medications for its duration. Changes in medication were only allowed in acute circumstances (reported in only two participants for the study). To the strength of the study, this well considered set of recruitment criteria, afforded a greater level of control over a range of confounding factors. Despite this, given the dynamic characteristics of RA, the wide age range, and the small sample size, it is unlikely that the two groups were completely homogenous. As it is impractical for research to control for all these variables. With this in mind the authors reported that the randomised allocation resulted in homogeneity between the groups in terms of age, gender, disease duration, medication, and physical characteristics (weight, fat%, lean mass) (de Boer *et al.* 2015).<br><br>The participants in the intervention group undertook strength training (sets of weight-bearing exercises for all major muscle groups) and performed systematic endurance-training on a cycle ergometer twice a week by comparison with the participants in the control group who only participated in stretching-exercises twice a week without additional resistance and were free to continue their recreational physical activity with the exception of strength-training and systematic endurance-training. Although these regimes were well described their complexity does detract from the ability to replicate this study in clinical practice (Heine *et al.* 2012). To add to this, the control represents a therapy arm and not a natural history comparative group. This unfortunately detracts from the study, as the control selected limits finding to relative effectiveness. Without clear benchmarking of the control efficacy of the programme cannot be determined.<br><br>Although there it is suggested that the patients were naïve to the interventions via the exclusion of those with previous research participation, yet this is not expressed plainly (Mouton 2006), furthermore, it is impossible to blind participants in terms of their getting a treatment programme as usual versus an unusual treatment programme because blinding in exercise and manual therapy studies is very limited (Moustgaard *et al.* 2014). This bias on the side of the patient may have been present, but the rheumatologist who assessed the clinical status of the patients and the research physician who assessed the outcomes were both indeed blinded according to the |

| | |
|---|---|
| | information provided in the study (Kendall 2003; Moustgaard *et al.* 2014). Although the lack of blinding of the patients seems to have detracted from the rigour the blinding of the assessors adds a greater rigour to the study.<br><br>With regard to the statistical analysis, no comment was made with regard to determining the sample size by means of the power analysis calculations and therefore it is not possible to tell if the study was underpowered (Kendall 2003; Mouton 2006; Christley 2010) or if it had the potential to suffer from type II error in its conclusions. This is increasingly possible when muscle strength in RA patients has negligible gains even over a six-month period (most muscle responses occur between six and eight weeks) (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019), which makes it even more difficult to detect small changes in a small sample size. In comparison other reported studies that required in excess of 20 patients per group in order to achieve the 80% power at a confidence level of 95 or 99%, it is likely that the outcomes of this study may be compromised. This is further compounded by the lack of intention-to-treat analysis reporting so, notwithstanding the five dropouts from the intervention group, it is left to the reader to assume that all participants completed the study without the need for intention-to-treat analysis computed. Further to this point, the study fails to inform the reader of the baseline comparability of the dropout against the completers. Given the small sample and the dropout suffered, the use of non-parametric statistics would have been required, thus limiting the study to the observation of trends without the ability to draw firm conclusions (Hoskin 2012; Bahreini *et al.* 2020). Given the weaker statistical outcomes, it is good to note that the compliance to training was high, but this may have been affected by the relatively individualised programmes within the allowed range negatively by outliers within a smaller sample group. The statistical analysis and thus the outcomes of the study are further limited to the treatment programme and not a single intervention, even though a recent systematic review indicated that high-intensity exercise may bring about more benefits (Stenström and Minor 2003). This may have been better measured with more considered timing of outcome measurements, but as these are not articulated clearly, it is not possible to determine its relevance or impact (Pater *et al.* 1998).<br><br>It must be acknowledged that the authors did allude to many of the limitations of the study as noted above, as well as presenting strategies for future studies to consider in order to improve the generalisability of the outcomes and to gain a better idea of the clinical impact of the intervention (as related to the MCID changes for the outcome measures (Bahreini *et al.* 2020)). In addition, a recommendation was made to include specific RA subgroups such as disabled patients and patients with active RA so as to allow for a better understanding of the clinical impact of these programmes on all types of RA patients. |
| Outcome | The study reflected in the outcomes several significant benefits of the programmed interventions. The intervention group showed trends toward significant improvement in disease activity (by DAS outcome) at six-months (p=0.06) but does not actually archive statistical significance. However, inflammatory markers (CRP and ESR) do not reflect significant change. The intervention group reflected significant change from baseline to six-months in pain reduction (p=0.05) and improvement in general health (p=0.04), while functional capacity (HAQ-DI) showed 16% improvement but failed to reach significance (p=0.06). With regard to strength outcomes, maximum strength and workload were insignificantly different between groups. But the intervention group did reach significant improvement from baseline to six-months in workload by 10% (p=0.00) and strength (leg press of 22% (p=0.01); bench press of 11% (p=0.06); bench pull of 9% (p=0.03)). The control reported no significant change. Body composition outcomes at six-months reflected weight loss for the intervention group (p=0.01) but not the control; fat % dropped in the intervention group (p=0.02) and control (p=0.03); lean mass increased in the control (p=0.01), but intervention group reported no significant change. These body composition outcome improvements were insignificantly different between groups. The study, therefore, concluded combination exercise therapy for six-months (strength and endurance) imparts benefit to the RA patient in terms of muscle strength, endurance, body compensation and functional ability. However, longevity of these findings is not known (Strasser *et al.* 2011). |
| Discussion | The lack of control over the above-mentioned confounding variables, along with the very small sample size, significant dropout in the intervention group, lack of blinding of the therapists, dubious naivety of participants, underpowered study, suggests that the outcomes may actually be a result of chance rather than a result of the interventions. Compounding factors were the with-in limits the prescription of an individualised care, dubious naivety of the participants, and the use of nonparametric statistics, with comparison to another intervention group, limits the studies capacity to draw definite conclusions. Theses and the limitations discussed above had significantly limits the generalizable of the outcomes. |
| Conclusion | Given the above analysis, from an external validity vantage point, this study suffers from a moderate degree of bias, and thus prevents the exclusion of effects of external influence/s. This is in line with the review by the reviewers, which rank the study by PEDro scale as 7/11 with 94% agreement, indicating that there is moderate methodological rigour and moderate risk of bias to the internal validity of the study. Therefore, this study at the best offers moderate evidence and required a degree of caution when used as a basis for management of RA. However, the field practitioner is restricted by the studies patient description (albeit fairly comprehensive) lacking in some RA dynamic variables. Thus, should be kept in mind when considering RA patients that would best benefit from the study intervention/s. As a final point, the study is one of relative effectiveness and not efficacy. This and the above-mentioned limitations require attention in future research efforts, in order to strengthen the evidence for RA therapy. |

Table 4.37a: Tabulated feedback of reviewer data

| Author/s | van den Ende, C., Breedveld, F., Le Cessie, S., Dijkmans, B., De Mug, A. and Hazes, J. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2000 | | | | | |
| Title | Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | No | No | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| | Total score | 7/11 | 9/11 | 8/11 | 8/11 | |
| | | | | | Overall percentage agreement | 94% |

Table 4.37b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | van den Ende CH, Breedveld FC, le Cessie S, Dijkmans BA, de Mug AW, Hazes JM. |
| Year | 2000 |
| Title | Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Swollen joints (primary)<br>ESR (inflammatory marker)<br>Pain VAS<br>Disease Activity Score (DAS)<br>Health Assessment Questionnaire (HAQ)<br>EPM-ROM (Joint mobility)<br>Isokinetic extension 60°/s<br>Isometric extension<br>50-foot walk test | Baseline and weeks -three, -six, -12, and -24. | 24-weeks | 64 consecutive admissions; experimental group (n=34) and control group (n=30).<br><br>Dropout suffered by week-24 (experimental group lost six, control group 10). With completers (outside of intention-to-treat) = 75% | Yes | Yes | Yes | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 94% |

| | |
|---|---|
| Limitations | The study utilises a convenience sampling method in that the recruited patients for this study are those of the 64 consecutively admitted eligible patients to a rheumatology clinic requiring multidisciplinary treatment because of active RA with a loss of functional ability. Thus, imparting risk of subpopulation of RA recruited into the study, limiting the generalisability of the study findings and influencing the likelihood of naivety and novelty in the sample (Wells *et al.* 2012; Jager, Putnick and Bornstein 2017).To be included the patients were required to conform to the 1987 American Rheumatism Association RA criteria (Arnett *et al.* 1988), to manifest six or more swollen joints (disease activity), to have had at least two of three (morning stiffness (> 45 minutes) / tender joint count (> nine) / ESR (>28 mm/1st h)), and to be between 20 and 80 years of age with the ability to walk 15 m indoors. The exclusion criteria covered arthroplasties of the knee joints and contraindications training such as serious cardiac or lung disease, previous or current surgical intervention and certain co-morbidities. Given that the above features are all important, a number of dynamic RA characteristics were not considered, nor were patient demographics, functional classification of RA, or variations in the medication prescribed per patient (England *et al.* 2019; Myasoedova *et al.* 2020).<br><br>Regarding the inclusion criteria, it is noted that patients were stratified by sex and randomly allocated by an independent person to the intensive or the conservative exercise programme by means of sealed, opaque envelopes (Mouton 2006). These two are in favour of the study improve its rigour (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019). The methodology was strengthened by an *a priori* analysis with the sample size set at 32 patients per group (Kendall 2003; Christley 2010) while allocation of numbers allowed for dropouts (Fogg and Gross 2000) and the utilisation of the intention-to-treat analysis (de Boer *et al.* 2015; Unbiased Research 2017). The post-study calculation also revealed that the sample size of 32 was sufficient to detect a 7% difference (0.5-unit DAS score) between the two groups on the DAS score and a 20% difference in swollen joint count at a 0.05 level of significance and a power of 90%. On computation of the baseline differences, none were found between the groups for sex, age, disease duration, serology (RF specifically), erosions, ESR, swollen joint count, HAQ, medication or for clinical measures with the exception of pain scored on a VAS. The significant difference in the VAS between groups ($p$ = 0.03) introduces a concern with regard to a confounding of the results as there are implied limitations to exercise and some measurement outcomes (Pocock *et al.* 2002).<br><br>The participation required hospital admission. In this context patients in both groups received primary nursing care and fixed rest hours per day. In addition, their supervised prone and seated activities included the ROM (active assisted smaller joints) and isometric muscle strengthening exercises for larger joints with strict progression requirements accommodating patient limitations. These were completed on alternate days four times a week; however, this was not clearly reported. Additionally, there was a weekly group session. Patients were informed about joint protection and provided with joint splints (if necessary); also, psychosocial problems related to RA were addressed. Optimised medical treatment consisted of NSAIDs, DMARDs, and intra-articular injections with corticosteroids if necessary. In addition to the above, the experimental group received a supervised intensive exercise regimen. This included working on the isokinetic strengthening of the knee extensors and flexors with training done on an isokinetic dynamometer (EnKnee, Enraf Nonius, Delft, The Netherlands), developing the shoulder girdle muscles against manual resistance and performing other muscle strengthening exercises such as bicycling on a home trainer. In general, the protocols were fairly well defined, particularly in the case of the experimental |

| | |
|---|---|
| | group, thus allowing reproducibility (Heine *et al.* 2012). However, there is lack of specificity in terms of the joints targeted in both groups as well as inconsistency in the administration of the exercises in the baseline care (i.e., four times a week as opposed to daily, both of which are reported in the study), while the multiplicity of treatment options almost make it possible for each patient to have had a different set of joints treated in a different manner (Kendall 2003; Heine *et al.* 2012). The factors make this a more a case series than an RCT. Furthermore, the study protocol did not change the patient's routine medication and any changes were recorded at each assessment. Although this may allow statistical control for the effect of medication, the complexity of a multidisciplinary treatment programme and the effects of synergistic or antagonistic effects (Robins, Hernán and Siebert 2004) can limit the ability of the study to draw firm conclusions regarding the treatment interventions, especially if this is not fully contextualised both in the statistical chance realm and also the realm of clinical significance through the MCID measures (Bahreini *et al.* 2020).<br><br>To the study's credit the outcome measures were completed by a single group-allocation-blinded observer. Evaluations were performed at baseline and at three, six, 12, and 24 weeks after admission. This frequency allowed for computation of intention-to-treat analysis for missing data and the multiple time points allowed for a more accurate prediction of patient outcomes in this model (de Boer *et al.* 2015; Unbiased Research 2017). This is further enhanced by the valid and reliable outcome measures which included disease activity (DAS) (Fuchs and Pincus 1994; Prevoo *et al.* 1995), ESR, VAS for pain, grip strength and isokinetic strength, HAQ (Fries *et al.* 1980; Siegert *et al.* 1984) and the patient's global assessment. Additionally, joint mobility was assessed by Escola Paulista de Medicina (EPM)-ROM scale (Ferraz *et al.* 1990; Vliet Vlieland *et al.* 1993), and walking 50-meter test. On the other hand, of the outcomes used there were limitations evident in this study regarding the grip and isokinetic strength in terms of reported pain. In these instances, data was computed as the lowest possible value resulting in a floor / ceiling effect for improvement in these measures after the resolution of the pain (Liang 2000; Coster 2013). Dropouts were recorded at 24 weeks and reasons recorded. There were only slight, statistically insignificant differences in demographic and clinical variables at baseline between the patients who completed the trial and the patients who were regarded as dropout; it does however show bias in that the loss would result in a regression to the mean. Additionally, the increased loss in the control group suggests that the perception of normal treatment may have influenced retention. |
| Outcomes | The implemented exercise programmes of varying intensity and exercise prescription in which the 100 participants comprising the study's sample took part over 12 weeks claim to have improved the aerobic capacity of a subset of participants by 17% with increases in joint mobility and muscle strength of 16% and 17% respectively resulting from the HIE group exercise programmes. These findings differed significantly from those of the other exercise groups. Despite the improvements measured at 12 weeks post-intervention by these gains had mostly disappeared by the 24-week follow-up. The HIE group reported increased $VO_2$ readings and muscle strength at 12 weeks, both from baseline and by comparison with the other exercise groups ($p$<0.001) and HE groups ($p$=0.02) respectively. The $VO_2$ improvement was not maintained at 24 weeks, yet muscle strength at 24-weeks was still significantly improved from baseline though but lower. HIE joint mobility at 12 weeks, though lower, was significantly improved from baseline (regarding palmar and dorsal flexion of the wrist, hip flexion, knee extension and plantar flexion of the ankle) by comparison with all other groups ($p$<0.001). Significance in improvement was not maintained at 24-weeks. The HIE functional capacity walk test reported significant improvement from baseline at 12 weeks (failure to report *p values)* and had maintained significance at 24 weeks (failure to report *p values)*. HIE joint count (swollen) had reduced significantly at 12 weeks from baseline and when compared with LIE-gr ($p$<0.01). HIE pain (VAS) had reported significantly increased pain at 24 weeks from baseline values, while the HIE group reported significantly increased pain at both 12 and 24 weeks from baseline values (failure to report *p values)*. Finally, the LIE-gr reported significant improvement in muscle strength at 24 weeks when compared with baseline values (failure to report *p values).* Furthermore, the study claims to have observed no exacerbation in disease activity during the study. In conclusion, the study found that HIE was more effective in improving aerobic capacity, physical condition, muscle strength and joint mobility than the other forms of exercise over a 12-week programme. However, effects of the exercise programme were mostly lost by 12weeks post-intervention (Van den Ende *et al.* 1996) |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, with a sample size allowing the calculation and reporting of parametric statistically defined causal links. However, the limitations noted above suggest that outcomes may have been affected by factors outside of the intervention. These reported limitations collectively affect the study negatively and hinder its capacity to both accurately and fairly compare the intervention groups without the risk of bias in the outcomes and conclusions. Additionally, despite the fair methodology reporting, aspects of the methodology lacking which influence the reader/s ability to analyse control and elimination of confounding and bias variables are lacking. The incomplete participant description limits the generalisability and implementation of the intervention programmes in clinical practice and the findings of the study are then less reliable, providing limited clinical guidance in terms of dynamic and weight bearing exercise programmes, although evidence is reported for the relative effectiveness of the intervention. |
| Conclusion: | Given the external validity of the programme as analysed through the limitations, this study shows a moderate possibility for bias. The internal validity review by the reviewers (with 97% agreement) indicates that the study presented moderate methodological rigour. These two evaluations place the study in moderate to low risk of bias. Thus, the outcomes show that there is moderate evidence for use of the interventions in clinical practice. |

Table 4.38a: Tabulated feedback of reviewer data

| Author/s | van den Ende, C., Hazes, J., le Cessie, S., Mulder, W. J., Belfor, D. G., Breedveld, F. C. and Dijkmans, B. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1996 | | | | | |
| Title | Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomised clinical trial. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | Yes | No | No | 67% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 8/11 | 7/11 | 7/11 | |
| Overall percentage agreement | | | | | | 97% |

Table 4.38b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | van den Ende, C., Hazes, J., le Cessie, S., Mulder, W. J., Belfor, D. G., Breedveld, F. C. and Dijkmans, B. |
| Year | 1996 |
| Title | Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomised clinical trial. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - Aerobic capacity ($V<V>O_{2max}$)<br>- Muscle strength (knee extensor and flexor)<br>- Joint mobility (EPM-ROM score)<br>Functional ability:<br>- HAQ<br>- Dutch-AIMS subscale<br>- 50 feet walk test<br>- Stair's test<br>Disease activity:<br>- Swollen joints count<br>- Ritchie articular index<br>- ESR, CRP and Platelet count<br>- VAS (pain, morning stiffness, tiredness, and disease activity) | Baseline, post intervention (week 12), and follow-up (week 24) | 12-weeks | Total sample of 100 assigned into four groups of 25.<br><br>Dropouts (n=10, 10%): three in the HIE, five in the LIE-gr, two in the LIE-ind.<br><br>Completers: HIE (n=22), LIE-gr (n=20), LIE-ind (n=23), HE (n=25). | No | No, the comparison group was a home exercise group with instructions for isometric and ROM exercises.<br><br>They claim this group as a control, but it is an intervention group. | Yes | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 97% |

| | |
|---|---|
| Limitations | The sample was recruited to a Rheumatology Outpatient Clinic at the Medical School Hospital of Leiden, Netherlands. Given that this facility of an outpatient facility, the impression provided is that the patients were recruited from the general population and not limited to more severe cases as seen in specialist centers (Wells *et al.* 2012). It is reported that 260 consecutive eligible participants were approached and screened against the following inclusion criteria, All participants were required to be diagnosed with RA, according to the 1987 ACR criteria (Arnett *et al.* 1988), be between 20 and 70 years, on stable medication for the three-months prior to the study and be able to use a home trainer bicycle. Participants were excluded if at recruitment they experienced high disease activity (according to their rheumatologist), had to change medication to slow acting antirheumatic drugs, were unable to tolerate the intervention due to co-existing cardiac or respiratory disease and / or they had weight bearing joint replacements. Even with the above criteria, the study did not consider the following variables to eliminate bias and confounding effects (Kendall 2003). These include, the number of other interventions over a specific time (Mouton 2006), the type and number of associated comorbidities (Kłodziński and Wisłowska 2018; England *et al.* 2019; Levitsky *et al.* 2019; Myasoedova *et al.* 2020), the type and number of associated concomitant diagnoses (Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011; Dagenais and Haldeman 2011), the type of contraindications to intervention (Durcan, Wilson and Cunnane 2014), the patient endocrine status (Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018) and other general factors (such as ethnicity and RA functional class) (Bartley and Fillingim 2013).<br><br>Of the 260 eligible participants 160 declined invitations. These withdrawals potentially limit the representation of the total RA population (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013), however the authors suggest that participant and non-participant populations did not differ significantly. As a result, the remaining 100 RA patients who were on stable medication, consisted of 63 females (37 males), a mean age of 52 and mean disease duration of 10 years reflected the sample adequately. The sample was randomised (method unknown) into four equal groups. The lack of description prevents analysis of randomisation and concealment methods, introducing risk of bias (Mouton 2006). The groups represented: high-intensity (HIE), low-intensity group-based (LIE-gr), low-intensity individual-based (LIE-ind), and home exercise (HE). These groups were insignificantly different in terms of age, sex, and RA disease duration; therefore, they are relatively comparable for the parameters controlled by the inclusion criteria, allowing the intervention to be tested above confounding variables (de Boer *et al.* 2015). The study controlled for allowed changes with routine medical therapy. However, with the claim that medication was "stable" throughout the study it is unclear to what extent the medication changes resulted in confounding or bias. This may be negligible as dropouts did not report changes in second line drug use over the 24-month study period. (Zhao *et al.* 2017; Heidari *et al.* 2019). |

| | |
|---|---|
| | The interventions designed for the four groups include variable exercise programmes carried out over 12-weeks at a set time and place (controlling for confounding effects of time and environment of the intervention), after which an additional 12-week follow-up period occurred. The duration of the intervention allows the study to provide the necessary expose to exercise causing measurable physiological change (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). The study offers a limited but clear definition of the programme, which limits the assessment of real difference between the groups (Heine *et al.* 2012). The unblinded therapists did not allow for controlling of their influence on intervention and outcome (Moustgaard *et al.* 2014). In addition to the variability of the programmes for the patients, introduces concerns regarding variable patient expectation, satisfaction, treatment novelty and naivety, patient blinding and how these impact the adherence to the exercise programme (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015; Ricker, Vergauwe and Cowan 2016). The balanced outcome measures used (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018); along with non-significant difference between the groups regarding rheumatoid factor (seropositive/negative), erosive status, swollen joint count, Ritchie index, HAQ, ESR, assists the study in being able to more directly measure the outcomes as a result of the interventions. The single unblinded assessor provided a degree of control over assessment variability; however, it does introduce of assessor bias (Kendall 2003; Moustgaard *et al.* 2014). Most of the outcome tools had citations indicating that there were valid and had reliability in the measurement of the respective outcomes, but few were validated for the RA population (Ferraz *et al.* 1990; Vliet Vlieland *et al.* 1993; Coster 2013; Bahreini *et al.* 2020). Another confounder is that although outcomes measures were detailed (protocol and control for variability) (Kendall 2003) details lack regarding time-modified confounding (Platt, Schisterman and Cole 2009) such as the diurnal effect (Nandgude, Hasabe and Kolsure 2018), testing environment (Richardson 2007), and the relationship between last intervention and assessment (Pater *et al.* 1998). The study made use of a power analysis for detection of a 20% improvement with regard to $VO_{2max}$ between high and low intensity programmes ($p$=0.05 and power of 90%). Given this, 22 participants per group were required (Kendall 2003; Christley 2010). With the considered dropouts some analyses could be completed with parametric (powered appropriately) and others with non-parametric (underpowered) statistics. To attain parametric statistics for the majority of variables, intention-to-treat analysis was employed (Heritier, Gebski and Keech 2003; Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017). This analysis presents risk to outcome reporting (positive or negative) (Fogg and Gross 2000; Kendall 2003; de Boer *et al.* 2015). Thus, for the primary outcome the study was underpowered (Mouton 2006; Christley 2010) and several outcome measures fell victim to the floor / ceiling effects (Liang 2000; Coster 2013). The incomplete evaluations were reported to display parity between the groups, and it was suggested that this had no effect on the outcomes. However, it did result in variable statistical analyses an only some measure could report causal links (Hoskin 2012; Bahreini *et al.* 2020). An additional consideration is that dropout fell out of the study because their RA got worse, but then it also needs to consider the intervention may have made the patients worse in terms of the RA presentation. Therefore, both options need to be adequately considered. Finally, the study does not compare outcomes to published MCID values to allow for understanding clinical impact (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020). |
| Outcome | The implemented exercise programmes of varying intensity and exercise prescription, in 100 participants over 12 weeks, claims to have improved the aerobic capacity of a sub-set of participants by 17%, increased joint mobility and muscle strength by 16% and 17% respectively as a result of HIE group exercise programmes. These findings when compared with the other exercise groups differed significantly. Despite the improvements, after 12 weeks post-intervention, these gains had mostly disappeared by the 24-week follow-up. The HIE group reported increased $VO_2$ readings and muscle strength at 12 weeks, both from baseline and compared to all other exercise groups ($p$<0.001) and HE groups ($p$=0.02) respectively. The $VO_2$ improvement was not maintained at 24 weeks, but the muscle at 24-weeks were still significantly improved from baseline but lower. HIE joint mobility at 12 weeks was significantly improved (in particular; palmar and dorsal flexion of the wrist, hip flexion, knee extension, and plantar flexion of the ankle) from baseline and compared with all other groups ($p$<0.001). Significance in improvement was not maintained at 24-weeks. HIE functional capacity walk test reported significant improvement both from baseline at both 12 weeks (failure to report *p values)* and maintained significance at 24 weeks (failure to report *p values)*. HIE joint count (swollen) had reduced significantly at 12 weeks from baseline and when compared with LIE-gr ($p$<0.01). HIE pain (VAS) had reported significantly increased pain at 24 weeks from baseline values, while the HE group reported significantly increased pain at both 12 and 24 weeks from baseline values (failure to report *p values)*. Finally, the LIE-gr reported significant improvement in muscle strength at 24 weeks when compared with baseline values (failure to report *p values).* Furthermore, the study claims to have observed no exacerbation in disease activity during the study. In conclusion, the study found that HIE was more effective in improving aerobic capacity, physical condition, muscle strength and joint mobility, than the other forms of exercise over a 12-week programme. These effects of the exercise programme were mostly lost after 12-weeks post-intervention (Van den Ende *et al.* 1996). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, with a sample size allowing the calculation and reporting of parametric statistically defined causal links. However, the limitations noted above, suggests that the outcomes may have been affected by factors outside of the intervention. These reported limitations collectively affect the study negatively and hinders its capacity to both accurately and fairly compare these intervention groups without risking bias in the outcomes and conclusions. Additionally, despite the fair methodology reporting, there are aspects of the methodology lacking, which influence the reader/s ability to analyse control and elimination of confounding and bias variables. The incomplete participant description limits the generalisability and implementation into clinical practice, and the findings of the study are then less reliable, providing limited clinical guidance in terms of dynamic and weight bearing exercise programs with evidence reported for the relatively effectiveness of the intervention. |
| Conclusion: | Given the external validity of the programme as analysed through the limitations, this study shows moderate possibility for bias. The internal validity review by the reviewers (97% agreement), indicates that the study presented with moderate methodological rigour. These two evaluations place the study in moderate to low risk of bias. Thus, the outcomes show that there is moderate evidence for use of the interventions in clinical practice. |

Table 4.39a: Tabulated feedback of reviewer data

| Author/s | Vljeland, T. V., Zwinderman, A., Vandenbroucke, J., Breedveld, F. and Hazes, J. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1996 | | | | | |
| Title | A randomized clinical trial of in-patient multidisciplinary treatment versus routine out-patient care in active rheumatoid arthritis. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | No | Yes | Yes | Yes | 67% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | No | No | No | No | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | Yes | Yes | Yes | 67% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | No | Yes | Yes | Yes | 67% |
| Total score | | 5/11 | 8/11 | 8/11 | 8/11 | |
| Overall percentage agreement | | | | | | 91% |

Table 4.39b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Vljeland, T. V., Zwinderman, A., Vandenbroucke, J., Breedveld, F. and Hazes, J. |
| Year | 1996 |
| Title | A randomized clinical trial of in-patient multidisciplinary treatment versus routine out-patient care in active rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - VAS (disease severity, pain and fatigue, morning stiffness (duration and severity)<br>- Physician's disease activity estimates<br>- Swollen joints count<br>- Ritchie Articular Index (RAI)<br>- Inflammatory markers (ESR and CRP)<br>- HAQ<br>- Grip strength<br>- AIMS<br>- Kellgren criteria (radiographs)<br>- Health care costs (records and interview) | Baseline, weeks two, four, 12 and 52. | 52-weeks | Total sample: 108<br><br>Randomised participants: 80 (group 1 n=39 and group 2 n=41)<br><br>Non-randomised group (n=28) | No | Yes, standard care | Yes, with exception to the "non-randomised" participants, supplementing the study. | 8/11, indicating a high methodological rigour and subject to low risk of bias within the internal validity of the study. | 91% |

| | |
|---|---|
| Limitations | A significant limitation of this study is that although it was available electronically, meeting the review criteria there were missing data pages. Attempts to contact the corresponding author were unsuccessful. This analysis is therefore based on the data and information available in those portions of the publication that were available to the researcher. The study conducted its recruitment through the Leiden University Hospital, rheumatology department (Netherlands). The department is a central referral point for the Leiden and surrounding areas. All participants in the study attended the clinic on an out-patient bases. Given the recruitment method, the reader is led to believe the study is at an increased risk of recruiting a sample that is prone to suffering a greater degree of severity in disease activity. As recruiting visiting outpatients comes with it the inference that they are seeking medical attention for some form of exacerbation. This would unintentionally limit the generalisability of the study. Furthermore, such a sample is unlikely to be naïve to or consider the intervention novel (Kendall 2003; Wells *et al.* 2012; Smith *et al.* 2013).<br><br>All patients were required to have active RA requiring medical assistance or a change of DMARDs. This limited the study in terms of the severity of the presenting RA and therefore its generalisation to the RA population generically. Additionally, the RA patients were required to fulfil the 1987 ARA criteria for definite RA (Arnett *et al.* 1988) and to be between the ages of 18 – 75 years with three or more swollen joints and two or more of the following: modified RAI (Ritchie et al., 1968) (>or = 9), morning stiffness (>45 mins) and (3) ESR (>28mm/h) (no citation offered). Although these inclusion criteria seem clear they allowed for a broad range of RA patient participation and thus were limited by the exclusion criteria which included previous hospitalisation for multidisciplinary treatment or medical need for immediate hospitalisation, ACR functional class I or IV (Hochberg *et al.* 1992) and the presence of major disability or severe joint damage requiring surgical correction. The study did not consider control for the RA serology the possible comorbid conditions that would influence outcomes or the patients' prescribed medication in the light of these criteria or the patient stratification, which may bias the comparison of the groups and thus not allow for effective treatment outcomes to be measured. This is compounded by the fact that all patients had DMARDs introduced or changed shortly after study entry or during the study. Also, NSAIDs were optimized during the study with intra-articular injections of corticosteroids administered if needed. This negates the influence of medication as the primary effector in the attained outcomes.<br><br>This study involved 80 active RA patients randomised into one of two groups (random cards in sealed envelopes with sex-based stratification). Additionally, there also seemed to be a group created for patients who objected to random allocation. The two randomised groups received either 11 days in-patient multidisciplinary treatment followed by standard out-patient care (n=39) or standard out-patient care only (n=41) and the non-randomised group (n=28). All clinical data for these groups were collected according to the study protocol. The inclusion of the non-randomised group, where the patient in essence got to choose the interventions, implies that this group was not free of bias resulting from naivety (Mouton 2006), the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015) or predefined ideas of the possible outcome of the study (i.e., a prior belief that they did not require some part of the intervention within each of the groups) (Mouton 2006; Moustgaard *et al.* 2014), any of which would have undoubtably influenced the outcomes attained clinically and thus the measurement of the interventions utilised. Additionally, the non-randomised group would also have had differences in their presenting baseline measures which the authors reported as being less active disease and a better functional and emotional status, the latter indicating better coping strategies among these patients compared to the other two groups. The randomisation was supported using an *a priori* analysis (36 patients per group). This means that the study met the requisite power and |

| | |
|---|---|
| | reduced the level of error in accepting outcomes that could be flawed (Kendall 2003; Mouton 2006; Christley 2010). The ability to attain the 36 measures per outcome was supported by an intention-to-treat analysis in those cases where data was missing and required imputation (Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017). Although plausible and acceptable, this can lead to over- or underestimation of the effect size and therefore result in regression to the mean, limiting the ability of the study to detect statistically significant change (Fogg and Gross 2000; Heritier, Gebski and Keech 2003; Kendall 2003; Gupta 2011; de Boer *et al.* 2015). The ability to reach statistical significance, although positive, does not imply that there is a clinically significant change; the study does not report the MCID for the various outcomes and therefore does not contextualise the outcomes clinically in the RA population (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020). In terms of the interventions, although they are spelt out clearly in terms of the different components of the programme, it is unclear what the exact exercise prescription was other than a programme of individual ROM and strengthening exercise as applied by a therapist (Heine *et al.* 2012). A unique feature in this study was that treatment goals were discussed at weekly multidisciplinary team meetings. This implies that there may have been a pragmatic approach to patient care and allow for adjustment to the intervention to meet treatment targets. This may take the form of progression to the exercise programme component and adjustment of NSAID and steroid use to optimise pharmaceutical management over time, but this is unclear and would not be repeatable in a future study. Alternatively, the lack of detailing a specific therapeutic structure, and the purpose of the multidisciplinary meetings could mean that specific goals were set between the therapists and the patient, and these are driving the therapeutic intervention. This mechanism would unduly influence the outcomes as patients would be less willing to report a lack of progress given that a specific outcome has been set in advance (Mouton 2006; Richardson 2007; Heine *et al.* 2012; Sedgwick and Greenwood 2015). Either way it is unclear as to how this process would have affected the measurement of the interventions. The reader can only say that the process would have biased the outcome of the interventions, thus limiting the ability of the study to measure the intervention. It is particularly important to consider this when the out-patient care group had similar treatments permitted within the groups, intervention framework, but whether to include and use them was left to the treating physician's discretion (as per normal practice), who did not specifically provide goals and outcomes which were clearly stated for the in-hospital group. In the light of the above, therefore, this study is at very best a pragmatic evaluation of normal in-hospital versus normal out-hospital care and does not appropriately confirm to the requirements of an RCT whose outcome is to measure the explicit effect of an intervention (Moher *et al.* 2012; Wells *et al.* 2012). This also speaks to the possibility that there is an unseen bias or expectation that an in-patient setting may have additional therapeutic value, which further increases bias from a patient reporting vantage point. To control for inter-assessor variability, the outcomes were all conducted by the same physician who was not involved in the management of the patients but was not blinded to the randomization status (Kendall 2003; Moustgaard *et al.* 2014). The use of a single assessor works in the study's favour, but the lack of blinding does imply that assessor bias was possible based on the perceptions held by the assessor in terms of the anticipated outcomes of the two programmes. This may have been countered by the broad range of outcome measures, but a fairly high reliance on subjective outcomes does not completely negate the possibility of bias (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). |
| Outcome | The study outcomes reflected improvements in several variables within the in-patient group when change from baseline was compared with the out-patient group. Pain, fatigue, the number of swollen joints and joint tenderness (RAI) significantly higher in the inpatient group versus the out-patient group at weeks two ($p$<0.05) and four ($p$<0.05). Additionally, the outcomes reflected more significant improvement in the in-patient group than out-patients regarding morning stiffness at weeks two, four, and twelve ($p$<0.05); grip strength at week two ($p$<0.05); anxiety at week twelve and fifty-two ($p$<0.05); depression at week twelve ($p$<0.05); and subjects reported disease activity for all assessments ($p$<0.05). According to the ACR criteria for clinical improvement, the in-patient group reflected a significantly better response to assigned intervention at all assessments when compared with the out-patient group ($p$<0.05). All other outcomes did not reflect any significance. The in-patient and non-allocated group comparison reported the non-allocated group response intervention imparted significantly less benefit in pain, RAI, HAQ, and depression at week two ($p$<0.05), and anxiety at weeks -two and -four ($p$<0.05). The data comparing non-allocated and out-patient groups reflected improvement in swollen joints count at weeks -two and -four, and CRP levels at week-twelve in favour of the non-allocated group ($p$<0.05). Finally, non-allocated patients made use of services (both medical and paramedical) less frequently than the other groups. The authors conclude that this short-term multi-disciplinary intervention in the in-patient setting, imparted beneficial outcomes in the active RA patient for disease activity parameters and emotional status (Vljeland *et al.* 1996). |
| Discussion | The lack of control for confounding variables in the study noted above suggests that the outcomes are at risk of reflecting change due to chance rather than to the intervention/s. Compounding factors introduced via the non-allocated group included the lack of intervention prescriptive details, the nature of reporting of outcomes risking regression to the mean, unblinded assessor and participants with heavy reliance on subjective outcomes, which are particularly vulnerable to bias introduced by lack of blinding. The study, therefore, was significantly restricted with regard to generalisability and consequently the recommendations stemming from these research efforts to address the issues raised by the study are needed to enhance the significance and application of the study's findings to improve clinical practice. |
| Conclusion: | Given the significant limitations noted above from the vantage point of external validity, this study suffers from a moderate to high risk of bias; the effect of the presence of external influence/s cannot be excluded. By contrast, the review of the study done by the external reviewers according to the PEDro scale ranks the study as 8/11 with 91% agreement, which indicates high methodological rigour and a low risk of bias to the internal validity of the study. This disjuncture between reviews is possible as strength in the rigour of the study may not reflect bias and confounding in this study is introduced by factors outside of the research process. Therefore, this study at best offers moderate to low evidence for the intervention/s. Until further research can be carried out to address the limitations noted above, this evidence should offer a baseline from which other evidence can indicate specific intervention component effectiveness or efficacy in RA management. |

Table 4.40a: Tabulated feedback of reviewer data

| Author/s | Westby, M. D., Wade, J. P., Rangno, K. K. and Berkowitz, J. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2000 | | | | | |
| Title | A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | No | No | No | No | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | No | No | No | No | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 7/11 | 7/11 | 7/11 | 7/11 | |
| Overall percentage agreement | | | | | | 100% |

Table 4.40b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Westby, M. D., Wade, J. P., Rangno, K. K. and Berkowitz, J. |
| Year | 2000 |
| Title | A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| Active joint count, ESR, HAQ, Fitness score, Caltrac, Bone mineral density (BMD) (hip and lumbar spine). | Pre/post | 12-months | 53 total sample: Usual care group (n=16), intervention group (n=14), comparison group (n=23).<br><br>Dropout of 15 (28%): Six at baseline from usual care group, three at baseline from intervention group, and six from comparison group.<br><br>Total completed: 38 | Yes | Yes, | Yes, methodology not stated | 7/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 100% |

| | |
|---|---|
| Limitations | The study reports recruiting participants from Rheumatology Clinics affiliated with the Arthritis Society, the Department of Medicine at the University of British Columbia and from outpatients of the Physical Therapy Department of the Mary Pack Arthritis Centre (Canada). The impression is that, given the source from which participants are drawn, there is an increased likelihood that the patients do not conformed to the general RA population and may indeed be more severe in terms of their presentation (Wells *et al.* 2012). As a result, the study unintentionally risks limiting the generalisability of the study. In such a scenario, the outcomes reflect external influence rather than intention.<br><br>The inclusion criteria for this study included: women with a RA diagnosis (ARA criteria / ACR (Arnett *et al.* 1988)); disease duration (> one year); ACR functional status (class I or II) (Hochberg *et al.* 1992); medication (continuous low dose prednisone) and not currently exercising. Exclusion criteria included known osteoporosis related fractures, significant cardiovascular disease, planned lower extremity surgery or joint replacement surgery within six months and medication including previous high dose prednisone use (the equivalent of 40 mg/day for a month or more) and methotrexate use within the past six months). These criteria, however, excluded RA serology, RA disease activity, patient age, pain rating and other interventions as well as comorbidities and contraindications to care (England *et al.* 2019; Myasoedova *et al.* 2020). Given what was noted, the baseline characteristics and distribution in this study showed significant differences among the three groups. These included cardiovascular and pulmonary disease ($p$=0.01), osteoporotic risk factors ($p$=0.03) and greater accumulative doses of corticosteroids ($p$=0.02)). With regard to other noted risk factors such as smoking, alcohol, calcium intake and parity the groups showed no significant differences. Such significant differences as occurred put the study risk of type II error due to the lack of homogeneity of the groups, which had the potential to bias the actual effects of the intervention being tested (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). In the RA context, medication requires very specific control (Zhao *et al.* 2017; Heidari *et al.* 2019). In this study, outside of the prednisone dose (a continuous low dose of 2.7 – 7.5 mg/day for six or fewer months), the other DMARDs were permitted and recorded for the duration of the study. In addition, all participants received 1000 mg of calcium carbonate and a multivitamin with 400IU of vitamin D daily, recorded at baseline, at six months and at 12 months. This rigour nevertheless failed to consider NSAID use and the use of Biologics. Additionally, the methods by which the medications were administered, the pharmacokinetic effects of the medications and their therapeutic concentration on outcome measures were also not considered specific to the RA patient. These parameters, in addition to the paucity of inclusion criteria, without stratification, decreased group homogeneity. Thus, the study was negatively impacted by several biases and its ability to measure the effects of the intervention/s was severely curtailed (Zhao *et al.* 2017; Heidari *et al.* 2019).<br><br>Another confounding factor and possibly a reason for the high dropout rate is that of the recruitment source and the previously bias associated with this. Together with the minimum 12 months disease duration, this reduced the likelihood of patients being exposed to novel treatments (Mouton 2006). This lack of novelty and the possibility that the patients had previously received a similar or identical intervention would have resulted in their perceptions being brought into the clinical interaction that allowed subjective outcomes to reflect perception instead of clinical reality. These concerns align with the authors' limitations which suggest that the attrition of patients may have been due to study duration, concerns over exercise aggravating symptoms, and the selection criteria.<br><br>In conflict with the inclusion criteria, the study groups all received advice to "continue with regular physical activities and therapy (physical and occupational) as needed". Additionally, there is a lack of a true "no intervention" group in this study, which prevents direct benchmarking of the outcomes against a placebo or known entity group. As a result, this study is really a relative effectiveness study of three groups with three different interventions on what is in part a similar population (i.e., one group did not take steroids). On a positive note, |

| | |
|---|---|
| | however, if the difference in medication use between the groups is overlooked the study used a form of unspecified randomization of participants in the usual care and aerobic training group. These groups consisted of 16 participants taking low does steroids, randomised into the usual care group, an intervention group, populated by 14 taking low does steroid randomised participants. And a separately recruited comparison group, populated by 23 participants meeting the inclusion/exclusion criteria, however they were not using steroids.<br><br>To the study's credit *a priori* power analysis was done (a sample size of 30 was required per group). As acknowledged by the authors, however, this was not achieved which thus decreases the study's capacity to detect statistically and clinically significant differences in outcome measures (thus, underpowered (Mouton 2006; Christley 2010)), which would have required the use of non-parametric statistics which can only infer trends and not causality (Hoskin 2012; Bahreini *et al.* 2020). Additionally, the underpowered study is also at risk of accepting type II errors. For example, BMD changes in the younger comparison group may have been due to their perimenopausal status and the increased risk of accelerated bone loss, which was in part statistically countered by the use of an intention-to-treat analysis (Gupta 2011). This controlled the dropout effect resulting from 28% attrition but caused a regression to the mean of the outcomes, increasing the likelihood of not detecting statistically and clinically significant differences in outcome measures. However, (and agreeing with the authors' limitation) because no regression analysis was conducted factors such as disease activity, disease duration and functional status, among others, may have confounded the results. Exercise adherence among the groups and a lack of accounting for this also affected the study's ability to discern improvement or lack thereof from the intervention.<br><br>In terms of outcomes measures, the study used a set of blinded assessors (one per category of outcomes), which controlled for bias from the assessor (Kendall 2003; Moustgaard *et al.* 2014). The study was further strengthened by the outcome tools being a mix of subjective and objective measures (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). The tools (Active joint count, ESR, HAQ, Fitness score questionnaire, Caltrac and DPX Bone mineral density) were mostly reliable and valid (Active joint count (Klinkhoff *et al.* 1988), ESR (no reference cited), DPX (no reference cited), HAQ (Fries *et al.* 1980; McDowell and Newell 1996), Fitness score questionnaire (Francis and Francis 1988) and suitable tools (Active joint count (Klinkhoff *et al.* 1988), ESR (no reference cited), DPX (no reference cited), HAQ (Hochberg *et al.* 1992), Fitness score questionnaire (Francis and Francis 1988), with the exception of the Caltrac, which was validated for healthy individuals (Montoye *et al.* 1983).<br><br>The baseline values attained for some of the outcome measures (HAQ being one of them) do not provide the ability for statistical or clinical significance to be attained as they are limited by the floor or the ceiling effect (Liang 2000; Coster 2013). This is perhaps compounded in that the evaluation environment is not detailed in this study to determine consistency of application between assessors or in relation to calibration measures (Richardson 2007; Kimberlin and Winterstein 2008). This may also have been impacted by the synergistic or antagonistic effects (Robins, Hernán and Siebert 2004) of the interventions within the stated programmes. Evaluation occurred at baseline, at six months (for checking disease activity with regard to active joint count, medication status and activity levels and for providing a reminder of the calcium and multivitamin supplements) and 12 months. |
| Outcome | The authors concluded that participants undertaking a dynamic weight bearing programme showed a non-significant trend for decreasing disease activity, while imparting significant statistical improvement of function and level of activity (both $p$=0.05). The BMD outcome reflected a significant reduction in hip BMD of the comparison group ($p$=0.004), while the usual care and intervention groups remained unchanged. BMD of the spine between steroid groups (intervention versus usual care) reflected an insignificant difference ($p$=0.09). As a result, it indicates that female RA patients on low dose steroid therapy can benefit in terms of physical function/activity, with fitness levels improving from a comprehensive management programme that includes dynamic, weight-bearing exercise, while not impacting on BDM nor exacerbating RA (Westby *et al.* 2000). |
| Discussion | The lack of control for several of the limitations outlined, the small sample size, the significant dropout suffered, the underpowered study, and the use of non-parametric statistics all raise the risk of type II error. In addition, the significant baseline differences, and lack of homogeneity between groups and the failure to control adequately for medication variables further enhance the risk of type II error. The study suffered floor/ceiling effect/s, thus imparting inaccurate outcome measures. Lacking no-intervention control limits, the study in presenting the intervention's relative effectiveness without a reference to efficacy, limits the clinical value of the findings. If all these are taken into consideration it cannot be excluded that the outcomes may be a result of chance rather than a result of the interventions. The limitations discussed above, together with the authors' acknowledgement of the study's significant limitations mean that the outcomes of this study are not generalizable to the RA population. |
| Conclusion: | Given the above analysis, from the external validity vantage point this study suffers from a moderate to high degree of bias and the reader cannot exclude the effect of an external influence/s. This is congruent with the review by the reviewers, which ranks the study as an 7/11 with 100% agreement, indicating that there is moderate methodological rigour and a moderate risk of bias to the internal validity of the study. Therefore, this study at best offers moderate to low evidence of the intervention and until further research can be carried out addressing all the above limitations, the reader cannot use the study's evidence as a basis to implement this intervention in the management of RA. |

Table 4.41a: Tabulated feedback of reviewer data

| Author/s | Williams, M. A., Williamson, E. M., Heine, P. J., Nichols, V., Glover, M. J., Dritsaki, M., Adams, J., Dosanjh, S., Underwood, M. and Rahman, A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2015 | | | | | |
| Title | Strengthening And stretching for Rheumatoid Arthritis of the Hand (SARAH). A randomised controlled trial and economic evaluation. | | | | | |
| Criterion | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| 1 | Eligibility criteria were specified | Yes | Yes | Yes | Yes | 100% |
| 2 | Subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | Yes | Yes | Yes | Yes | 100% |
| 3 | Allocation was concealed | Yes | Yes | Yes | Yes | 100% |
| 4 | The groups were similar at baseline regarding the most important prognostic indicators | Yes | Yes | Yes | Yes | 100% |
| 5 | There was blinding of all subjects | No | No | No | No | 100% |
| 6 | There was blinding of all therapists who administered the therapy | No | No | No | No | 100% |
| 7 | There was blinding of all assessors who measured at least on key outcome | Yes | Yes | Yes | Yes | 100% |
| 8 | Measures of at least on key outcome were obtained from more than 85% of the subjects initially allocated to groups | Yes | Yes | Yes | Yes | 100% |
| 9 | All subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by “intention-to-treat” | Yes | Yes | Yes | Yes | 100% |
| 10 | The results of between-group statistical comparisons are reported for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| 11 | The study provides both point measures and measures of variability for at least one key outcome | Yes | Yes | Yes | Yes | 100% |
| Total score | | 9/11 | 9/11 | 9/11 | 9/11 | |
| Overall percentage agreement | | | | | | 100% |

Table 4.41b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Williams, M. A., Williamson, E. M., Heine, P. J., Nichols, V., Glover, M. J., Dritsaki, M., Adams, J., Dosanjh, S., Underwood, M. and Rahman, A. |
| Year | 2015 |
| Title | Strengthening And stretching for Rheumatoid Arthritis of the Hand (SARAH). A randomised controlled trial and economic evaluation. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| Primary outcome<br>The Michigan Hand Outcome Questionnaire (MHQ) overall hand function subscale score<br><br>Secondary outcome<br>The full MHQ, pain, health-related quality of life (Short Form questionnaire-12 items), impairment (grip strength, dexterity and range of motion) and self-efficacy. European Quality of Life - 5 Dimensions (EQoL-5D),<br><br>Medication and healthcare | Baseline, and at follow-up was at four and 12 months | 12-months (12-week intervention and follow-up at months four and 12. | Total: 490 patients, usual care (n=244), exercise programme (n=246).<br><br>Completed: usual care 93%, exercise programme 75%.<br><br>Outcomes were obtained for 89% of participants at 12 months (222 for usual care, 216 for exercise programme). | Yes | Yes, the control received usual care. | Yes, by means of a central telephone randomisation service at Warwick Clinical Trials Unit, University of Warwick. | 9/11, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 100% |

| | |
|---|---|
| Limitations | The study followed a multi-centred approach, with 17 UK NHS hospitals and centres located across the UK participating. Recruitment was through clinic staff approaching patients during their normal clinic visits and providing them information of the study. A supplementary recruitment was through postal invitations to patients listed for rheumatology consult within the NHS. If patients agreed, appointments were made for screening and informed consent was completed. This sourcing provided the study greater representation of the general RA population, when compared to other articles in this review (Wells *et al.* 2012). The study however, reported gaining broader external validity in terms of geographic and size of recruitment. Additionally, the groups where comparable at baseline in terms of participants source of recruitment, and according to the authors did not carry any effect on the study outcomes. It therefore appears that recruitment did not influence the outcomes (Williams *et al.* 2015).<br><br>Following a well-defined set of recruitment criteria, participants were limited to RA meeting the ACR clinical and immunological criteria (Newsome 2002). In addition, patients needed to report pain and dysfunction of hand and / or wrist, comply with medication, be >18 years of age, report no surgery or fractures in six months prior to study and no planned surgery during study and pregnancy. These inclusion criteria did not include several parameters related to hand dominance, the duration of the RA, the serological components of RA and or the clinical presentation, which may represent confounders to the outcome of the study. The recruitment and inclusion identified 1606 participants, of these 1042 were eligible, 32% were willing (n=512). At initial screening, 22 were excluded (reasons recorded), leaving a sample of 490. These participants underwent baseline analysis prior to randomisation. The demographic, clinical assessment findings and primary outcomes and several of the excluded factors are provided control over and with this, protecting the homogeneity and offering clarity to the type of patient involved in the study.<br><br>A strength of this study was the consort diagram provided (Begg *et al.* 1996; Kendall 2003; Moher *et al.* 2012), and indicates that there was blinding for the assessors (but not the patients or the therapists) (Kendall 2003; Moustgaard *et al.* 2014). Randomisation was carried out by third party and concealed (Mouton 2006). Stratification for location/sites in which the studies were completed was utilised to control for confounding factors of the sites (as is stated by the authors, minimising staff anticipation of and alteration to allocation of participants) (Torgerson and Sibbald 1998; Fogg and Gross 2000; Kendall 2003; Nair 2019).<br><br>By means of a power analysis prior to the study (directed at the primary outcome), the study was able to calculate and ensure that the numbers of patients in the study were sufficient to attain statistical power and thus allow for outcomes to be better than chance alone (Kendall 2003; Christley 2010). In addition, the study controlled for |

| | |
|---|---|
| | dropout of participants by incorporating provision into the recruited sample size (Fogg and Gross 2000), and carried out dropout effect analysis reporting limited effect on the measured outcomes implying that the outcomes provided a direct measure of the intervention. This was paired with a comprehensive and well described baseline comparison covering demographics and baseline clinical measures as well as a post hoc baseline analysis, which showed that there was no difference between the groups as a result of any differences in the baseline measures, and in the comparison of dropouts versus those that completed the study. This further reinforced that the outcomes are more accurate measures of the intervention, as does the statistical analysis of the effects of adherence to treatment effects as well as the analysis of the outcome measures' sensitivity (Fogg and Gross 2000; Kendall 2003; Coster 2013; de Boer *et al.* 2015; Bahreini *et al.* 2020).<br><br>The well-defined treatment was delivered by a group of trained therapists (Heine *et al.* 2012). It was carried out over a sufficient period to allow for the intervention to manifest effects on outcomes (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012). Specialisation consistency, quality control and a support system were provided for the duration of the study in order to ensure that the therapist treatment protocols were adhered to in a similar manner across all patients (Kendall 2003). Therapists completed therapy logs providing data on each participant including data on progression. The intervention group received six therapist visits (the first was an assessment and education session, with the other five being exercise session). The participants also carried out prescribed home exercise programme daily, this was facilitated by information sheets and education (several forms, among them arthritis handbook, joint protection advice) with provision of joint splinting if deemed necessary. The intervention group received supplementary exercise therapy, in depth detail was offered regarding load and volume progression and documented in exercise diaries. The control was issued usual care treatment involving one to three sessions of therapy, information sheets and education (several forms, among them arthritis handbook, joint protection advice) with provision of joint splinting if deemed necessary. The measures of outcomes the outcome assessors were blind to group assignment and independent of treatment delivery. Additionally, the analysis was controlled for assessors identifying the possibility of group allocation; blinding analysis was conducted and found not to be better than chance and equal in both groups. Lastly, all outcomes were compared to the MCID for the respective outcome measures in order to reflect clinical significance of outcomes (Bahreini *et al.* 2020). |
| Outcome | The study randomised 490 patients (244 to control; 246 to intervention). Compliance within allocated group intervention was 93% of control and 75% of intervention. Outcomes were obtained for 89% of participants at 12 months (222 for control, 216 for intervention). There was a statistically significant difference in favour of the intervention for the primary outcome at 4 and 12 months [mean difference of 4.6 points, 95% confidence interval (CI) 2.2 to 7.0 points; and a mean difference of 4.4 points, 95% CI 1.6 to 7.1 points respectively]. There were no significant differences in pain scores or adverse events. The estimated difference in mean quality-adjusted life-years (QALYs) accrued over 12 months and was 0.01 greater (95% CI –0.03 to 0.05) in the intervention group.<br><br>The qualitative study found the intervention to be acceptable to participants and highlighted the importance of the therapist in enabling patients to establish a routine and incorporate the exercises into their lives. Thus, the authors concluded that a strengthening and stretching programme for RA of the hand is clinically effective when compared with usual care alone, with no adverse effects being associated with the intervention (Williams *et al.* 2015). |
| Discussion | The study, for the most part, acknowledged and attempted t to control limitations. It offered significant control over and analytical evidence for the effects of confounding and bias factors in the study. This strengthens the external validity as the study offers evidence against the outcomes being influenced by external factors. The one specific limitation is that of patient description which lends the field clinician limited aid in identifying potential patients to benefit. The Intervention programme offered a degree of variability in the design, which imparts an element of individualised care, which would reflect a pragmatic clinical setting. However, structure was maintained over the tailoring of the intervention by means of a protocol and implemented progression requirements to maintain physiological loading required in muscle strength and growth. The limitations the study could not control were blinding of the therapist and patient; given the paradigm of the research, this is particularly difficult if not impossible to instil in such a study. However, this does not negate the effect this may have on the outcomes, especially the subjective outcomes used in this study (e.g., the primary outcome). The study uses a comparative intervention group as opposed to a natural history group (or stable medication-only group, which is probably as close to natural history research in this area as can be ethically achieved). This results in outcomes reflecting relative effectiveness but limits the overall clinical value of the findings. Finally, the primary outcome fell short of clinical difference (even though statistical significance was reached), based on MCID values determined for post-surgical cohorts. Thus, it is plausible that the MCID values are not applicable in the conservative therapy paradigm (Bahreini *et al.* 2020) ; therefore, the results of this study need to be applied with caution in clinical practice. |
| Conclusion | Given the above analysis, from the external validity vantage point this study suffers from a low degree of bias as the authors' present evidence that the outcomes of the study are unlikely to be by chance/natural history/external influence/s. This is consistent with the review by the reviewers, which ranks the study as a 9/11 with 100% agreement, indicating that there is high methodological rigour and a low risk of bias to the internal validity of the study, in terms of the PEDro scale. Therefore, this study offers the reader high-level evidence of the interventions' relative effectiveness and until further research can be carried out to determine the efficacy of the interventions, the findings could be applied to the management of RA, though with some caution. |

### 4.4.2.2 nRCTs

Table 4.42: List of table numbers for nRCT feedback and analysis

| Tabulated feedback of reviewer data | Properties of the study, outcome, and discussion | Author/s: | Year of publication: | Title of research paper: |
|---|---|---|---|---|
| Table 4.43a | Table 4.43b | Brorsson, Hilliges, Sollerman and Nilsdotter | 2009 | A six-week hand exercise programme improves strength and hand function in patients with rheumatoid arthritis. |
| Table 4.44a | Table 4.44b | Field, Diego, Delgado, Garcia, and Funk | 2013 | Rheumatoid arthritis in upper limbs benefits from moderate pressure massage therapy. |
| Table 4.45a | Table 4.45b | Gizińska, Rutkowski, Romanowski, Lewandowski and Straburzyńska-Lupa | 2015 | Effects of whole-body cryotherapy in comparison with other physical modalities used with kinesitherapy in rheumatoid arthritis. |
| Table 4.46a | Table 4.46b | Häkkinen, Hannonen, Nyman, Lyyski and Häkkinen | 2003 | Effects of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: comparison with healthy subjects. |
| Table 4.47a | Table 4.47b | Hawkes, Care, Dixon, Bird and Wright | 1986 | A comparison of three different physiotherapy treatments for rheumatoid arthritis of the hands. |
| Table 4.48a | Table 4.48b | Hurkmans, Van den Berg, Ronday, Peeters, Le Cessie, and Vliet Vlieland | 2010 | Maintenance of physical activity after Internet-based physical activity interventions in patients with rheumatoid arthritis. |
| Table 4.49a | Table 4.49b | Karatepe, Günaydin, Türkmenand Kaya | 2011 | Effects of home-based exercise program on the functional status and the quality of life in patients with rheumatoid arthritis: 1-year follow-up study. |
| Table 4.50a | Table 4.50b | Law, Saynor, Gabbitas, Jones, Kraus, Breslin, Maddison and Thom | 2015 | The effects of aerobic and resistance exercise on markers of large joint health in stable rheumatoid arthritis patients: a pilot study. |
| Table 4.51a | Table 4.51b | Lowman | 1958 | Rehabilitation of the rheumatoid cripple: a five-year study. |
| Table 4.52a | Table 4.52b | Noreau, Martineau, Roy and Belzile | 1995 | Effects of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. |
| Table 4.53a | Table 4.53b | Oh and Seo | 2003 | Decreasing pain and depression in a health promotion program for people with rheumatoid arthritis. |
| Table 4.54a | Table 4.54b | Romanowski, Straburzyńska-Lupa, Romanowska and Lorenc | 2014 | AB1125 A Comparison of the effects of kinesiotherapy and post-isometric relaxation on knee pain in patients with rheumatoid arthritis. |
| Table 4.55a | Table 4.55b | Rønningen and Kjeken | 2008 | Effect of an intensive hand exercise programme in patients with rheumatoid arthritis. |
| Table 4.34b | Table 4.34c | Scholten, Brodowicz, Graninger, Gardavsky, Pils, Pesau, Eggl-Tyl, Wanivenhaus and Zielinski | 1999 | Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. |
| Table 4.56a | Table 4.56b | Spiegel, Spiegel, Ward, Paulus, Leake and Kane | 1986 | Rehabilitation for rheumatoid arthritis patients. A controlled trial. |
| Table 4.57a | Table 4.57b | Zernicke, Kedor, Müller, Burmester, Reißhauer and Feist | 2016 | A prospective pilot study to evaluate an animated home-based physical exercise program as a treatment option for patients with rheumatoid arthritis. |

Table 4.43a: Tabulated feedback of reviewer data

| Author/s | Brorsson, S., Hilliges, M., Sollerman, C. and Nilsdotter, A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2009 | | | | | |
| Title | A six-week hand exercise programme improves strength and hand function in patients with rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 0 points (b) | 0 points (b) | 1 point (a) | 0 points (b) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (c) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 2 points (a+b) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (c) | 0 points (c) | 0 points (d) | 0 points (c) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Total score | 2/10 | 6/10 | 6/10 | 5/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.43b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Brorsson, S., Hilliges, M., Sollerman, C. and Nilsdotter, A. |
| Year | 2009 |
| Title | A six-week hand exercise programme improves strength and hand function in patients with rheumatoid arthritis |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - EX-it<br>- Grippit<br>- VAS pain<br>- Grip Ability test (GAT)<br>- DASH<br>- SF-36<br>- muscle cross section thickness<br>- muscle thickness<br>- muscle pennation angle<br>- muscle contractile pattern<br>- muscle contractile time | Baseline (two occasions), week- six, -12 and -18. | 18 weeks | Total sample of 40: 20 RA patients in group and 20 healthy controls in group.<br><br>4 dropouts (2 per group)<br>Completers 18 per group = 36 total. | No | Yes, a healthy control. | No | 5/10, indicating a moderate to low methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | With the patient sample taken from the Spenshult Rheumatic Hospital outpatients' clinic, the perception from the outset exists that the sample consists of patients that have previously sought specialist care, implying greater severity of RA disease activity than the average RA patient (Wells *et al.* 2012). No clarity was provided on where the healthy population was drawn from. Additionally, there is also a limitation to generalisability in that the patients were further limited to only female patients.<br><br>Inclusion criteria required that the RA patients comply with the 1987 ACR criteria (Arnett *et al.* 1988), have an RA duration of greater than one year and with the capacity to fully extend fingers and no exclusion criteria for the RA patients were listed. The healthy controls were excluded if they presented with inflammatory or muscle diseases or previous hand and arm injury. There were no specific inclusion criteria (which may have limited the differences between the groups with the exception of the RA disease). Further, the lack of specific dynamic RA disease characteristics and the lack of control for medication and concomitant treatments does not clearly allow the extrapolation of data to the general RA patient population in practice and the lack of power analysis limits the ability to draw firm conclusions in an underpowered study. These concerns are touched on by the authors in addition to their comment that the small sample, despite the comparability to other studies with similar sample sizes (O'Brien *et al.* 2006), does put the study at risk of bias. Even with this risk of bias, the RA disease duration ranged from two to 40 years (age range of 33 to 70) compared to the healthy 36- to 73-year-old group. In terms of this comparison the groups seemed equal where all other baseline measures were different in keeping with the RA versus the healthy controls.<br><br>The intervention lasted 12 weeks (Flatt 1974), with participants completing the explicit exercises five times a week (sessions separated by at least one day), with variable resistance and unclear bilateral or unilateral application. This description creates doubt, given that five exercises sessions cannot fit in a week, given the rest period in addition to the variability in resistance for each patient. On the limitation of bias, the participants used exercise diaries to record their exercises and adherence, in addition to the study being of sufficient length to actually detect the change due to exercise (Fraser 2008).<br><br>All measures were standardised to help control for bias and confounding, with the exception that the grip ability test (GAT) seems to only have validity in normal patients and the limited protocol for the use of ultrasound as a measurement tool. Therefore, the standardisation as well as the mix of objective and subjective outcome tools provides depth and range to the study data and allows for higher quality external validity to the study, notwithstanding the limitations of two outcome measures. However, it must be noted that the assessor was not blinded to the group allocation and introduces risk of bias of skewed data by the assessor or influence over the participant such as the Hawthorn effect or Placebo effect (Kendall 2003; Moustgaard *et al.* 2014).<br><br>From an outcome/s vantage point, the CSA of the EDC muscle showed a significant difference between the two groups before the period of hand exercise ($p<0.05$), but after six weeks no significant difference was observed between the groups ($p=0.4$), thus it was concluded that there was a greater response in the RA group assumed to be related to the intensity of the training programme and the therapeutic putty resistance. |

| | |
|---|---|
| | There was no citation of minimal clinical important difference (MCID), the results of the given study cannot be read in terms of the clinical importance (Fogg and Gross 2000; Rothwell 2006; Bahreini *et al.* 2020). |
| Outcome | The study data reflected that finger force (thus strength) had increased significantly from baseline in the RA group for extension (of the fingers) at six weeks ($p$<0.01) and continued to improve at 12 weeks ($p$<0.01); for the control for extension (of the fingers) at six weeks ($p$<0.05) and continued to improve at 12 weeks ($p$<0.01). For flexion at six weeks ($p$<0.05) and continued to improve at 12 weeks ($p$<0.01); for the control for extension at six weeks ($p$<0.05) and continued to improve at 12 weeks ($p$<0.01). Regarding grip strength the Gat outcomes reflected both groups had significantly improved at six and 12weeks ($p$<0.01) while reflecting no significant increase in pain ($p$>0.05). According to DASH there was reported improvement in hand function in the RA group at 12weeks ($p$<0.05), but this was not seen in the control. CSA of the extensor digitorum communis at baseline was significantly different between groups ($p$<0.05), and through the study the outcomes reflected significant change at six weeks for the RA group ($p$<0.01), and at 12weeks for both groups ($p$<0.01). By the six-week assessment the significant difference between groups seen at baseline was no longer reported as significant ($p$=0.40). The MCT significantly improved in both groups at six weeks ($p$<0.01) and 12 weeks ($p$<0.05). The shape of the muscle produced change at six weeks for the RA group ($p$<0.01) and control ($p$<0.01,) but ($p$<0.01) at 12 weeks for the RA group and for control ($p$=0.05). The muscle thickness of the control only had reflected significant improvement at both six weeks ($p$<0.05) and 12weeks ($p$<0.01). Pennation angle and SF-36 produced no significant findings. In short, the study provided evidence that a hand exercise programme of six weeks duration saw significant improvement hand function and force for rheumatoid arthritis patients. |
| Discussion | To conclude, the authors indicate that the study reflected significant and favourable outcomes for the RA hand function and strength. These findings were limited both by the outcomes of the fairly congruent (75% agreement) review of the article by reviewers from the vantage point of internal validity and methodological rigour, reflecting the moderate to limited strength of the methodology with high risk of bias to the internal validity of the study. Compounded by areas of significant weakness in the external validity in the context of this review (focused on RA), the study is at high risk of the outcomes reflecting more than just the effects of the intervention, given that the design prevents the exclusion of natural history / chance imparting influence on the outcomes measured. The study findings are reported from a small sample size and lacking the context offered by a power analysis. The sample cannot be defined as sufficiently powered; following the dropout suffered, the findings reported above are assumed with fairly good reasoning to be underpowered and drawn from non-parametric statistics. The lack of clarity regarding several methodologically important factors imparts significant risk of bias, confounding and error in outcome measures. The study overlooks utilising a sufficient set of recruitment criteria to offer a degree of protection and / or control over several RA specific variables and thus reducing the risk of missing data and type II error (an example of this would be the exclusion of hand deformities likely to fall short of some outcome measure, viz., the floor effect (Liang 2000; Coster 2013)). Regarding the comparison group, the study makes use of a healthy control that significantly differs in some baseline variables, so the groups offer little regarding an RA population beyond the context of improving towards a healthy norm, and even from that vantage point the study introduces bias through the demographic differences. The RA sample is poorly described, offering a wide range of disease duration, an all-female sample and only one participant capable of full hand extension (through inference implying a restriction on hand deformity and dysfunction). The reader is left to assume very niche yet poorly described generalisable capacity, in clinical practice. The outcomes are assessed by an unblinded assessor (despite the giving merit of for a larger part standardised protocol) with validated outcome measures (omitting GAT) in the RA population. The reader cannot exclude the risk of bias introduced through this. The intervention lacks the needed clarity derived solely from the publication thus preventing not only clear analysis of the intervention but also preventing replication in clinical setting and future studies. Finally, the lack of MICD values in the study prevents clear evaluation for clinical purposes. It is necessary for the study to be re-done with specific attention to the above-mentioned limitations, to provide data that can provide a bases for clinical and prognostic application. |
| Conclusion | In conclusion, when the evaluation of the internal validity and methodological rigour of this study is considered, the fairly congruent (75% agreement) review of the article by reviewers reflects the moderate to limited strength of the methodology with moderate risk of bias to the internal validity of the study. Taken into consideration with the limitations notes in this table, the external validity is at moderate to high risk of reporting outcomes imparted by factors other than intervention. This is evident and further exacerbated by the weakness highlighted in the discussion above, which results in, the study's providing at best moderate to low evidence for the use of strengthening exercise for the RA hand but limited to no evidence for the application of this study's programme in the management of RA of the hand or in clinical practice.<br><br>Finally, as noted in the discussion above, the need for re-testing of the intervention with greater consideration for the following: a wider population representation, a clear set of recruitment criteria, the provision of a patient description adequate for generalisability for clinical purposes, the use of a natural history control, the attainment of a sample size determined by power analysis with prevision for drop-out and for following the RCT paradigm with a clearly defined intervention programme. The evidence derived could then provide a base for evaluating the true effect of the intervention and its prognostic value, both needed for confident implementation into RA management. |

Table 4.44a: Tabulated feedback of reviewer data

| Author/s | Field, T., Diego, M., Delgado, J., Garcia, D. and Funk, C. G | | | | | |
|---|---|---|---|---|---|---|
| Year | 2013 | | | | | |
| Title | Rheumatoid arthritis in upper limbs benefits from moderate pressure massage therapy. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (c) | 0 points (b) | 0 points (c) | 0 points (c) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (e) | 0 points (e) | 0 points (e) | 0 points (e) | 100% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 0 points (b) | 0 points (b) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 1 point (a) | 0 points (b) | 1 point (a) | 1 point (a) | 67% |
| | Total score | 4/10 | 3/10 | 3/10 | 3/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.44b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Field, T., Diego, M., Delgado, J., Garcia, D. and Funk, C. G |
| Year | 2013 |
| Title | Rheumatoid arthritis in upper limbs benefits from moderate pressure massage therapy. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Self-administered:<br>- VITAS (pain)<br>- Perceived grip strength<br>- STAI<br>- POMS<br>- The sleep disturbance scale<br>Objective tools:<br>- Grip strength<br>- ROM | For all the measures, before and after the massage sessions on the first and last day of the four-week study. | Four-week period | Recruited 42 participant. 2 withdraw and total participants 40 (20 per group).<br><br>Dropout (20%) 8 (4 per group), thus 32 completers (80%) (16 per group). | The self-administrated outcomes where not blinded. While physical (objective) measures had no description on how it was measured so it is assumed that it was no blinded. | No control, two intervention groups (light and moderate massage). | Yes, the participants were randomised into one of two intervention groups, however the mechanism employed is not described. | 3/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | The article fails to explicitly specify the location of the study. However, the recruitment was achieved by means of email advertising through the medical school. The principal researcher is associated with the Touch Research Institute at the Medical school of the University of Miami (USA). The recruitment strategy leads the reader to assume the participants of a medical school data base are more likely to seek advanced / specialist care. Following this assumption, it would imply that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). This would inherently limit the studies generalisability to only this population group.<br><br>Recruitment criteria appear to be limited to adults with diagnosed RA (no criteria reported) that had affected their upper extremity/s (joints of the wrist, elbow, and shoulder). The inclusion criteria are too vague, and this fails to control for any RA dynamic characteristics variability (McInnes and Schett 2011; Yang *et al.* 2017). Additionally, handedness / dominance, deformity, comorbidities, medications, disease duration (chronicity) and activity, among others (described in Section 4.4.1.1). This unclear recruitment and description introduce/s bias into the study prohibiting reliable generalisability of outcomes and can confound / bias analysis, by increasing type II error reporting in relation to the effectiveness of the intervention. At baseline, the study reports demographic description as a full sample including sex distribution, age, socioeconomic status, and ethnicity. They only state that these variables did not differ between groups, implying group homogeneity for these limited demographics.<br><br>The sample was based on a power analysis (Kendall 2003; Christley 2010) and loosely based on Field *et al.* (2011) study. No study however provides the details of the analysis. Given this, if 20 participants is taken as the threshold for the use of parametric less than this would require the use of non-parametric statistics; the sample is likely to have been underpowered after drop outs are accounted for (Fogg and Gross 2000; Mouton 2006; Christley 2010) and no computation by intention-to-treat analysis was reported (Heritier, Gebski and Keech 2003; Gupta 2011; de Boer *et al.* 2015; Unbiased Research 2017). Thus, with a final sample of <20 per group, the use of non-parametric statistics (Hoskin 2012; Bahreini *et al.* 2020), result in under powering and thus, the study can only make comment on trends not causation (treatment – outcome effect).<br><br>Both groups, receive a form of massage one with light and one with moderate pressure, respectively. Massage occurred as weekly unblinded therapist sessions and self-massage daily at home. The latter provides a clear confounding effect of pressure variability, failure to conduct the protocol correctly and inconsistency in the application (both in timing and frequency/adherence), even though adherence was monitored telephonically. By contrast the therapist-based sessions and self-administered sessions appear to be the same protocol aimed at the affected limb, requiring 15 minutes, in this providing a level of consistency. Another is the reporting of the intervention application, although it is appreciated that therapists and patients have different understandings of pressure and RA patients may have been compromised in applying pressure if they presented with any hand deformities.<br><br>The premise of the study was based on the expectation that massage provide relief of pain (Field, Diego and Hernandez-Reif 2007), where moderate pressure massage has positive effects while light pressure did not (Diego *et al.* 2004). Thus, expecting that moderate pressure massage would improvement in pain, grip strength and ROM makes an assumption that pain reduction will change grip strength and ROM. This assertion should first have been tested. By contrast, a strength of the study was the use |

| | |
|---|---|
| | of a mixed set of subjective and objective outcomes (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Most of the outcomes, measures had proven reliability and / or validity reported, although the objective measures had some concerns around them (Pugatch, Haskell and McNair 1969; Spielberger 1970; McNair, Lorr and Droppleman 1971a; Spielberger 1972; Snyder-Halpern and Verran 1987; Field *et al.* 2004; Field *et al.* 2011; Coster 2013; Bahreini *et al.* 2020). These concerns may result in bias, as they may not adequately reflect clinical outcomes in this specific population (Kimberlin and Winterstein 2008), given rheumatoid cachexia, hand deformities, and changes in strength (Liang 2000; Coster 2013; Bahreini *et al.* 2020). With respect to the outcomes assessment, inconsistency in the environment (Richardson 2007) and timing (influenced by the diurnal effect or flare-up of the condition (Kaur, White and Bartold 2012; Walker *et al.* 2013; Nandgude, Hasabe and Kolsure 2018)) has the ability to confound outcomes which could have been controlled (Platt, Schisterman and Cole 2009). Additionally, the timing between assessment and the intervention has the capacity to change assessment interpretation (immediate, short / intermediate or long-term effects) (Pater *et al.* 1998). This may be illustrated by the fact that strength training may require >8 weeks compared to the studies four weeks for clinical manifestations to be measurable (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). Lastly, the study failed to offer a baseline analysis of outcome measures between groups, prohibiting the reader insight into possible bias from differences between the groups.<br><br>From the underpowered perspective, the study made use of MANOVA, ANOVA and followed up with a post-hoc Bonferroni t-test for all outcome measures (Fogg and Gross 2000). The *post-hoc* Bonferroni *t*-test was used for interaction effect analysis in repeated measurements, meaning that the authors attempted to control for variance of the outcomes and thus improve the analysis. In addition, the study reports only means and not ranges, so it is difficult to compare the groups and understand the corrections that the *post-hoc* Bonferroni *t*-test controlled for.<br><br>Additional factors that may have impeded on this study include, a lack of a description of previous exposure to massage (applied and self-administered) in addition the patient perception of care and self-care in the RA context. Pain management in RA patients through pharmaceuticals influences pain perceptions and thus pain and ADLs outcomes. A lack of accounting for RA medications in addition to inflammatory management, if not accounted for can influence the outcomes. Similar concerns arise with anti-depression medications and its effects on outcome measures. Lastly, in terms of interpretation of the outcomes the authors make claims that the intervention improved ROM all joint in the upper limb, but these were not all assessed and reported on. This is exacerbated by the study only assessing one movement plane in the assessed joints. |
| Outcome | The authors noted, across the four-week intervention that the use of moderate versus light pressure massage produced significant improvement in both groups over time. They go further in reporting that the moderate pressure massage significantly decreased pain, improved perceived grip strength, and decreased depression in the pre-post intervention; whilst increasing objective grip strength, and improving ROM across the upper limb joints (namely wrist flexion, elbow flexion, and shoulder abduction). These findings, suggest that massage administration (self and therapist based) in patients capable of learning to self-administer, is both effective, would be cost-effective (for pain and ROM improvement) potentially improving QoL and enhancing activities of daily living (Field *et al.* 2013). |
| Discussion | The authors indicated that the outcomes of this study were favourable for both groups, with moderate pressure massage reflecting greater degree of improvement both intra- and inter- group comparison/s, even though the sample size was small and required the use of non-parametric statistics. This limitation, along unclear baseline measures, lack of control of specific RA dynamic factors, patient expectation and ability to perform at home treatment; may suggest that the outcomes are a reporting of natural history of RA / the outcome by chance alone and / or a regression to the mean given the variance between patients.<br><br>These collectively negatively affect the study's ability to accurately and fairly compare intervention groups without risking bias in outcomes and drawn conclusions. Additionally, failure to report methodological detail required for analysis of control and elimination of confounding and bias variables limits the readers capacity to evaluate the degree of risk in the study. The vague participant description prevents generalisability and implementation into clinical practice. The study findings are then not reliable and provide very little clinical guidance or evidence of relative effectiveness as an intervention to be incorporated into the management of RA. This study provides a basis for follow-up RCT studies to determine the true causal links between massage and its influence on the RA patient (subjectively and objectively). |
| Conclusion | As a result of the limitations noted in the table above, there is a possibility of a high degree of bias as it relates to the external validity of the study. In addition, there is ambiguity reflected in the review by the reviewers, indicating that the study's presentation lacked the clarity for accurate and clear understanding by all reviewers. Thus, when reviewing the statistically significant outcomes within the groups that are underpowered with strong evidence for bias (3/11) for internal validity in addition to the external validity concerns, the suggestion that one form of massage is better than that of another is limited at the very best. The studies capacity to provide evidence for the intervention nearly impossible. At best therefore, it is concluded that this study provides no evidence for moderate pressure massage (other they that provided by the underlying hypothesis) and no evidence for the support of its use in RA patients. This is based on the internal validity and external validity limitations and its impact on the ability to provide sufficient and clear patient description so as to facilitate the ability of field practitioners to identify which RA patients would benefit most. It is recommended that this study be re-done with a larger clearly defined sample size (clear description of the power analysis), with a viable comparison group, and clear criteria controlling for the RA clinical variability. This will also limit bias and indicate actual treatment effects avoiding natural history and / or chance alone in the outcomes. |

Table 4.45a: Tabulated feedback of reviewer data

| Author/s | Gizinska, M., Rutkowski, R., Romanowski, W., Lewandowski, J. and Straburzynska-Lupa, A | | | | | |
|---|---|---|---|---|---|---|
| Year | 2015 | | | | | |
| Title | Effects of Whole-Body Cryotherapy in Comparison with Other Physical Modalities Used with Kinesitherapy in Rheumatoid Arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 point (b) | 0 point (b) | 0 point (b) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 point (c) | 0 point (b) | 0 point (b) | 0 point (b) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 point (b) | 0 point (-) | 0 point (-) | 0 point (-) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 0 point (-) | 0 point (-) | 0 point (-) | 0 point (-) | 100% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (a) | 0 point (d) | 1 point (a) | 1 point (a) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 point (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 point (b) | 0 point (-) | 0 point (-) | 0 point (-) | 67% |
| | Total score | 3/10 | 2/10 | 3/10 | 3/10 | |
| | | | | | Overall percentage agreement | 75.25% |

Table 4.45b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Gizinska, M., Rutkowski, R., Romanowski, W., Lewandowski, J. and Straburzynska-Lupa, A |
| Year | 2015 |
| Title | Effects of Whole-Body Cryotherapy in Comparison with Other Physical Modalities Used with Kinesitherapy in Rheumatoid Arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - VAS (pain & fatigue)<br>- DAS 28<br>- HAQ-DI<br>- 50-meter walk test<br><br>Blood analysis<br>- ESR<br>- RF levels<br>- TNF-alpha<br>- IL-6 | Pre/post design | Two weeks | 44 participants in total, 25 in the whole-body cryotherapy and 19 in the traditional rehabilitation group.<br><br>No details on drop-out or intention-to-treat. | No | Not officially 2 intervention groups (1) 25 in the whole-body cryotherapy and (2) 19 in the traditional rehabilitation group. | No description of randomisation | 3/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | The Poznan University Medical School Hospital had the patients admitted to the Rheumatology Department for this study. This provides a perception that the sample consisted of patients that have previously sought specialist care or that the RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). This influences the generalisability of the study findings to this population only and must be kept in mind when applying these in a clinical setting. This aligns with the fact that patients that were contraindicated to physical treatment, were not eligible for and excluded from the study. As reported in the study, participants enrolled were female, post-menopausal, diagnosed with RA (1987 ACR criteria (Arnett *et al.* 1988)). The sample description included age, BMI, disease duration, serological type (RF positive yes/no), medication used (NSAIDs, DMARDs, Glucocorticoids and Prednisolone) and DAS28 (joint pain and tenderness, disease activity and patients' global health). All were distributed with no significant difference between the groups. Although it appears that the use of Prednisolone differed significantly (66.7% versus 84.2%). The study failed to control for a number of other dynamic factors of the RA (such as concomitant conditions, lifestyle factors, ethnicity, functional class and disability, imagining), which may have skewed data outcomes (type II error).<br><br>Forty-four participants created two groups, one comprised of 25 participants receiving whole body cryotherapy the other comprising 19 participants receiving traditional rehabilitation therapy. Thus, there was no true natural history control. The study fails to consider a power analysis, did not need to account for drop out and use of intention-to-treat analysis. These factors would underpin whether the study was underpowered or sufficiently powered to detect change (Kendall 2003; Mouton 2006; Christley 2010). Thus, the reader assumes that the study was underpowered. The two groups received a comprehensive baseline therapy including pharmacotherapy, supervised kinesiotherapy (individualised (dosage and type) according to functional capabilities, overall health, age, and severity of disease), and physical modalities. In addition, group one also received a prescribed whole-body cryotherapy and group 2 received a programme of therapy including a range of physical agents (magnetotherapy, electrotherapy, ultrasound therapy, and laser therapy), based on the international organizations against rheumatic disease (Braun *et al.* 2009; Hurkmans *et al.* 2011). The poorly described protocol of interventions in this study were administered daily on weekdays with breaks over the weekend, for a two-week period. The lack of detail prevents the reader understanding the effect of this programme on the reported outcomes, with particular emphasis on context of the study, parameters of these modalities or who receive them. The lack of clearly defined context and application, leads to concerns of bias associated with the Placebo effect, Hawthorne effect, touch therapy effect, patient expectation and patient naivety. Additionally, there is also a concern about the lack of blinding of both participant and therapist.<br><br>A positive for the study is the use of subjective and objective outcomes. Both pain and fatigue were measured by means of VAS (Collins, Moore and McQuay 1997; Brosseau *et al.* 2014), DAS28 (Kelly 1998), HAQ-DI (Prevoo *et al.* 1995), 50-meter walk test (time and step count) (Pincus, Callahan and Vaughn 1987; Bruce and Fries 2005; Callahan *et al.* 2008; Jastrząbek *et al.* 2013). These citations don't all match up with the outcome measures, and despite the likelihood of referencing error, it does take away from the validity and reliability the citations would impart (Coster 2013; Bahreini *et al.* 2020). Laboratory outcomes on fasting blood samples, from 16 in group one and 14 participants in group two; included ESR, RF, TNF-alpha, and IL-6. The application and measurement of the protocols followed were all well described, but lacked exercise |

| | |
|---|---|
| | intensity and effort measures, the measurement environment, the blinding of the assessor (therapist assumed to be unblinded) (Moustgaard *et al.* 2014), the timing of the assessment in relation to the intervention (even though it was a pre/post design set up) and of the day (diurnal effect) are not detailed (Pater *et al.* 1998; Nandgude, Hasabe and Kolsure 2018). In terms of the outcomes specifically, the study appears to not have been long enough to allow effective change in some outcomes, even though almost all findings have positive outcomes in this study (Fraser 2008). It does question the effects of participant bias (e.g., participant blinding, treatment novelty and effect of touch therapy) and / or research / assessor bias as influencing the results introduced into this study.<br><br>Given the sample size and the differential size of each group, parametric and non-parametric statistics had to be used (Hoskin 2012; Bahreini *et al.* 2020). This together with the non-normal distribution of some of the data, the floor / ceiling effect (as seen with the HAQ) (Liang 2000; Coster 2013) and the lack of reporting alongside MCID values (Bahreini *et al.* 2020) for the different measures make it difficult to tell what the outcome truly is in relation between the groups, over time and whether either were in fact clinically significant in terms of the application. This also brings uncertainty to the drawn correlations between variables as there were several significant factors that were not taken into account in the inclusion criteria and group homogeneity that may impact positively or negatively on these correlations and may result in type II errors of reporting.<br><br>The study seemed to suggest that there two groups improved significantly to the same degree over time, but where not significantly different from each other. This questions the effect of the pre-study treatment protocol that may have influenced the outcome. It further suggests that the traditional therapy group showed better improvements over the other group, which may imply that cryotherapy, but there is no measure to show this. Other suggestions include that all modalities including exercise reduce pain facilitating improvement. Which implies that neither group showed a higher degree of analgesic effect, thus leaving the authors stating that "this may suggest that comprehensive, monitored rehabilitation programmes are important, regardless of the type of physical treatment". This in essence contradicts the discussed effects of the cryotherapy. These inherent contradictions and small sample size, forces the authors to emphasize the need for caution in interpreting the presented results. |
| Outcome | Following two weeks of intervention programmes incorporating whole body cryotherapy (in group one) or traditional therapy (in group two), the study demonstrated that in RA patients, results indicate that regardless of physical therapy received significant improvement (with insignificant inter-group difference) was noted in multiple outcome measures including pain and fatigue severity, morning stiffness, step count and time taken during a 50-meter walk test, and DAS28 scores. Both groups reported functional status (via HAQ-DI) improvement to a significant degree, with traditional therapeutic groups responding significantly better than that of the cryotherapy group. Biological outcomes including TNF-alpha and IL-6 demonstrated significant improvement in both groups but not between them (Gizińska *et al.* 2015). |
| Discussion | The authors indicated that the outcomes of this study were favourable for both intervention groups. This conclusion is offered despite the admittedly small sample size and use of parametric and non-parametric statistics and thus providing a mix of observations some of which can only be reduced to trends. In addition, when considering the limitations noted in this table major areas of concern include a vague and poorly controlled inclusion criteria, poorly defined intervention protocol, and large range of modalities included into the intervention programme; the outcomes may have been affected by significant confounding or simply by natural history of the condition, by chance or with time. Finally, the results are reported in the context of a comparison between the two intervention groups that are not equal in size and showed baseline difference in some of the outcome measures. These cumulative limitations, negatively affects the study's ability to compare the two intervention groups resulting accurately and fairly in a biased outcome. This is further compounded by a lack of clarity in intervention programmes and a several bases for introducing risk of patient effects (around novelty, naivety and past experience with the intervention, lack of blinding, hawthorn and placebo effects) that would likely to have skewed the subjective outcome measures. |
| Conclusion | Given the limitations noted in this table, it is clear that there are several threats to the external validity of this study, placing it in a high risk of bias category. This is also reflected in the reviewers' review, which indicates that the study presentation lacked the clarity in the methodology and intervention/s. This introduces bias as the understanding by various readers is thus congruent and the ability to repeat the study becomes compromised. This indicates that the internal validity of the study only attained a 3/10 from the reviewers, again placing the study at a high risk of bias.<br><br>Therefore, the study does not provide evidence in favour of the intervention protocols better than chance or natural history, therefore future appropriately, designed trials are required in order to validate or refute the study's outcome and enable the outcomes to be appropriately placed in a setting that has applicability to clinical practice. |

Table 4.46a: Tabulated feedback of reviewer data

| Author/s | Hakkinen A., Hannonen P., Nyman K., Lyyski T., and Hakkinen K. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2003 | | | | | |
| Title | Effect of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: comparison with healthy subjects. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 2 points (a+b) | 2 points (a+b) | 2 points (a+b) | 83% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (e) | 0 points (c) | 1 point (a) | 0 points (-) | 0% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 0 points (b) | 0 points (b) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Total score | 4/10 | 7/10 | 6/10 | 6/10 | |
| | | | | | Overall percentage agreement | 73% |

Table 4.46b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hakkinen A., Hannonen P., Nyman K., Lyyski T., and Hakkinen K. |
| Year | 2003 |
| Title | Effect of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: comparison with healthy subjects. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - Grip strength<br>- Trunk extension & flexion<br>- Bench press<br>- Leg press<br>- Leg extension<br>- Max walking speed (10-meter walk and 10-stair climb test)<br>- Dynamic explosive force of low limb<br>- $VO^2$max<br>- Heart rate<br>- Blood pressure<br>- Blood lactate concentrations<br>- Aerobic and anaerobic threshold<br>- Borg's rate<br>- Maximum cycling time<br>- VAS (pain and general health)<br>- HAQ<br>- BDI<br>- Joint index<br>- ESR<br>- Hb<br>- Fat % | Weeks -2, 0, 7, 14, 21.<br><br>Aerobic capacity was only measured pre/post. | 23-weeks | Total sample (RA + control) = 35<br>23 female patients with RA: early RA (n=12) and long-standing (L-S) RA (n=11)<br>The control of healthy participants (n=12) [age, sex, weight, and height matched].<br><br>Drop-out: one longstanding RA after five weeks, one early RA due to tenderness of the small joints of the feet. one early RA could not attend $VO^2$max testing at 21 weeks due to pregnancy. | Yes | Yes, healthy participant made up the control that was age, sex, weight, and height matched to the RA participants. | No | 6/10, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 73% |

| | |
|---|---|
| Limitations | The participants were recruited from the Rheumatology Unit at Jyväskylä Central Hospital, Jyväskylä, Finland. This creates the impression that the sample consists of patients that have previously sought specialist care, implying that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). The healthy control was recruited via flyer from the city of Jyväskylä. The bias associated with recruitment context raises concern around limitation to generalisability of the study findings.<br><br>Inclusion criteria required participants with established stable RA [1987 ACR (Arnett *et al.* 1988), and functional class I or II]. The disease duration of those with L-S RA had a mean of 14.5 years (SD 4.5, range 8-21) and the short-term RA group reported mean of 2.9 years (SD 0.6, range 2-4). These also had pre-diagnosis symptoms at 15.5 years (SD 4.8, range 10-24) and at 4.3 years (SD 1.9, range of 3-9 months) respectively. In terms of medication control, all patients were on DMARDs. NSAIDs and steroid injections were variable but consistent between the groups. This detail assists in excluding medication as a confounding variable; however, dosages, administration and compliance are not considered or documented in the study. In comparison the three groups did not differ significantly in terms of age, height, weight, and body fat percentage according to the authors although age reflected as $p$= 0.059. In terms of RA dynamic characteristics, the L-S group's mean age was 49 years (SD 7), while early RA and control had means of 41(SD 9) and 42 (SD 9) respectively. Thus, homogeneity is potentially possible subject to the complete spectrum of RA dynamic characteristics having been considered.<br><br>Given the above group description, no detail of a power analysis was provided. This increases the risk of an underpowered study (limited to non-parametric statistics), which is acknowledged by the authors who go on to state that clinically relevant changes in outcome measures could have been missed or go undetected; by contrast the study was of sufficient length to detect physiological and anatomical effects. These factors influence outcomes and conclusions that may be taken into clinical practice. Despite dropouts, intention-to-treat analysis was not conducted. The statistical limitations are however not compounded by patient naivety, as the participants reporting of exercise participation at |

| | |
|---|---|
| | baseline was limited to all and the types of exercises and prior strength training was limited but common between the groups. Therefore, it is unlikely that prior exposure and thus patient perception bias would have influenced the outcomes. In addition, a two-week pre study period was required to stabilise the "regular" exercises performed by participants, which provided a unique to control confounding associated with previous exercise exposure. The actual programme of supervised exercises involved concurrent strength and endurance training (over 21-weeks), with each (strength and endurance) administered three times a fortnight (two weeks). The exercises were performed in a clinical setting, controlled for intervention timing, adherence, and protocol consistency). All procedures were explained orally and in text. The supervision prevented blinding, it enabled consistency in programme. The lack of blinding may have influenced the results due to effects such as Hawthorne and variable encouragement.<br><br>The exercise programme was well defined allowing for replication of the intervention (Häkkinen *et al.* 1998), although it was noted that there was a protocol variability, allowing for individualisation of the exercises to maintain patient interest and interdigitate with RA change introduces bias - as the group now becomes a case series as opposed to being a homogenous group against which one intervention is tested. For protocol adherence, the study suggests that this was no concern for bias, this may be because missed sessions could be "caught up". This raises concerns of protocol consistency / individualisation and the effect on timeframe between assessments and outcome measures, as this will affect recorded data.<br><br>All assessors were blinded to participant group allocation and outcomes included a balance of objective and subjective outcome tools. The protocol was detailed and thus reproducible for muscle strength measurements and grip strength (Rantanen, Airaksinen and Penttinen 1994; Häkkinen *et al.* 1998), dynamic assessments, RA disease assessments (Scott and Houssien 1996), aerobic performance, perceived levels of exertion, blood work and other assessments (Durnin and Rahaman 1967; Borg 1973). The grip strength measures were however summed to create a data point, which prevents assessment of hand-dominance or variable RA in each hand. Of the subjective outcomes, the VAS (Price *et al.* 1983), HAQ (Fries 1983), Beck's depression index (Keltikangas-Järvinen and Rimon 1987) were all relevant, appropriate, and appropriate. The structured reporting of the outcomes are different time intervals takes into account the periods in which detectable change can be anticipated (e.g., >eight weeks for muscle changes). These data were then compared to the period between baseline and week two assessments, which served as the control for the study (i.e., no intervention). This latter provided the study a unique control over previous exercise exposure not often employed in physical modalities studies. Thus, the biggest limitation is that due to the design of the study and utilisation of a programme (aerobic and endurance training) causal link cannot be formally made between an outcome and a component of the intervention. Furthermore, whether the different parts of the programme where synergistic/antagonistic or did not influence the results cannot be determined. Additionally, a clinical limitation is that outcomes were not reported in comparison to published MCID values to aid the reader in contextualising outcomes for a clinical setting (Bahreini *et al.* 2020). |
| Outcome | The study reported that female participants with either early or longstanding RA, receiving a 21-week exercise programme (of concurrent strengthening and endurance exercises), one and a half times a week; showed significant improvement in maximal strength, walking speed, vertical squat jump height, reduction in body fat percentage and maximal aerobic capacity. Thus, concluding both early and longstanding RA patients with stable disease can safely improve all characteristics of their physical fitness profile using a progressive concurrent strength and endurance training protocol (Häkkinen *et al.* 2003). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though the sample size was small. Taken in addition to other considerations with the limitations noted the table above, suggestions may be made for outcomes to have been affected by factors such as natural history of the condition, chance, time and / or be a result of a regression to the mean in terms of outlier effects in small samples. These negatively affect the study's ability to compare the intervention groups accurately and fairly, resulting in a biased outcome. |
| Conclusion: | As a result of the limitations noted in the table above, the study is at moderate risk of bias due to the external validity issues. This concurs with the reviewers' outcomes that suggest 6/10 ranking or moderate bias based on the internal validity assessment, thus indicating a moderate level of evidence in support of a strength and endurance training. Given the fairly well-defined population in the study, this data may be extrapolated to patients seen by field practitioners, providing clinical relevance.<br><br>Nevertheless, it is recommended that despite this evidence this study be re-done with a larger sample size (incorporating power analysis), tested on multiple different RA patient groups, and referenced to MCID values to further limit bias and indicate actual treatment effects are better than natural history and / or chance alone in the intervention groups. |

Table 4.47a: Tabulated feedback of reviewer data

| Author/s | Hawkes J, Care G, Dixon JS, Bird HA, Wright V. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1986 | | | | | |
| Title | A comparison of three different physiotherapy treatments for rheumatoid arthritis of the hands. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 point (c) | 0 points (b) | 0 point (c) | 0 point (c) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 0 point (-) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (a) | 1 point (b) | 1 point (a) | 1 point (a) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 point (b) | 1 point (a) | 0 point (b) | 0 point (b) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 point (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Total score | 3/10 | 6/10 | 5/10 | 5/10 | |
| | | | | | Overall percentage agreement | 79% |

Table 4.47b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hawkes J, Care G, Dixon JS, Bird HA, Wright V. |
| Year | 1986 |
| Title | A comparison of three different physiotherapy treatments for rheumatoid arthritis of the hands. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Grip strength<br>- Joint size<br>- Pain VAS<br>- Articular Index<br>- ROM<br>- Timed task<br>- Check list of activities | Baseline and at end of weeks one, two and three. | Three-weeks | 30 total (10 per intervention group), no reported drop out. | Yes | No | Yes, by means of an unknown method. | 5/10, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 79% |

| | |
|---|---|
| Limitations | The Regional Rheumatology Centre at the Royal Bath Hospital, served as a base for this study and was the region from which the sample was recruited. This provides the perception that the sample consists of patients that have previously sought specialist care, implying that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). Thus, clinical application of outcomes may only be limited to a similar group.<br><br>Once approved for entry into the study by their hospital consultant, the first 30 participants who met the inclusion criteria of: having RA (American Rheumatism Association criteria classic or definite RA (Ropes *et al.* 1959)) and specifically RA of the hands (with painful, swollen joints limiting ROM) where included. During the study period, participants were required not to change their pharmaceutical therapy and were prohibited from receiving steroid intra-articular injections. These criteria provided a limited control over the RA and its dynamic presentation. It omits RA disease activity and duration, RA functional class, cultural and social variables that influence hand use and disease progression, patient endocrine status, surgical history, and comorbidities, along with specific control over medication (stable pre-study regiment, number, type, dosage, and administration of medications (DMARDs/NSAIDs/etc), prior treatment washout periods (medicinal and non-medicinal) (England *et al.* 2019; Myasoedova *et al.* 2020). These factors may be a causative agent for change unrelated to the actual study and set the scene for a type II error in accepting / not accepting a planned hypothesis in the study (Bartley and Fillingim 2013; de Boer *et al.* 2015).<br><br>The study fails to report an a priori analysis and used a sample of 10 participants per group, under powering the study from the start and providing limitations for statistical analysis and conclusions drawn (Kendall 2003; Mouton 2006; Christley 2010). The sample was randomised (method unknown) with no noted concealed allocation into one of three intervention groups (group A (exercise and wax bath); group B (exercise and ultrasound); and group C (exercise, ultrasound, and faradic hand baths)). These groups where reportedly well matched and showed no statistically significant difference at baseline (diagnosis of RA confirmed, dominant hand noted (27 of 30 were right-handed), occupation recorded (of which are not detailed nor reported). Thus, the authors report group homogeneity which was maintained as there were not reported dropouts (Wells *et al.* 2012). An assumption was made that participant/s were not blinded, as this remained unreported, and the participants all received exercise component (Moustgaard *et al.* 2014). The supervised exercise component was defined as “standard” and “routine”, but no details of this therapy were given. In addition to this, Group A received wax hand baths, the technique used was detailed. Group B received therapeutic ultrasound, the technique and dosage were detailed. Group C received both ultrasound (like that of group B) and faradic hand baths, the technique of which was “standard” and described. All treatment groups received parallel interventions, even though clear exercise parameters were unpublished (Heine *et al.* 2012). The details of administration of intervention including environment, timing, therapist/s, blinding of therapist/s, or other protocol variables were not defined (Kendall 2003; Moustgaard *et al.* 2014). Given these interventions it is clear that each has the ability to achieve clinical changes in three weeks with the exception of exercise (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Piva *et al.* 2019). Additionally, patient response to touch therapy effect (not all received this), the placebo effect, Hawthorne effect and treatment novelty and naivety effects are not clearly delineated in the study. Finally, the synergistic or antagonistic effect of the coupled interventions is unknown and referenced material is provided to allow the reader context and the study may thus have benefitted from a control group allowing comparison to natural history (Robins, Hernán and Siebert 2004).<br><br>To the study’s credit, a balance of subjective and objective outcomes were used. Were applicable the single blinded assessor, created consistency in assessment protocol and readings. However, the study failed to detail the environment, the timing of the assessment (relationship between intervention and the measurement, and |

| | |
|---|---|
| | the time of the day accommodating diurnal RA pattern). But the outcome tools did measure grip strength, joint size, pain, articular index, ROM, timed task, and checklist of activities which covered a broad scale of outcomes, although reliability and validity of these tools for the RA population was unreported.<br><br>The study fails to report a dropout of participants nor the utilisation of the intention-to-treat analysis for outcome measures. For this reason, the reader is led to believe that no loss was experienced by the study, which mitigated against further impact to the already small sample size, restricting the statistical strength of the study. All recorded data was presented as a mean, where outcomes were bilateral the reporting was an average of both for all participants in the group, with the exception of participant numbered 9 – from group C, and participant numbered 18 – from group A; where the data was stratified to report means of the hands (right and left) separately. This was because two participants were reported as outliers to the data, one as a result of a RA flare-up during the third week of the intervention and the other due to non-response. As noted in the PIP circumference and the ROM outcomes the data were averaged for the analysis, which raises the concern of regression to the mean and small changes mitigate against larger changes. After data capture, the study made use of the Kruskal-Wallis one-way analysis of variance, to compare treatment group variability. This analysis reported no significant difference in all outcomes at baseline. When comparing the assessments at baseline and each subsequent week assessment for intra-group difference. Improvement was observed in a majority of outcome measures for each of the three groups, with a reported statistically significant improvement in half of all outcomes at the week one assessment. At the end of the study the results of the three intervention groups provided variable outcomes with statistical significance ($p$<0.05) between the groups present for activity score only. Here group A and B reporting insignificant differences, while group C reported a significant difference. The authors claimed this outcome as a side effect of the small sample size.<br><br>When considering hand dominance, the changes from baseline to week three for the four outcomes (grip strength, PIP circumference, ROM, and articular index) was analysed by the Wilcoxon rank sum test. There was no reported significant difference in response for any treatment. However, when conducting this analysis in the overall comparison within a measurement; grip strength reported better responses in the dominant hand ($p$<0.05). This was not seen in the other three outcomes. Additionally, none of the changes in any group for any outcome was compared to published MCID values to contextualise the clinical changes due to the interventions (Bahreini *et al.* 2020).Lastly, although good for the study, the use of therapeutic ultra-sound (in clinical practice) is normally tailored to the response of the individual patient (in terms of dosage/intensity and duration). For this study however, to control for the modality across the participants, its parameters were set and limited to ensure identical exposure. |
| Outcome | The study of 30 RA participants receiving one of three intervention programmes of five-days a week for three-weeks, revealed that statistically significant improvement was gained for half of the outcome measures at week one, but this was not different between groups. All outcome measurements (including grip strength, joint size, pain, articular index, ROM, timed task, and checklist of activities) reported significant improvement for all groups by the end of the third week (with exception of ROM and checklist of activities in group C) within the groups. At the end of the programme (week three) statistical differences were only noticed between group for grip strength (after dominance correction) and articular index (between the groups) (Hawkes *et al.* 1986). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though the sample size was very small and would have required the use of non-parametric statistics, thus only allowing for trends to be observed. This further limitation, of a semi-detailed inclusion criteria and a lack of baseline demographic data comparisons impact on outcomes as they enable a type II error the acceptance or non-acceptance of the stated hypotheses.<br><br>The use of mean value data does not allow the reader to assess the impact of outliers and can result in regression to a mean due to external factors nor monitored in this study (e.g., RA flareup). This impacts on outcomes statistically and may change conclusions drawn. Thus, accurate and fair comparisons may be obviated in the study. Further to this the type of intervention programmes and lack of a natural history control group, which would provide a base for the control of patient effects (around novelty, naivety, and past experience with the intervention – particularly when participants were recruited at a specialist centre and share a base intervention (exercise) and compliance effects. |
| Conclusion | As a result of the limitations noted in the table above that stem from the analysis of an external validity of the publication, there is a moderate to high levels of bias that have been identified to impact on the study outcomes. This is not dissimilar to the score of 5/10 from the collective reviewer recording for the internal validity evaluation, which signals a moderate level of bias. Thus, collectively the study suffers with a slightly more the moderate level of bias and thus can only provide a limited level of evidence in terms of the interventions tested on the RA population.<br><br>Therefore, to attain a better insight into the outcomes of this study and their veracity, it is recommended that the study be repeated on a larger scale, with consideration of the limitations of this study addressed and with a formal control group. This will then either refute or confirm the findings and allow for recommendations for the use of these therapies in clinical practice. |

Table 4.48a: Tabulated feedback of reviewer data

| Author/s | Hurkmans, E. J., Van den Berg, M. H., Ronday, K. H., Peeters, A. J., Le Cessie, S. and Vliet Vlieland, T. P. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2010 | | | | | |
| Title | Maintenance of physical activity after Internet-based physical activity interventions in patients with rheumatoid arthritis | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (c) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (-) | 0 points (-) | 0 points (-) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 0 points (-) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (a) | 0 points (d) | 0 points (d) | 0 points (d) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Total score | 3/10 | 4/10 | 3/10 | 4/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.48b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Hurkmans, E. J., Van den Berg, M. H., Ronday, K. H., Peeters, A. J., Le Cessie, S. and Vliet Vlieland, T. P. |
| Year | 2010 |
| Title | Maintenance of physical activity after Internet-based physical activity interventions in patients with rheumatoid arthritis |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation | Ranking | Total Percentage Agreement |
| Physical activity questionnaire, HAQ, RAQoL, SF-36 – RAND<br>Health care: Medication / steroid / surgery (prior year) by medical record/s, use of medical services (prior 3 months - questionnaire. | Baseline and months 12 and 24. | Two-years (One year intervention) | 160 total sample: group 1 (n=78) and group 2 (n=82).<br>At 12 months: group 1 (n=75) and group 2 (n=77), At follow-up: group 1 (n=54) and group 2 (n=56). | No, outcomes are solely self-administered and subjective. | No, two intervention groups. | Yes, method not stated (see comment in limitations box about this) | 4/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | This study follows on from an initial multi-centred RCT conducted from 2002 to 2004. The sources for participants included the same three hospital outpatient clinics for rheumatology. This follow-up study evaluates the maintenance of effects reported in the forementioned RCT and thus utilises participants RCT. The location and recruitment strategy thus leads the reader to assume the participants are more likely to be seeking advanced / specialist care. Following this assumption, it would imply that their RA disease activity and severity may be more severe than in the average RA patient (Wells *et al.* 2012). This would inherently limit the study's generalisability. This study tracks the participants of a previous RCT conducted (Van den Berg *et al.* 2006). The original size of the sample recruited in the RCT was that of 160 participants. The inclusion criteria employed in the initial recruitment included an RA diagnosis by means of the ACR 1987 criteria (Ropes *et al.* 1959; Arnett *et al.* 1988), physical activity less than a level of moderate intensity (defined as 30 minutes of continuous training for five days a week), capacity to undergo assessment on bicycle ergometer, and having computer and internet access. The exclusion criteria for this recruitment were not detailed in this publication. However, the current study excluded two participants from the follow-up study because of their inability to meet the standard of moderate intensity physical activity as is recommended in public health. These two excluded participants did, however, meet the recommendations of vigorous physical activity as per public health requirements. Following the initial study's recruitment, these participants were randomised into one of two groups (method of randomisation is not detailed in this publication – Van den Berg *et al.* (2006) that it was conducted by means of "random digit generator, using blocks of four participants with stratification for centre and sex. This process was carried out by an independent physical therapist who was not responsible for recruiting the participants"). Group 1 was the general training and Group 2 the individualized training. Group 1 initially had 78 participants while group 2 had 82 participants. Of these, 152 completed the RCT (75 participants in group 1 and 77 participants in group 2), thus the drop-out during the intervention stage was 8 (5%)). 110 participants (54 in group 1 and 56 in group 2, representing 69% of baseline sample and 72% of completers) of these completers of the 2002 to 2004 RCT were eligible for the follow-up (the 2010 study) at 24 months, thus the follow-up study suffered a dropout of 40 (from baseline 25%) and 32 (from completers – 21%). These dropouts were reported with reasoning.<br>These 110 participants did not differ significantly from those who were lost-to-follow-up (at baseline and at 12-month / post-intervention assessment) in terms of socio-demographics and disease variables (Group 1 and group 2: age (median of 51.0 with a interquartile range of 10.9 and median of 50.6 with a interquartile range of 13.1), sex (40 and 43 females respectively), disease duration (median of 5.9 with a interquartile range of 12.4 – years and median of 8.0 with a interquartile range of 8.1), BMI (median of 24.9 with a interquartile range of 6.7 and median of 25.7 with a interquartile range of 5.5), co-living (7 and 10 living alone respectively) and employment status (31 and 32 maintaining employment respectively), and co-morbidities (20 and 26 respectively)), and outcome measures (this data was not shown in the publication but stated). For the above reasoning, the reader can assume groups-maintained homogeneity. A major limiting factor in this study for the generalisability of outcomes is acknowledged by the authors, in that the entire sample is made up of sedentary RA population. As is noted by the authors, when the initial sample size is compared with the sample size available at the follow-up at 24-months, almost one quarter of the initial sample are not available. This will influence the generalisability of the outcome of this study. However, this is a follow-up from another (Van den Berg *et al.* 2006) and has detailed the interventions poorly in this publication making reference to the prior study publication. Despite this, the general training group (group 1) consisted of general information delivery which covered physical activity and was delivered to the participants of the group via email and website. The individualised training group (group 2) consisted of individualised guidance delivered to the participant via email and website, in addition they received a bicycle ergometer with group contacts. The intervention phase of the study was 12 months in duration. Unfortunately, this publication does not offer any further details regarding the intervention, leading the reader to assume several confounding and bias factors at play. These include protocol description and consistency, effect of group participation (seemingly limited to group 2), blinding of a therapist in the group setting, the content of the individualised email and its synergistic/antagonistic effect on the ergometry, and the failure of publication to allow for repetition of the treatment protocol (Robins, Hernán and Siebert 2004). As was the case with the initial study (Van den Berg *et al.* 2006) |

| | |
|---|---|
| | the outcomes were carried out in the same way at 24-month follow-up. This was achieved by means of questionnaire outcome tools; thus, outcomes are solely subjective. These questionnaires enquired into several areas. (1) Physical activity measured in reference to the proportion of participants meeting the recommendations for public health (Ooijendijk, Hildebrandt and Stiggelbout 2000). This data was achieved by means of two questions, the first of which enquired as to the frequency of moderate intensity physical activity per week. Defined as thirty minutes of continuous exercise, achieved by any means of physical activity resulting in small increase in respiration or heart rate during the past three-months. The second of like that of the first but enquired into vigour's exercise. Defined as twenty minutes of continuous exercise of any form that resulted in large increase in respiration or heart rate during the past three-months. This introduces (as is acknowledged by the authors) a risk of memory recall / decay bias (however, the authors note that an accurate and valid means of assessing this is still a major challenge). Derived from these questions the study calculated the proportions of exposure to these two intensities of exercise; Functional ability via HAQ (Spitz and Fries 1987); RAQoL questionnaire (de Jong *et al.* 1997), supplemented by validated (SF-36) (RAND) (Hays, Sherbourne and Mazel 1993); Health care was measured by medical records documenting medication changes (DMARDs and / or NSAIDs) and steroids use and surgery, during the preceding year to follow-up assessment. As well as medical services use in the three-months prior (questionnaire). The second part of health care measures was subject to memory recall / decay bias (Mouton 2006; Coster 2013; Ricker, Vergauwe and Cowan 2016; Bahreini *et al.* 2020). A major confounding factor in most physical therapy studies in RA is that of medication. It represents a difficult factor that is unethical to enforce cessation of pharmaceutical therapy during a study particularly for longer studies. The best that can be done in this context is to control for change and effect of these medications. This study made one of the outcomes focuses medical care, which included medication tracking. These outcomes enabled tracking of these parameters; DMARDs and NSAIDs (although limited data expressed regarding these medications) did not significantly changes in these medications, along with steroid injections (intra-articular) or surgery during the follow-up period. Furthermore, medical services used by participants who did and did not meet the public health exercise recommendations (moderate intensity), during the preceding three-months (to follow-up at 24 months) did not significantly change. These findings are consistent with the that of any time point during the 24 months. The study fails to report use of power analysis in this publication, did not make use of intention-to-treat analysis, and reports on MCID values for the readers reference for clinical context (Bahreini *et al.* 2020). However, baseline differences (as reported above) were insignificant. The study utilised a mix of parametric and non-parametric analysis tools to determine both inter and intra group differences and change at the difference assessment points. This analysis also looked at outcomes according to exercise intensity not only within the randomised group allocation. |
| Outcome | The study followed an intervention programme run over 12-months, after which follow-up assessment was undertaken 24-months after enrolment. 152 of 160 participants completed the intervention period, of these 110 participated in the follow-up. At this follow-up, the proportion of these participants meeting recommendations of public health for moderate exercise intensity was significantly ($p < 0.05$) increased when compared to baseline in both groups (group 1: 24% increase and group 2: 19% increase). But the proportion reported of participants meeting the vigorous exercise intensity recommendations as per public health, where insignificant in group 1 and significantly increased in group 2 ($p < 0.05$) when compared with the baseline. However, no significant difference could be found between group relating to proportion of participants meeting either moderate or vigorous exercise intensity recommendations at follow-up. Finally, the study reported that group 2 reported significantly higher scores of the RAQoL at follow-up compared to baseline. This was lacked significance, both intra- and inter- group analysis for QoL and functional ability. Thus, suggesting effectiveness of one-year moderate intensity exercise, improving physical activity in sedentary RA patients were maintained at one-year post-intervention. |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, despite the significant loss-to-follow-up and lack of clarity regarding intervention content. The study was able to conduct parametric statistical analysis, but the observed outcomes should be cautiously read. Limitations, along with the limited inclusion criteria and lack of comparison (natural history) groups, suggest that the outcomes could have been affected by factors such as natural history of the condition / chance / time, and be the result of a regression to the mean regarding the comparison between the two groups. This negatively affects the study's ability to compare the two intervention groups accurately and fairly, thus resulting in a biased outcome. Furthermore, the type of intervention programmes and the exclusive use of subjective self-administered outcomes provided a base for the introduction of patient effects (around novelty, naivety, previous experience with the intervention, and memory recall/decay bias) that are likely to have affected patient compliance and responses to the outcome measures. |
| Conclusion | The limitations noted in the table above as well as the ambiguity reflected in the review of the study by the reviewers indicate that the study presentation lacked the clarity required to allow for an accurate and clear understanding of the methodology. This alone introduces bias as the understanding by various readers is thus compromised and significantly hinders the ability to repeat the study and attain the same results. Thus, in respect of the outcomes, some of which were statistically significant, the suggestion that neither programme is better than the other must be cautiously considered; both these interventions are significantly limited regarding the study's the ability to provide evidence for any component of either programme, particularly given that the interventions are very poorly reported. At best, therefore, it is concluded that this study provides no evidence for any single intervention and very limited evidence to support either programme for RA patients. In addition, the study's limited patient descriptions and acknowledged limitations in generalisability due to its recruited sample, compromising the field practitioner's ability to identify which RA patients would benefit most. It is therefore recommended that this study be re-done with a clear natural history control group, with clearer inclusion criteria to assist in limiting bias, with the interventions described in far greater detail (despite the reference made to van den Berg et al. (2006)) and a clear indication as to whether treatment effects are better than natural history and / or chance alone in either or both of the intervention groups. |

Table 4.49a: Tabulated feedback of reviewer data

| Author/s | Karatepe, A. G., Günaydin, R., Türkmen, G. and Kaya, T. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2011 | | | | | |
| Title | Effects of home-based exercise program on the functional status and the quality of life in patients with rheumatoid arthritis: 1-year follow-up study. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 0 points (b) | 1 point (a) | 0 points (b) | 0 points (b) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (c) | 0 points (c) | 0 points (c) | 0 points (c) | 100% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (e) | 0 points (d) | 0 points (d) | 0 points (d) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (-) | 0 points (b) | 0 points (b) | 67% |
| | Total score | 2/10 | 2/10 | 1/10 | 1/10 | |
| | | | | | Overall percentage agreement | 83.5% |

Table 4.49b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Karatepe, A. G., Günaydin, R., Türkmen, G. and Kaya, T. |
| Year | 2011 |
| Title | Effects of home-based exercise program on the functional status and the quality of life in patients with rheumatoid arthritis: 1-year follow-up study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - HAQ<br>- RAQoL<br>- DAS28<br>- Duration of morning stiffness | Baseline, weeks four and 52 (one-year). | The intervention was four weeks long with a one-year follow-up. | 44 patients started the initial study. of which 28 were available at 1-year follow-up.<br>The 44 participants were completers of the four-weeks exercise protocol. | Not mentioned | No control for the one-year follow-up article. | No randomisation | 1/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 84% |

| | |
|---|---|
| Limitations | There is no clear description of study location, and the recruitment context is not clearly described in this follow-up publication, seems to be a long-term follow-up on a short-term study that employed an exercise programme. With the lead author was associated with the Department of Physical Medicine and Rehabilitation in Turkey at the Izmir Bozyaka Training and Research hospital, it may appropriate for that to be considered the setting. Given the limited clarity of the context of this study, the reader is left to assume participants of this study at this hospital are more likely to seek advanced / specialist care. Following on from this assumption, it would imply that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). This would inherently limit the studies generalisability to patients of a similar type in practice.<br><br>Seeing as this study may be a follow-up from an initial study, from which completers were asked to continue exercise for an additional next year. The article makes reference to 44 participants recruited with RA according to a set of unspecified ACR criteria for which the inclusion criteria outlined stable RA disease for at least three months prior to recruitment; between low and moderate RA disease activity (criteria were not defined); a functional class of I or II according to the ARA (not cited by study); no significant cardiovascular disease, joint replacement; planned lower limb surgery and did not participate in regular exercise programmes (outside of the initial study which required exercise for four weeks prior entry into the long term arm of the study). Following the four-week intervention, completers of the previous study were invited to a control visit. These participants were advised to continue forward with the same exercise programme to a one-year follow-up (this study), planning for medical treatment was undertaken in the case that disease flare-up or side effects / adverse reaction to pharmaceutical therapy occurred. In terms of the efforts to enhance adherence to intervention were ceased.<br><br>At baseline / end of the four-week study, results showed 68% engaged in training that involved more than one joint. The frequency of joints trained include the hand (68%), wrist (68%), knees (68%), shoulder (45%), elbow (25%), foot (20%), ankle (16%) and hips (11%). Of the 44 participants, 28 completed the one-year follow-up (64%) implying a 36% drop out, six had surgical intervention while the remaining 10 could not be reached. At follow-up, the participants available for follow-up assessment were compared to the loss-to-follow-up completers. The age, sex, disease duration, morning stiffness duration, DAS28, HAQ, and RAQoL were all not significantly different. It is not clear from the statistics as to whether an *a priori* analysis was done either prior to the initial study or with concern for the more long-term study that is reported here. Therefore, this study does run the risk of being under powered (Kendall 2003; Christley 2010).<br><br>The intervention was tailored to the physical examination results of the participants, taking into consideration RA active joints and those previously affected. The intervention consisted of strength and ROM exercises for which each participant was trained and given written instructions (with illustrations). Each carried out with 10 to 15 repetitions, twice a day, five times a week for four weeks (Heine *et al.* 2012). To record compliance an intervention chart was completed by participants and weekly adherence calls were made to the participants along with exercise chart collections at the end of the intervention. Thermotherapy (specifically application of heat prior to exercise and cold application after) was advised, along with advice that participants must report to their physician should swelling and warmth or pain of joints post exercise last more than an hour. Unfortunately, the current publication (without access to the preceding and cited study) does not provide additional detail required to repeat the intervention in either a clinical setting or research. Despite the participants controlling the frequency of participation in exercise, those assessed at follow-up assessment “almost all” reported having continued exercising albeit less frequently. Given that the exercise frequency decreased, changes in compliance over the course of the year may also have waxed |

| | |
|---|---|
| | and waned (in line with RA fluctuations), thus the implications on measuring the actual effect of an annual exercise programme should have had better controls (Kaur, White and Bartold 2012; Walker *et al.* 2013).<br><br>Assessment of the participants occurred at baseline, post-intervention (at four-weeks), and a one-year follow-up (the last the reason for this publication). These assessments were all conducted with the same set of outcome tools by a physician (clarity of consistency and blinding of the assessor in not reported). The outcome tools included the duration of morning stiffness (tool not clearly defined), RA disease activity and swollen and tender joint count (by means of the DAS28), functional status (by means of the Turkish version of the HAQ) (Fries *et al.* 1980; Küçükdeveci *et al.* 2004), and health related QoL (by means of the Turkish version of the Rheumatoid Arthritis QoL Questionnaire (RAQoL)) (Kutlay *et al.* 2003). Given this, the outcome tools were appropriately chosen and implemented making the outcomes valid and reliable with a good balance between subjective and objective outcomes. It must be noted though that the HAQ was evaluated the week prior to the assessment risking subjective memory recall/decay bias. But then many subjective outcomes are also at risk of placebo and hawthorn effect as well as memory recall bias (Mouton 2006; Ricker, Vergauwe and Cowan 2016).<br><br>In terms of the statistical analysis, the study made use of Kolmogorov-Smirnov test (normal distribution of data), descriptive statistics were employed to define the demographic and clinical descriptors of the sample and the independent sample *t* test, Mann-Whitney U test and Chi-square test were used for between group differences. An ANOVA using the Geisser-Greenhouse correction to correct for violations of the sphericity assumption, and Bonferroni correction for number of significant differences; indicated that the data could have skewed and required statistical correction to allow approximation to a norm within the group. This is problematic as there was no control group for the annual study, thus not providing for a comparator in terms of the outcomes. Furthermore, intention-to-treat analysis for dropouts does not seem to have been considered. |
| Outcome | Following a four-week exercise-based intervention, completers of this four-week period went on to self-regulated exercise participation for the subsequent year. At the four-week assessment the outcomes reported significant improvement in functional status (HAQ: $p<0.05$) and health-related QoL (RAQoL: $p<0.05$); while DAS28 proved insignificantly altered ($p=0.305$). At the one-year follow-up, there was a reported maintenance of outcomes with no significant change between week four and 52 assessments in functional status, health-related QoL, and DAS28. The comparison of the one-year follow-up assessment against baseline, still showed statistically significant differences in HAQ ($p = 0.015$) and RAQoL ($p=0.037$). However, DAS28 did not achieve a statistically significant difference between baseline and one-year follow-up. The authors concluded that "we observed significant improvements in the functional status and health-related quality of life by 4-week home-based exercise programme. Furthermore, these improvements were maintained at 1-year follow-up. Further studies are needed to confirm the usefulness of the home-based exercise therapy" (Karatepe *et al.* 2011). |
| Discussion | The authors of this study indicated that the outcomes of this study were not only significantly favourable in two of three outcome measures, but these were maintained at the one-year follow-up. These findings are despite the concerns around the sample size, significant drop out, no comparison group, in addition to protocol adherence and compliance. The vague methodology of the first four-week intervention study (as reported in an inaccessible publication), suggests that the outcomes are subject to confounding and bias factors that may be subject to each or a combination of natural history of the condition, chance and / or time. Furthermore, the reporting of the data without a control / comparison group, negatively affects the study's ability to compare the intervention accurately and fairly to a baseline, resulting in a potentially biased and misleading outcome. Further to this the type of intervention (home-based exercise followed by a self-defined exercise exposure over the follow-up period) provided a base for the introduction of patient effects (around novelty, naivety and past experience with the intervention) and protocol inconsistency, exercise variables necessary to induce physiological change all of which are likely to have affected patient compliance and responses to the outcome measures and the true effects of the intervention on the outcomes. |
| Conclusion | As a result of the limitations noted in this table (above), it is suggested that the article and its outcomes are subject to a significant bias, which implies that the outcomes are not a clear indication of the intervention and may be coloured by several factors (accounted for and unaccounted for). This is also reflected in the reviewers' assessment of the study (1/10), which indicates that the manner in which the study was presented lacked clarity to allow understanding of the methodology and intervention. This may have been because the prior study was not available for review resulting in the review being harsh both from an internal and external vantage point. This therefore renders the degree of evidence that this article provides as at best limited and worst no evidence is provided in the publication of this article.<br><br>Given these concerns, it is important that the study is repeated to overcome the limitations and short comings of this study. This should be in the form of an RCT or a highly structured n-RCT addressing natural history comparison group, clearly described sample, power analysis and clear description of the intervention programme. This would facilitate the ability of field practitioners to identify which RA patients would benefit most. |

Table 4.50a: Tabulated feedback of reviewer data

| Author/s | Law, R. J., Saynor, Z.L., Gabbitas, J., Jones, J., Kraus, A., Breslin, A., Maddison, P. J. and Thom, J.M | | | | | |
|---|---|---|---|---|---|---|
| Year | 2015 | | | | | |
| Title | The effects of aerobic and resistance exercise on markers of large joint health in stable rheumatoid arthritis patients: a pilot study. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 1 point (a) | 1 point (a) | 67 |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67 |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100 |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100 |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 1 point (a) | 2 points (a+b) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (b) | 0 points (d) | 0 points (d) | 0 points (d) | 67 |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67 |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (-) | 0 points (-) | 0 points (-) | 67 |
| | Total score | 6/10 | 5/10 | 6/10 | 6/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.50b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Law, R.-J., Saynor, Z.L., Gabbitas, J., Jones, J., Kraus, A., Breslin, A., Maddison, P. J. and Thom, J.M |
| Year | 2015 |
| Title | The effects of aerobic and resistance exercise on markers of large joint health in stable rheumatoid arthritis patients: a pilot study. |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - sCOMP<br>- CF<br>- sCRP<br>- Pain intensity scale<br>- Borg scale (peripheral and central rating of perceived exertion) | Study 1:<br>Baseline and 0, 0.5, 1, 2, 6 and 24 hours post intervention.<br><br>Study 2:<br>Baseline, one-hour post-intervention and weeks 0, 1, 4, and 8. | Study 1: 30-minute intervention with 24-hours assessment.<br><br>Study2: eight-week intervention with assessment during the period. | Study 1: total: 16 patients: eight RA patients and eight healthy matched control patients<br><br>Study 2: nine RA patients.<br>Neither study suffered dropout. | Yes | Study 1 had a healthy matched control group. drawn from local area with no knee symptoms.<br><br>Study 2 did not. | No | 6/10, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 75% |

Study 1 – focuses on the acute effects of aerobic and resistance exercise on cartilage breakdown (sCOMP), synovial inflammation, systemic inflammation (sCRP) and knee joint pain.
Study 2 – investigates the effects of an 8-week combined aerobic and resistance exercise training intervention on these markers of joint health.

| | |
|---|---|
| Limitations | Convenience sampling with participants from the local hospital's outpatient Rheumatology Clinic. This suggests that the participants were more likely to be actively seeking advanced / specialist care, implying that their RA disease activity and severity may be more severe than in the average RA patient (Wells *et al.* 2012). The healthy control participants were recruited by word-of-mouth and poster advertising within the same community. As acknowledged by the authors, the sample has limited generalisability because, unlike the wide variety of presentations of RA, this study comprised only of participants with well controlled RA disease. The inclusion criteria required the RA participants to have a diagnosis (ARA 1987 revised criteria (Arnett *et al.* 1988)), along with clearance from their treating rheumatologist. Participants with medical conditions that placed them under significant risk relating to the intervention were excluded (e.g., cardio-pulmonary, and renal conditions and metabolic or gastrointestinal conditions considered to be uncontrolled). In addition to these baseline requirements:<br><br>- Study 1 required RA participants to have symptomatic knee complaints 12 months preceding the study, while study 2 did not. Participants in the healthy control utilised in study 1, matched by age and gender and only differing in BMI (24.8(SD4.1) and 27.4 (SD3.8) and MHAQ (0.6(SD0.4) and 0.1(SD0.1). The study approached 21 RA participants and 11 healthy controls for study 1. Of these, eight RA and eight healthy controls participated at baseline, all of whom completed the programme. In terms of disease characteristics in study 1, the RA participants were stable in terms of treatment with low disease activity.<br>- For study 2 the authors approached 35 RA participants, of which 11 were included in the study with nine completing the programme. In terms of the medication status of the participants in study 2, all but two of the nine were stable.<br><br>In terms of the interventions (Heine *et al.* 2012):<br><br>- Study 1 had two intervention programmes (a 30-minute high-intensity low- impact aerobic and a lower body resistance exercise, one week apart). Two randomised groups undertook a crossover design with a one-week washout period. Prior to the start of the experiment a standard instruction set was offered. The intervention included a familiarisation visit followed by a 30-minute aerobic exercise session (an interval walking exercise), then a rest period, followed by a 30-minute resistance exercise session (a lower body resistance exercise). For the aerobic intervention, the protocol by Ebbeling *et al.* (1991), was followed, to maintain and record heart rate (HR) and RPE (defined as 70-90% and 40-50% of max HR respectively). For the resistance intervention, participants carried out eight maximum repetitions (RM) of leg curls, extensions and presses (Whaley *et al.* 2006), from which 80% load was calculated (Brzycki 1993). The interventions were preceded and followed by five-minutes of low-intensity walking as a warmup / cooldown.<br>- Study 2's exercise programme consisted of progressive resistance and walking over an eight-week period. Familiarisation occurred no more than two days before the start of the eight-week intervention. These supervised exercise sessions (an hour per session three times a week, minimum of 48 hours rest between sessions) followed |

| | |
|---|---|
| | a predetermined protocol. A warm-up and cool-down were included (comprising walking as in study 1 and flexibility exercises for the lower body). The aerobic part of the intervention consisted of half an hour of interval treadmill walking and an equal duration of resistance exercises for the lower limb (as in study 1). Appropriate modification to the load followed developmental milestones. This intervention was sufficient in terms of load and duration and had a sufficient adherence to induce physiological change. Adherence was reported as a mean of 21.8 out of 24 sessions.<br><br>Baseline assessments:<br><br>- For study 1 each assessment had blood samples were drawn (for CRP and sCOMP analysis) and bilateral knee ultrasound was done (CF). The protocols used in these outcomes are detailed in the publication and ensured assessor blinding. To ensure intra-rater reliability all the imaging of five RA participant and five control participants was analysed by another researcher and a radiologist with musculoskeletal sonography expertise was consulted.<br>- Study 2 utilised the same primary and secondary outcomes and included additional outcomes (aerobic capacity, physical function, and maximum strength). Blood sample and CF assessments were preceded by half an hour rest period. In addition, VO2 max, leg press, extension and curl 8-RM were established, and two functional lower body tests were completed (eight-foot-up-and-go and the 30-second sit-to-stand) (Rikli et al., 2001). Assessments were undertaken at a standardised one-hour-post-exercise time at weeks one, four and eight for blood and ultrasound outcomes.<br><br>The lack of power analysis implies that it is likely that this study is underpowered study (Kendall 2003; Mouton 2006; Christley 2010). In addition, the reporting of missing data indicates that intention-to-treat analysis was not utilised (de Boer *et al.* 2015; Unbiased Research 2017). Furthermore, the study sample size limited the study to non-parametric statistics, thus leading conclusions to report trends at best (Hoskin 2012; Bahreini *et al.* 2020). |
| Outcome | The study shows that well-controlled RA patients had higher sCOMP and CF on average when compared with a health control. When participating in an exercise-based intervention (aerobic exercise or resistance exercise or both), no significant change was noted in the RA group from baseline to post-intervention assessment times. These findings suggest that no acute effect was observed on the knee joints of RA participants following exercise. This was also reported in the eight-week intervention study, which also report positive effects, in particular physical fitness, and daily activity function (Law *et al.* 2015). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though the sample size was very small and would have required the use of non-parametric statistics, thus only allowing for trends to be observed. This limitation, along with the vague inclusion criteria, suggests that the outcomes may have been affected by factors such as natural history of the condition / chance – because of the low participant numbers not excluding the possibility of outliers and or patient variance in RA presentation. This negatively affects the study's ability to compare the two intervention groups accurately and fairly, thus resulting in a biased outcome. Furthermore, the type of intervention programmes and the gross over procedure also provided a base for the introduction of patient effects (regarding novelty, naivety, and past experience with the intervention) that are likely to have affected participants' compliance and responses to the outcome measures. |
| Conclusion: | The review of the study by the reviewers indicates that the study lacks the necessary clarity in its presentation to allow for an accurate and clear understanding of the methodology through the attainment of 6/10 on the NOS review tool. This moderate rating indicates that the study's internal validity or methodological rigour is moderately influenced by bias. This is not dissimilar to the review of the external validity bias's present in the study as noted above in this table.<br><br>At best, therefore, it is concluded that this study provides moderate evidence for the exercise-based interventions tested in the programme, which is not associated with patient Ra aggravation It is therefore recommended (as the authors acknowledge) that this study be re-done with a larger sample size, without crossover procedure and with clearer inclusion criteria (incorporating a wider range of RA disease status), along with more clinical outcome measures so as to limit bias and indicate whether actual treatment effects are better than natural history and / or chance alone in either or both of the intervention groups. |

Table 4.51a: Tabulated feedback of reviewer data

| Author/s | Lowman, E.W. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1958 | | | | | |
| Title | Rehabilitation of the Rheumatoid Cripple: A Five-Year Study | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 0 points (b) | 1 point (a) | 0 points (b) | 0 points (b) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 0 points (-) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (d) | 0 points (d) | 0 points (e) | 0 points (d) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (b) | 1 point (a) | 0 points (b) | 67% |
| | Total score | 2/10 | 3/10 | 3/10 | 2/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.51b: Properties of the study, outcome, and discussion

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Author/s | Lowman, E.W. | | | | | | | |
| Year | 1958 | | | | | | | |
| Title | Rehabilitation of the Rheumatoid Cripple: A Five-Year Study | | | | | | | |
| Study properties: | | | | | | | | |
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| ADL | Baseline, post establishment of medication, pre/post physical therapy / rehab, and reportedly carried out several follow-ups throughout the remained of the five years, but frequency and timing are not indicated | Five-years | 38 patients were included in the initially study.<br><br>"Arbitrarily" divided into 2 groups: Group 1 (n = 17) – severely disabled and Group 2 (n = 21) – less disabled.<br><br>Drop out of 8 per group dropped out (Group one (47% drop-out) and Group two (38% drop-out)) | No mention | No, two intervention groups | No | 2/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |
| Limitations | Participants were recruited from New York as the study was conducted at the Bellevue Medical Centre of the New York University. This opens the population to a general RA population and not only more severe RA patients (Wells *et al.* 2012).<br><br>A self-selected sample of 38 participants with active disease and disability RA, who perceived the need for medical attention (medical, physical, and vocational rehabilitation), represented the simple inclusion criteria. This implies that the study did not account for the majority of RA disease dynamic and patient demographic characteristics (Kendall 2003; England *et al.* 2019; Myasoedova *et al.* 2020), introducing variance, outliers and the high likelihood of type II error in the outcomes and conclusions (Bartley and Fillingim 2013; de Boer *et al.* 2015). In addition, participants where "arbitrarily divided" into groups based on disability severity. This system was based on a score of >40% on the ADL measure after being medicated (group one of 17 severely disabled patients) and conversely a score of <40% on the ADL measure after being medicated (group two of 21 moderately disabled participants unable to carrying out jobs / housework).<br><br>The demographic description of patients included 59% females, with an average age of 46 years, an average 13-year disease duration (group one) and 62% females, with an average age of 40 years, an average 8-year disease duration (group two). This reflects that the disability allocation was achieved. A drop-out of eight per patients' group was noted for reasons of continued hospitalisation, death, and home-based confinement. This implies the results are from "cherry-picking" (Mayo-Wilson *et al.* 2017) as the worst of the participants self-excluded allowing only reporting of responses from less severe cases. Thus, only 22 participants completed the study through the follow-up period, nine from group-one and 13 from group-two. These concerns link to the lack of a power-analysis, compensation for drop-out and the intention-to-treat analysis, in order to optimise outcomes and support appropriate statistical analysis (Fogg and Gross 2000; Mouton 2006; de Boer *et al.* 2015; Unbiased Research 2017).<br><br>The five-year study consisted of a multidisciplinary programme (dynamic programme of medical therapy combined with intensive rehabilitations), with the initial two-year intervention followed by a follow-up period. For the first two years, all participants were hospitalised and placed under the care of multidisciplinary team members of the team varied per patient as needed in the study. After the participant had gained a perceived maximum effect of rehabilitation in the in-patient setting, they were then discharged. The description of each component of the programme went unreported (e.g., frequency, intensity, combination), which may be secondary to the individualisation for each participant being too onerous to report (Heine *et al.* 2012). This makes it impossible for the reader to understand how patient factors of naivety, novelty, blinding in addition to the effects of Hawthorn, touch therapy and others impact on the responses of patients to the questionnaire on "Activities of Daily Living" (ADLs) (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015; Senderovich *et al.* 2016). This is critical as the outcomes are only ADLs which are subjective outcomes and heavily reliant on patient response. Further and given the use of the (ADL) sheet as an only outcome measure, it is surprising that neither reliability nor validity of the tool were cited or discussed in the publication, other than to state that it was chosen based on common use in practice (Coster 2013; Bahreini *et al.* 2020). | | | | | | | |

| | |
|---|---|
| | Statistically, the participants were assessed at baseline, after establishment of medical therapy but before physical rehabilitation was commenced. Then after the rehabilitation intervention the patient was re-assessed with multiple follow-up assessments over the five-year study period (viz., both in hospital and after discharge). The frequency and timing of post-in hospital intervention and during discharge assessment is not clear, thus the reader cannot assess whether the outcomes are immediate, short term and / or long term. Regarding these assessments, the procedures, blinding of the assessor and environment where not detailed in anyway; so their effects cannot be delineated. This stands in contrast to the detailed report of findings, which show that improvement/change by category (Pater *et al.* 1998; Kendall 2003; Richardson 2007; Moustgaard *et al.* 2014).<br><br>From an analysis vantage point, the study provides no clearly described method of statistical analysis. It appears that the study reports average change (in percentage form) scores on the ADL outcomes. There is no MICD values comparison and neither have p values been generated, so that the reader can contextualise the outcome in terms of clinical and statistical significance respectively (Bordens and Abbott 2002; Bahreini *et al.* 2020). Additionally, some reported data contradicts the initial layout of the study. An example of this is that the initial period of in hospital care was two years, but on reporting the average hospitalization in days for group one was 351 days, and group two was 241 days. Perhaps this incongruence is related to the manner in which the study was presented, but it does raise concerns around the attention to detail in the programme as well as in reporting the programme (Kendall 2003). This is a significant negative for the study. The irony of the previous statement is reflected in the following statement extracted from the study, where the author state that "In a study of this type many observations are of importance: long-term steroid therapy effects, the deterioration accompanying a progressive rheumatoid disease process, the deleterious intra-articular effects of activities of daily living, the value of corrective orthopaedic surgical procedures, the influence of psychologic factors on a rehabilitation goal, the value of close follow-ups medically as well as socially, and the social and vocational factors modifying attainment of goals. These and many other aspects of the rehabilitation are of importance but cannot be discussed in this report". Given this statement and acknowledgement of the author, it leaves the reader wondering why this was not controlled for in this study. |
| Outcome | With the initial multidisciplinary in-patient intervention (including medical, physical, and occupational therapy, rehab, and psychological and social interventions), the RA participants were able to make improvements in their ADLS or self-efficiency.<br><br>Of the 38 enrolled participants, 17 showed severe disease induced disability, while 21 showed moderate disease induced disability. Group one showed improvement in ADL scores from baseline after stable medical intervention was achieved (from 51 to 47%), followed by improvement at discharge following the extensive multidisciplinary intervention programme (improved to 15%). At the one-year follow-up these findings were maintained, but at the three-year follow-up showed regression (21%). Six of the initial 17 in group-one reported complete self-sufficiency during the intervention with five maintaining this at the three-year follow-up, while five achieved partial self-sufficiency with four maintaining this. Six required custodial care despite intervention. By contrast, group two showed an improvement in ADL scores from baseline after stable medical intervention was achieved for those that completed (from 28 to 13%), post-intervention participants average additional improvement to 4%. This was maintained at the one-year follow-up. Looking at the total sample, 17 obtained self-sufficiencies during the intervention programme, with 12 maintaining this status. Four achieved partial self-sufficiency, only one maintained this (Lowman 1958). |
| Discussion | The author indicates that the outcomes of this study were favourable, even though the sample size was very small. The presentation of only percentages suggests that no statistical analysis was performed.<br><br>The lack of clear inclusion and control of cofounding variables, along with the use of only a subjective outcome measure without control for patient related impact factors presents some significant shortcomings of this study. Further, the lack of a power-analysis, compensation for dropout and the intention-to-treat analysis, in order to optimise outcomes and support appropriate statistical analysis undermines the study significantly in the context of the fact that seemingly no statistical analysis was performed. This therefore does not allow for the exclusion / control of many variables that could lead to a type II error in hypothesis acceptance or non-acceptance. These controls currently do not exist in the study and thus the outcomes are nothing better than chance.<br><br>Additionally, a comparison of the groups is in appropriate as the magnitude (duration and severity) of the RA was different and thus comparative outcomes have not value. |
| Conclusion: | The analysis of the external validity of the study suggests that there is significant bias that has been introduced, such that there are very high levels of bias. This is reflected in the analysis by the reviewers, where the internal rigour was rated at best 3/10, indicating that there is a significant level of bias. Thus, the level of evidence that this study provides no evidence in support of the unpublished protocol of intervention as represented by this study. Thus, there is no clinical benefit for the patients and the information cannot be translated into a clinical scenario. This study should be repeated but given the lack of particular information this may be impossible, so at best studies approximating this study should be considered in the future. |

Table 4.52a: Tabulated feedback of reviewer data

| Author/s | Noreau, L., Martineau, H., Roy, L. and Belzile, M. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1995 | | | | | |
| Title | Effect of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 0 points (b) | 1 point (a) | 1 point (a) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (-) | 0 points (b) | 0 points (b) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (a) | 0 points (d) | 0 points (d) | 0 points (d) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (-) | 0 points (b) | 0 points (b) | 67% |
| | Total score | 5/10 | 3/10 | 4/10 | 4/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.52b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Noreau, L., Martineau, H., Roy, L. and Belzile, M. |
| Year | 1995 |
| Title | Effect of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - AIMS<br>- POMS<br>- Simple index of physical activity<br>- Clinical examination (risk analysis, medication use, joint pain and swelling)<br>- Dynamic strength (knee) peak torque (Nm) and total work (joule)<br>- Haemoglobin (Hb) and ESR<br>- 50-foot walk-test<br>- Body mass and height<br>- Blood pressure and resting ECG<br>- Peak $VO_2$<br>- Max power outlet<br>- Exhale analysis ($O_2$ and $CO_2$)<br>- Stress ECG and blood pressure | Baseline, post-intervention (week-12) and follow-up (week-36) | Intervention 12-weeks<br>Follow-up assessment six-months after the end of the 12-week exercise programme. | 29 total (19 intervention and 10 control). The study makes no effort to declare exact completers or dropout numbers. | No | Yes | No | 4/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | By means of convenience sampling of 29 individuals were drawn from the Quebec City metropolitan area, this implies that the sample population is more likely to reflect the general RA population as opposed to samples drawn from specialist hospitals and outpatient clinics. (Wells *et al.* 2012). However, the convenience sampling does limit generalisability of results as representativeness is often assumed and may not be present (Jager, Putnick and Bornstein 2017). The patients met the following criteria: a history of RA (unknown diagnostic criteria), functional classes I or II (no clear classification tool); free of unstable cardiopulmonary disease; without symptoms of an acute joint/s and being capable of graded exercise test(s) (bicycle ergometer). The study did not report exclusion criteria. These criteria lacked control of a number of confounding variables (Kendall 2003; Wells *et al.* 2012), which included RA serology, RA disease duration and disease activity, age, pain severity, imaging studies, medication, and other physical modalities in use that other studies have variably considered (England *et al.* 2019; Myasoedova *et al.* 2020).<br><br>Medication (important in RA studies) is not controlled in the criteria, but the study recorded medication at each of the assessment periods. Medication changes occurred in only a few participants during the 12-week intervention period. Medication used included intra-articular steroid injections in 17 experimental and all control participants; whereas cortisone was increased for one participant in the experimental group, without change to anyone else. NSAID use remained consistent in 16 experimental and eight control participants, one experimental and two control participants recorded a decrease in use while the remaining two experimental participants reported increase in use during the intervention periods. Finally, “Remittive agent” use remained unchanged in 17 experimental and 8 control participants. While two experimental and one control reported reduction in use, and one control reported an increase.<br><br>These partly and unconsidered variables have the likelihood to influence outcomes, be a cause of type II errors in analysis and deters the study from being able to apply outcomes to specific RA populations (Kendall 2003). Additionally, the criteria should have addressed contraindications to the interventions and assessments to protect against participants from failing to complete assessments and thus results in missing data (Kendall 2003). The sample was non-randomly allocated to one of two groups, raises risk of bias in the groups under study, including non-comparability of RA groups; control of patient naivety to intervention (Moustgaard *et al.* 2014); the management of patient (de Boer *et al.* 2015) and failure of concealment (researcher bias) (Mouton 2006). The groups - an exercise (n=19) group and a control (n=10), were not significantly different in terms of age; RA disease duration; joints count/s; aerobic power; functional class; sex; and medication use (limited to articular injections, cortisone, NSAIDS and “remitive agents”). This assists in addressing some concerns noted at the beginning of this paragraph, which is important given the sample size (N=29). |

| | |
|---|---|
| | Statistically the small sample suggests that the study is underpowered (acknowledged by the authors); which implies that the outcomes may not be better than chance alone, particularly as there were dropouts (albeit low). These small numbers mitigate against the sample being able to overcome the effects related to Hawthorne, intervention novelty, patient naivety and assessor / patient blinding (Kendall 2003; Mouton 2006; Moustgaard *et al.* 2014).<br><br>The supervised exercise protocol continued twice a week for 12 weeks. This was sufficient time for physiological change due to exercise and thus be measurable (Bell 1998; Fraser 2008; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). The exercise component included 15 to 30 minutes of aerobic dancing with heart rate (calculated with Karvonen formula) at 50% of maximum for weeks first three-weeks; 70% of maximum for the remainder of the 12 weeks. The aerobic dance was preceded and followed by warm-up (10 minutes - undetailed) and cool-down (10 minutes – flexibility and stretching exercises). However, this programme lacked detail regarding specific exercises and their load, repetitions, progression, which prevents study replication and clinical interpretation and application of the programme (Heine *et al.* 2012). By contrast the educational psychological counselling and discussion (individual and group) sessions (weekly), focussed on providing material and address areas of difficulty, whilst recording experiences and share coping mechanisms. With the various possible combinations of the interventions, the outcome only relates to the programme as a whole and not a single treatment intervention/s (Robins, Hernán and Siebert 2004). Additionally, the agonistic or synergistic effects of various intervention protocols used in this study are not determinable and therefore their impact on the study cannot be accounted for (Robins, Hernán and Siebert 2004). The control intervention is unclear and undefined; thus, it is impossible to interpret any data comparisons and may be the reason for significant improvements in the control group. Then, in terms of adherence, the exercise group reported an 83% participation (20 out of 24 sessions). Even with this report, the study stated that exercise group had most subjects not remaining physically in the post 12-week exercise period. This seems contradictory and potentially misleading.<br><br>Balanced outcomes measures are a strength of the study (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), for immediate, short, and possibly intermediate term, no longer term outcomes can be concluded. The subjective and objective outcomes seemed valid and reliable however not always contextually applicable to RA (McNair, Lorr and Droppleman 1971b; Meehan RF 1980; Godin and Shephard 1985; Harkcom *et al.* 1985; Liang 2000; Coster 2013; Bahreini *et al.* 2020). Because of the disparity between the two study groups in terms of the findings as suggested in the previous paragraph, there is an indication that there are uncontrolled variables that are impacting on the outcomes and resulting in changes that are not anticipated given the description of the interventions and the manner of testing. This is possibly also complicated by the lack of a reported power analysis resulting in an underpowered study, that cannot control adequately for extraneous variable effects. This is further complicated by the drop out of participants. |
| Outcome | Improved depression, fatigue, anxiety, tension (all $0.06 \leq p \leq 0.08$), vigour ($p \leq 0.01$) was noted, along with maximum aerobic power, maximum workload, maximum heart rate, 50-foot walking time ($p \leq 0.01$) and peak torque and total work of knee flexion of a ($p \leq 0.05$). Non-significant changes in joint status, despite improvements of joint pain was associated with mobility and household activity improvement (all $0.04 \leq p \leq 0.11$). Thus, the authors state that the evidence suggests favour for aerobic exercise in the management of RA patients. The programme did not result in any adverse events (short term), long term is beyond the scope of the study. The medium-term outcomes suggested limited change at the 36 week follow-up as many outcomes like $VO_2$max and mood states had tended to return to baseline values (Noreau *et al.* 1995). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable, even though the study was underpowered and could only report non-parametric calculations. This limitation, along with lack of variable control, accounting for participant factors, inexplicit intervention versus the control group and the unclear period after the intervention ceased, suggests that the outcomes may not be better than chance alone. This reduces the study's ability to compare intervention group accurately and fairly to the control without risking bias in outcomes and conclusions. This is further compounded in that the outcomes may not be generalisability and implementable into clinical practice. |
| Conclusion: | Given the limitations provided in the description and discussion of the study above, the external validity is highly biased, and the outcomes may not be directly due to the intervention. This is equally reflected in a poor rating by the reviewers with a 75% agreement, which indicates that there is methodological rigour and a high degree of bias.Thus, at best this study provides no evidence for dance-based aerobic weight-bearing exercise programme/s for the support of its use in RA patients. It is thus recommended that the study be repeated with a larger sample (determined by power-analysis) with a clearer description of participants (inclusion criteria), control of variables and appropriate parity between the intervention groups, in order to provide a basis to determine the true causal link and application in the clinical setting with prognostic capacity. |

Table 4.53a: Tabulated feedback of reviewer data

| Author/s | Oh, H. and Seo, W. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2003 | | | | | |
| Title | Decreasing pain and depression in a health promotion program for people with rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 0 points (b) | 0 points (b) | 1 point (a) | 0 points (b) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 0 points (-) | 0 points (-) | 0 points (-) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 2 points (a+b) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (d) | 0 points (d) | 0 points (d) | 0 points (d) | 100% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 1 point (a) | 0 points (-) | 0 points (-) | 0 points (-) | 67% |
| | Total score | 3/10 | 4/10 | 4/10 | 3/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.53b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Oh, H. and Seo, W. |
| Year | 2003 |
| Title | Decreasing pain and depression in a health promotion program for people with rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Degree of pain<br>Depression<br>Functional disability<br>Pain management<br>Exercise<br>Coping skills | Pre / post | Intervention one week<br><br>Total duration seven weeks. | Total sample (N = 36): the intervention group (n = 18), non-randomly selected among patients who did not apply to participate in this programme (n = 18). | No | Non-random matched control. | No | 3/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | The 36 participants comprising the sample were recruited from RA outpatients who regularly visited the RA clinic at a University Hospital in South Korea. The recruitment was advertised and thus patients came at their convenience for screening and possible inclusion (Wells *et al.* 2012; Jager, Putnick and Bornstein 2017). The only requirement for inclusion was that participants met the American Rheumatoid Association criteria as assessed by a rheumatologist. The selected participants were non-randomly matched which allowed the control to be matched with the intervention group at baseline in terms of demographic and socioeconomic factors, including sex, age, marital status, income, education, job status and RA characteristics, such as duration of illness and the year of appearance of clinical symptoms (Pocock *et al.* 2002; de Boer *et al.* 2015). The homogeneity of the patients decreases the bias from this source and improves the ability to measure the effects of the CHPPRA more directly. The authors, however, noted that with this underpowered study (Mouton 2006; Christley 2010) the findings may not be generalized. In addition, the use of two instructors could have introduced further bias (Kendall 2003), but by contrast the CHPPRA teaching methods improved skill compliance, thus reducing patient variability and bias. The appropriateness for use in clinical practice, would require that the CHPPRA be more fully detailed (Heine *et al.* 2012).<br><br>The interventions included exercise strategies within the Comprehensive Health Promotion Programme for Rheumatoid Arthritis (CHPPRA) teaching methods group discussion, lectures, demonstrations, contracts, diaries, role-playing and weekly feedback). Patient naivety to these is not reported on (Mouton 2006). It is also noted that there is a tension between participants in the comparison group that were not exposed to the intervention (nor had they participated in any health promotion programme previously), but they expressed a desire to participate in the CHPPRA programme, which could have resulted in under reporting and thus patient effects (e.g., Hawthorne in the reverse), thus influencing the outcomes of this study. A noted concern in the study was that the programme schedule used for practicing exercise strategies may have had an impact on the outcomes because of the timing of the assessment of clinical outcomes (Pater *et al.* 1998). Therefore, the authors state that "the influence of exercise strategies on functional ability might have been evaluated without providing sufficient time for these strategies to be employed", even though outcome measures for both groups were measured at the same time (Kendall 2003). Compounding this, the authors also note that the intervention did not influence functional ability. This could be secondary to the seven-week duration of the study being insufficient time to consolidate the gains of exercise (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). |
| Outcome | According to the study's findings, participants in CHPPRA experienced significant improvement in group difference in favour of the CHPPRA group pain ($p$<0.05) and depression ($p$< 0.05) but did not show improvements in functional disability ($p$=0.84), while pain management ($p$<0.00) and psychosocial coping skills ($p$<0.03) were significantly improved in the intervention group, but exercise was not significantly altered as a result of participating in CHPPRA (Oh and Seo 2003). |
| Discussion | The study lacks control of several limitations, thus placing the study hostage to its own shortcomings, viz., the small sample size, the potential underpowering, lack of sufficient exposure particularly to the practical application of exercise in the intervention programme, the lack of a follow-up period, and the sample conveniently recruited from a specialist centre, all of which combine to suggest that the outcomes may actually be a result of chance rather than a result of the interventions. This is compounded by lack of clarity as to the contents of each component of the study and the limited generalisability of the study. Thus, the study offers limited evidence for the given focus of this systematic review. |
| Conclusion: | Given the significant limitations noted above, from the external validity vantage point, this study suffers from a high degree of bias and the reader cannot exclude the effect of an external influence/s, let alone the lack of change in outcome in relation to the study focus. These limitation line up with the review of the internal validity and methodological rigour via the PEDro scale offered by the reviewers, reflected a combined ranking of 3/10, expressed in agreement of 75%. This ranking infers, a low methodological rigour and subject to high risk of bias within the internal validity of the study. Therefore, this study offers this systematic reive no evidence and until further research can be carried out, this study and its intervention are not clinically applicable for management of RA. |

Table 4.54a: Tabulated feedback of reviewer data

| Author/s | Romanowski, M., Straburzyńska-Lupa, A., Romanowska, A. and Lorenc, R. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2014 | | | | | |
| Title | AB1125: A comparison of the effects of kinesiotherapy and post-isometric relaxation on knee pain in patients with rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 0 points (c) | 0 points (-) | 0% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (c) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (-) | 0 points (-) | 0 points (-) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 0 points (-) | 1 point (a) | 0 points (-) | 0 points (-) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (d) | 0 points (e) | 0 points (e) | 0 points (e) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (c) | 0 points (-) | 0 points (c) | 0 points (c) | 67% |
| | Total score | 2/10 | 3/10 | 1/10 | 2/10 | |
| | | | | | Overall percentage agreement | 59% |

Table 4.54b: Properties of the study, outcome, and discussion

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Author/s | Romanowski, M., Straburzyńska-Lupa, A., Romanowska, A. and Lorenc, R. | | | | | | | |
| Year | 2014 | | | | | | | |
| Title | AB1125: A comparison of the effects of kinesiotherapy and post-isometric relaxation on knee pain in patients with rheumatoid arthritis. | | | | | | | |
| Study properties: | | | | | | | | |
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Knee pain (VAS) | Pre/post | Intervention lasted 10 days | Total of 38 participants, with each group containing 19 participants. No mention of power analysis, dropout rate nor intention-to-treat analysis | No | No | No | 2/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 59% |
| Limitations | The inclusion of this abstract was enforced by the systematic review scope, efforts were made to gain access to the full study. Unfortunately, the author reported none existed. As a result, according to the principles of analysis, this study has significant and vast underreported areas of internal and external validity. What is evident, has been analysed below. For participants to be included in this study, the patients were required to be admitted to a hospital, implying likely recruitment of a subpopulation of RA – poorly representative of the general RA population, influencing generalisability (Wells *et al.* 2012), and be diagnosed with RA (no specified criteria cited). The specifics of the diagnosis and the exact requirements in terms of the RA dynamic characteristics were not specified (Kendall 2003; England *et al.* 2019; Myasoedova *et al.* 2020). From the outline of the participants provided, there did not seem to be a randomised allocation to the groups within the study and it is unclear whether (Mouton 2006), with the exception of age, BMI and DAS28 which were reported as insignificant, the groups were actually similar with regard to their baseline characteristics (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015). The participants were allocated either to group one (kinesiotherapy) [n = 19] or group two (post-isometric relaxation) [n= 19]. The participants in the kinesiotherapy group were expected to perform exercise once a day for 30 minutes while those in group two were required to perform the group one exercise regime. Group two in addition received post-isometric relaxation on three identified muscle groups (knee flexors and extensors and plantar flexors of the foot). This brief outline of the protocol did not provide detail to allow for replication of the study (Heine *et al.* 2012). The fact that the total sample size was only 38, as acknowledged by the authors implies the use of non-parametric statistics, which allow for the computation of trends (Hoskin 2012; Bahreini *et al.* 2020), necessary in this case seeing that the underpowered study (Mouton 2006; Christley 2010) would be likely to suffer from the effects of a lack of patient naivety since the participants were RA patients and hospital based and as such were likely to have been influenced by unreported blinding (Mouton 2006), while the use of only subjective outcome measures as well as the possibility of the Hawthorne effect if the practitioners were familiar to the patients in the hospital setting could have introduced the risk of bias (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015). Additionally, the use of exercise therapy would ideally have required a period greater than 10 days for the effects of such interventions to be measured objectively (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). Moreover, the changes in pain could well have been reduced by the increased mechanical stimulation of the exercise and not the actual physiological effects of the exercise on the muscle. The changes in primary outcome (VAS) from pre to post were reported as > 20 points on the VAS which meets the MCID for clinical significance in group two (post-isometric relaxation group) (Bahreini *et al.* 2020). | | | | | | | |
| Outcome | The study suggests that there was a significantly positive statistical effect for these interventions on pain and kinesiotherapy alone ($p$ =0.0016), while the effects kinesiotherapy and post-isometric relaxation as a combined intervention were even more effective ($p$ =0.0000). Furthermore, the combined therapy of kinesiotherapy and post-isometric relaxation reflected a significant statistical improvement over that of kinesiotherapy alone ($p$ =0.0147) but owing to the relatively small groups it was difficult to draw firm conclusions (Romanowski *et al.* 2014). | | | | | | | |
| Discussion | Given the lack of available information necessary to provide an in-depth analysis of the quality and risk to the internal and external validity of the study, the reader (like this systematic review) fails to clearly evaluate the control or lack thereof over confounding variables within the research. The small sample size and the lack of power analysis suggests that the outcome may be underpowered and the result of chance rather than a result of the interventions. This is compounded by the lack of clarity in several important variable baseline group distributions (viz., the lack of a clear patient description, as well as the lack of a description of the protocol of the intervention/s, the participants' potential under-exposure to exercise in the form of kinesiotherapy), which contribute to the risk that the study's results may fail to manifest detectable outcomes as uniform for the purposes of comparing their outcome. The above limitations (acknowledged by the authors) and the areas of methodology that cannot be evaluated represent significant limitations for the generalisability of the study's outcome. | | | | | | | |
| Conclusion: | Given the significant limitations noted above, from the external validity vantage point this study suffers from a high degree of bias and the study cannot offer the reader a full analysis because no full publication of the research is accessible. The reader therefore cannot exclude the effect of external influence/s. This is congruent with the review by the reviewers, which ranks the study as a 2/10. Nevertheless, the internal validity analysis rendered significant inter-review disagreement in that there was consensus in terms of the low total score, but one category had a 59% agreement. Collectively the PEDro scale analysis of the publication reflected low methodological rigour and a high risk of bias to the internal validity of the study. Therefore, for this study to offer any valuable evidence toward RA therapy, it needs to be repeated with the results published in full. At this point no evidence can be drawn from this publication in support of the intervention. | | | | | | | |

Table 4.55a: Tabulated feedback of reviewer data

| Author/s | Rønningen, A. and Kjeken, I. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2008 | | | | | |
| Title | Effect of an intensive hand exercise programme in patients with rheumatoid arthritis | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 0 points (c) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (-) | 0 points (b) | 0 points (b) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 0 points (-) | 2 points (a+b) | 2 points (a+b) | 2 points (a+b) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 1 point (a) | 0 points (d) | 1 point (a) | 1 point (a) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 0 points (b) | 1 point (a) | 1 point (a) | 1 point (a) | 67% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non-respondents described<br>c) Rate different and no designation | 0 points (b) | 0 points (b) | 0 points (b) | 0 points (b) | 100% |
| | Total score | 3/10 | 5/10 | 6/10 | 6/10 | |
| | | | | | Overall percentage agreement | 79% |

Table 4.55b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Rønningen, A. and Kjeken, I. |
| Year | 2008 |
| Title | Effect of an intensive hand exercise programme in patients with rheumatoid arthritis |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Primary outcomes:<br>Grip and Pinch strength<br><u>Secondary outcomes:</u><br>Joint mobility (goniometer)<br>Hand pain (VAS) during 1st evaluation<br>Functional ability<br>GAT<br>MHAQ<br>Function in daily activity<br>SEHF<br>Patient's global assessment:<br>Disease activity (VAS), Pain (VAS), Fatigue (VAS)<br>Patient's experience (Interview (14 weeks)<br>Adherence | Baseline, week two and week 14 | 14 weeks | Total of 60 participants; 30 to conservative exercise programme and 30 to intensive exercise programme.<br><br>Completers: 24 to Conservative exercise programme and 26 to intensive exercise programme. | No | Two intervention groups. Conservative exercise programme and intensive exercise programme. | No | 6/10, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 79% |

| | |
|---|---|
| Limitations | The participants for this relative effectiveness study were recruited from the Rheumatology Department at Lillehammer Hospital for Rheumatic Disease, Norway (Wells *et al.* 2012). All were required to have RA diagnosis by ACR criteria 1987 (Arnett *et al.* 1988) of greater than one year's duration, to be of an age between 18 and 70 years and to be able to communicate in writing and verbally. The only bases for exclusions noted were functional problems related to diseases other than RA, pregnancy, hand surgery six months before or during the study, mental or cognitive deficits which limited feedback and / or attendance at less than five training sessions during the hospital stay. These patient characteristics are broad and increase the likelihood of heterogeneity of sample groups. In addition, the recruitment of patients from one setting is likely to make these patients less naïve to the intervention types available for RA as they will have originated from the same hospital setting or at least potentially have done so. The participants were assigned to group based on order of recruitment, the first 30 participants were allocated to the conservative exercise group (CEP), while the subsequent 30 participants were allocated to the intensive exercise group (IEP). To the study's credit, the medication consumption, type and dosage were recorded both at baseline, where no differences between the groups were noted, and again at follow-up, which assisted in making it possible to regulate physically and statistically for the effect of medication on the clinical outcomes (Smith *et al.* 2013). Despite these concerns the groups were comparable in terms of the demographic data (sex, age, occupation status), clinical characteristics (comorbidities), disease activity, fatigue, pain, and medication consumption. However, they differed in term of disease duration (the conservative therapy group had a significantly longer disease duration) (Smith *et al.* 2013). Furthermore, not all the RA dynamic characteristics that may have been different between the groups were taken into consideration. In terms of functional ability, the two groups were comparable except for the dominant hand extension being significantly larger in the conservative therapy group (see Section 4.4.1.1). The study was designed as a clinical controlled trial but suffered from a lack of randomisation (viz., patients were consecutively assigned first to the conservative exercise programme and then to the intensive exercise programme) (Fogg and Gross 2000; Mouton 2006; Mansournia *et al.* 2017). This bias cannot be excluded and its impact on the outcomes of the study cannot be negated as the manner and system of recruitment as well as the size of the population would have dictated that the more reluctant patients (potentially less severe cases of RA and / or referred by a physician as opposed to a self-referral) would have joined later while more patients lacking access to medical care would have joined sooner (potentially more severe cases of RA and in their own minds in need of treatment). These different perspectives and the lack of naivety may have had an impact on the outcome measure when comparing the groups to each other (Mouton 2006; Moustgaard *et al.* 2014). This may have further been compounded in that neither the patient nor the therapist was blinded in this study, thus allowing both parties to have undue influence over the largely subjective secondary outcome measures (Kendall 2003; Moustgaard *et al.* 2014). |

| | |
|---|---|
| | In terms of the sample size, it was reported that a power analysis was conducted, which required a sample size required of 30 per group (Kendall 2003; Christley 2010). It is unclear whether the 60 participants all completed the study or whether dropouts were replaced but it was reported that an intention-to-treat computation was used to import missing data (de Boer *et al.* 2015; Unbiased Research 2017). Reasons for loss-to-follow-up were reported, with the only clinical variable between completers and non-completers being grip strength in the dominant hand, which was significantly stronger among the completers. Given the small numbers in the study, despite meeting the minimum requirements for parametric statistics the study dropouts resulted in underpowered calculations for the study. However, it is worth noting that the study utilised missing data analysis (intention-to-treat) to salvage the power through analysis of a complete sample. To the authors' credit, however, the study suggests that the statistical analysis was done with control for variables such as disease duration. Clinical data was recorded by four different assessors, but there was no reporting of assessor training in order to standardise the application of measures to ensure consistency (Kendall 2003; Heine *et al.* 2012). The interventions were well described, and the outcomes measures well detailed. In essence, the difference between two groups lay in the increased repetitions with instructions in the respective exercise programmes being organized as a daily group session led by a therapist. At discharge, the patients in the conservative exercise programme were instructed to exercise according to their usual training routines, while the patients in the intensive exercise programme were encouraged to exercise once a day for a minimum of five days a week (each for 14 weeks). Participation was recorded both in hospital and after discharge with a median attendance of training sessions (seven sessions) in both groups and self-reported training sessions (reported at the week 14 follow-up) indicating a non-significant difference ($p$=0.04) between the groups.<br><br>The outcomes may have been affected by the difference between contacts with the therapists (three times a week for the CEP group compared to four times a week for the IEP group). This may have also been the reason why some participants complained about the exercise period being too long in terms of duration per session and over the 14 weeks); those who had more contact with the therapists may have felt that they were obliged to participate. This could possibly have influenced these participants to report greater improvements in order to please the therapist / assessor (Hawthorn effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015)). It was further noted that the study started to adapt the programme to the patients' needs – perhaps in response to these complaints. This, however, then makes the treatment programmes more variable with the result that the outcomes become negatively affected insofar as there is a natural tendency for the outcomes to aggregate to the mean for the collective group and no definitive improvements are reached. |
| Outcome | This study concluded that a two-week exercise-based intervention programme imparted significant but different levels of improvement between the groups and reported this difference as favouring the intensive programme over the conservative programme in respect of the following measures: the pinch strength of the dominant hand ($p$=0.01), while both the hand grip ($p$=0.04) and pinch strength ($p$=0.05) improved significantly in the non-dominant hand. The difference in grip strength improvement in the non-dominant hand was maintained at a 14-week assessment ($p$=0.04) in favour of the IEP group. A trend towards a decrease of pain was reported for the IEP group, with increasing pain reported for the CEP group. Significant inter-group difference was calculated for measures of pain at the two- and 14-week assessments. Regarding joint mobility and function, very few significant differences were reported between the groups. The study indicates that its intensive programme was not only well tolerated but also imparted greater improvement in the hand function of RA patients in the IEP group by comparison with the results for the CEP group's conservative exercise programme (Rønningen and Kjeken 2008). |
| Discussion | The authors of the study indicate that its outcomes reflect significant improvement pain reported in components of hand function and pain. However, as noted in the limitations above, the study suffers several risks to such findings. As a result of these limitations, therefore, the possibility that confounding factors and natural history of the condition / chance / time may have produced the results reflected in the outcomes cannot be excluded. Furthermore, the study salvages the limited sample power through intention-to-treat analysis, raising concerns of misleading data and the risk of its enhancing/inhibiting effect being exacerbated in a small sample such as this one. The statistics used in this study are parametric allowing causation to be reported on. Therefore, the analysis employed with regard to the study's findings is not exempt from regression to the mean. As a result of this, the study's ability to compare the two intervention groups accurately and fairly is negatively affected, thus producing bias in the outcomes. The design failed to provide a natural history/no treatment control, and differences reported in the groups such as disease duration, strength of the dominant hand extension and difference in therapist contact time, among others, provide a base for patient- and therapist-related effects on the outcomes (e.g., around novelty, the lack of naivety, and past experience with the intervention, the Hawthorne effect as well as the risk to therapy consistency owing to the lack of therapist training) to influence the outcomes measured. Therefore, the limitations noted in the table above result in moderate bias from an external validity vantage point, while the review of the study by the external reviewers at a 6/10 ranking (with 79% agreement) indicates that the study presented outcomes that were subject to a moderate risk of compromise to the methodological rigour. Thus, this study provides moderate evidence supporting the tested interventions. |
| Conclusion | Due to the limitations noted in the table above as well as the fairly congruent response in the review of the study by the external reviewers both the external validity and methodological rigour of this study are of moderate strength with a moderate risk of bias at best. Therefore, it is concluded that this study provides relatively moderate evidence for IEP over that of CEP for RA patients but that it is limited to a rating of no evidence for each of the individual interventions as a result of failing to incorporate a natural history control; consequently, it fails as a stand-alone publication to offer evidence for this programme's efficacy. Therefore, efficacy studies for each of the interventions are required in order to establish what contribution they make to the body of evidence supporting specific exercise regimes for RA patients. |

Table 4.56a: Tabulated feedback of reviewer data

| Author/s | Spiegel J.S., Spiegel T.M., Ward N.B., Paulus H.E., Leake B. and Kane R.L. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1986 | | | | | |
| Title | Rehabilitation for rheumatoid arthritis patients: A controlled trial. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 1 point (a) | 0 points (b) | 0 points (b) | 0 points (b) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 1 point (a) | 0 points (-) | 0 points (b) | 0 points (-) | 0% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (e) | 0 points (d) | 0 points (d) | 0 points (d) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 0 points (b) | 1 point (a) | 0 points (b) | 0 points (b) | 67% |
| | Total score | 6/10 | 5/10 | 4/10 | 4/10 | |
| | | | | | Overall percentage agreement | 75% |

Table 4.56b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Spiegel J.S., Spiegel T.M., Ward N.B., Paulus H.E., Leake B. and Kane R.L. |
| Year | 1986 |
| Title | Rehabilitation for rheumatoid arthritis patients: A controlled trial. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| The 50-foot walk test,<br>The grip strength test,<br>ESR,<br>Joint count evaluation.<br>Estimate of global arthritis activity, A5-point deformity rating (44-joints)<br>Self-administered functional and health status questionnaire (Rand Corporation's Health Insurance Experiment)<br>The Arthritis Impact Measurement Scales (AIMS) | Baseline, week-six, and months-six and -12.<br><br>An evaluation of deterioration of functional ability in the three-months prior to entry into the study was required as an inclusion criterion. | One-year | Total of 92 patients, 49 experimental (in-patients) and 43 control (out-patients) participants.<br><br>Loss to follow-up = three. One experimental participant died during follow-up period (cancer) and two controls did not complete the study. | No | Yes | No | 4/10, indicating a low methodological rigour and subject to high risk of bias within the internal validity of the study. | 75% |

| | |
|---|---|
| Limitations | Recruitment was from the areas around the Cedars-Sinai Medical Centre or community clinics. This range of recruitment sites offers access to a variety of patients with variable RA presentations. This is different from recruitment from specialised departments or hospitals. This implies that the sample for this study may be more representative of RA in general (Wells *et al.* 2012). The sample (n=92) was required to have classic or definite RA (Ropes et al., 1958), for >one-year, functional deterioration (i.e., acute joint inflammation >two-weeks), subacute persistent joint inflammation or inactive synovitis with permanent joint deformities >3 three months before the study, moderate pain or difficulty with daily activities and be English proficient. Any other medical concerns or functional limitation sequelae resulted in exclusion.<br><br>Age, sex, basic ethnicity, education, income, co-living status, employment, comorbidities, disease duration, joint deformity, medication use were compared at the baseline. The only differences between the groups were significantly higher joint deformity in the experimental group and significantly higher gold therapy use in the control group. Medication use, type and administration was left unreported and was thus uncontrolled. This study did not employ randomisation to the study groups, which implies that there is greater reliance on the need to have homogeneity between the groups. Most outcomes were similar with the exception of joint deformity, which was significantly different between the groups. This may create a bias to the control group or decreased ability to show comparable differences / similarities between the groups. This is compounded, by differences between the two referral site options with the 49 experimental group patients, admitted to the Rheumatology Rehabilitation Unit of the University of California at Los Angeles Medical Centre, where the remainder were referrals to the unit via community physicians. This suggests that the two RA patient types may have been different (pre-surgical versus long-standing low-grade RA). This bias is further exacerbated, in that the experimental group spent 13 days in hospital accessing multidisciplinary care within a specialised rheumatology rehabilitation clinic to develop individualised management programmes for participants. The range of treatment options, although standard in terms of the type of intervention, prevents replication and generalisability of the study procedure and results respectively. Thus, although pragmatic and more reflective of clinical practice it does imply that the outcomes are really based on a collection of cases (e.g., case series) rather than a group where the same intervention is tested. Further to this and after the hospitalisation period, the patient management continued under the physician who referred them into the study (this thus resulted in variance from this point forward as approximately half of these physicians were UCLA physicians and the remainder community-based physicians).<br><br>This compares with the control group, who continued care under their rheumatologist (these practitioners were aware of modalities of treatment available at the rehabilitation centre, but it was not indicated whether they had the potential to provide the same interventions). This implies that the rheumatologist may / may not have used treatment modalities that were available at the rehabilitation centre, thus placing the control at a disadvantage (i.e., did the control differ from the experimental |

| | |
|---|---|
| | because of the health care environment or the actual treatment). The only positive aspect of the interventions is that once the interventions were set, they did not change for that patient for the duration of the study.<br>In terms of the outcome measures, the objective and clinical outcome measures were administered by two unblinded physicians, with participants maintaining the same assessor. The inter-assessor variability was controlled with the use of the American Rheumatism Association developed protocol (Mainland 1965). Intra-assessor bias based on group allocation was not controlled for (Kendall 2003) or limited but may have been mitigated by the balance of the subjective and objective measures (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018), limiting influence from patient factors. However, to compound assessment, it was not clear that all tools were validated, reliable and contextually appropriate for RA patients (Coster 2013; Bahreini *et al.* 2020).<br><br>In terms of the statistics, the baseline data included a functional ability (five scales) screen at three months prior to the study, at baseline and on termination. It is unclear whether the initial recordings required retrospective (recall) data as this may have introduced bias (Mouton 2006; Ricker, Vergauwe and Cowan 2016). The statistics were also affected by differential access to physicians and intervention plan implementation between the groups. Based on the above inconsistencies of baseline comparison, allocation, interventions, and bias associated with each, it is noted that a number of *post-hoc* adjustments were done during the statistical analysis. This alone is also subject to introducing analysis bias, as it removes the reader from an accurate picture in terms of the clinical outcomes**.**<br><br>Practically in regard interventions post this research, given the location of the experimental group, the study reports in the follow-up periods, that the experimental group had significantly more orthopaedic surgical intervention (given that they had significantly higher joint deformity). Whilst the less severe control group were admitted into hospital for arthritis complaints (but non-surgical). Thus, given the severity of the experimental group the likelihood for increased presence of co-morbid disease is likely the cause for admission to hospital for non-arthritis complaints. These are however only conjectures based on literature suggestion but may also have been linked to the baseline variables not controlled for in this study. |
| Outcome | This study aimed to demonstrate the effect of inpatient multidisciplinary care within a rheumatology rehabilitation unit, compared to continuing care under participants rheumatologist. Forty-four of the participants in the multidisciplinary care were admitted of 13 days on average, while 43 participants continued current private care under their rheumatologist. When controlling statistically for group differences, the study showed significant improvement ($p$<0.05) in functional status, disease activity and mental health measures, as a result of the multidisciplinary care at the one-year assessment compared to the control group (Spiegel *et al.* 1986). |
| Discussion | The authors indicated that the outcomes of this study were favourable for functional status, disease activity measures and mental status, even though there was a lack of homogeneity between groups, a lack of power analysis (potentially under powering the study). These principal concerns were additionally affected by the lack of blinding of assessors and patients and patient related factors were not adequately controlled. These collectively negatively affect the study's ability to accurately and fairly compare intervention group to the control group. Additionally, the lack of clarity on the specific intervention types as well as control over confounding factors limits the authors capacity to evaluate the degree of risk to the study and then too to evaluate the risk of type II error. By contrast had the above concerns been addressed, then the participant description would have provided for a degree of certainty regarding generalisability and implementation into clinical practice. Currently the study findings provide limited clinical guidance for the use of a multidisciplinary intervention that incorporates physical modalities in the treatment of RA patients in improving functional status, disease activity and mental health. |
| Conclusion: | Given the limitations and concerns around the external validity discussed and noted above, there is a high degree of bias and thus the study is expected to have very little ability to be able to translate into clinical practice. This in conjunction with the 4/10 rating by the reviewers for the internal / methodological rigour of the structure of the study, indicates moderate to high bias, means that the collective outcome of the study is one of moderate to high bias resulting in no or at best limited evidence that may be extrapolated to clinical practice. Therefore, the evidence is to be read with significant caution. This study requires to be repeated with a far greater detailed methodology including a published power-analysis and an inclusion of a larger representation of the variable RA presentations, in order to provide a basis to determine the true causal link and application in the clinical setting with prognostic capacity. |

Table 4.57a: Tabulated feedback of reviewer data

| Author/s | Zernicke, J., Kedor, C., Müller, A., Burmester, G.-R., Reißhauer, A. and Feist, E. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2016 | | | | | |
| Title | A prospective pilot study to evaluate an animated home-based physical exercise program as a treatment option for patients with rheumatoid arthritis | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| Selection | Is the case definition adequate?<br>a) Yes, with independent validation *<br>b) Yes, e.g., record linkage or based on self-reports<br>c) No description | 1 point (a) | 0 points (b) | 1 point (a) | 1 point (a) | 67% |
| | Representativeness of the case<br>a) Consecutive or obviously representative series of cases *<br>b) Potential for selection biases or not stated | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67% |
| | Selection of controls<br>a) Community control *<br>b) Hospital control<br>c) No description | 1 point (a) | 0 points (-) | 0 points (-) | 0 points (-) | 67% |
| | Definition of controls<br>a) No history of disease (endpoint) *<br>b) No description of source | 0 points (b) | 0 points (-) | 0 points (-) | 0 points (-) | 67% |
| Comparability | Comparability of cohorts on the basis of design or analysis<br>a) Study controls for _____ (select the most important factor.) *<br>b) Study controls for any additional factor * (This criterion could be modified to indicate specific control for a second important factor.) | 1 point (a) | 2 points (a+b) | 1 point (a) | 1 point (a) | 67% |
| Exposure | Ascertainment of exposure<br>a) Secure record (e.g., surgical records) *<br>b) Structured interview where blind to case/control status *<br>c) Interview not blinded to case/control status<br>d) Written self-report or medical record only<br>e) No description | 0 points (d) | 0 points (d) | 1 point (b+d) | 0 points (d) | 67% |
| | Same method of ascertainment for cases and controls<br>a) Yes *<br>b) No | 1 point (a) | 1 point (a) | 1 point (a) | 1 point (a) | 100% |
| | Non-response rate<br>a) Same rate for both groups *<br>b) Non respondents described<br>c) Rate different and no designation | 1 point (a) | 1 point (a) | 0 points (b) | 1 point (a) | 67% |
| | Total score | 6 /10 | 5/10 | 4/10 | 5/10 | |
| | | | | | Overall percentage agreement | 71% |

Table 4.57b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Zernicke, J., Kedor, C., Müller, A., Burmester, G.-R., Reißhauer, A. and Feist, E. |
| Year | 2016 |
| Title | A prospective pilot study to evaluate an animated home-based physical exercise program as a treatment option for patients with rheumatoid arthritis |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Semi-structured face to face interviews<br><br>Patient reported outcomes:<br>1) physical function (HAQ-DI)<br>2) patient's global assessment (VAS)<br>3) short form 36-Item questionnaire.<br><br>Physical function tests:<br>1) dynamometer (isometric measurement)<br>2) respiratory function test (peak flow meter)<br>3) 6-minute walk test | Baseline, week 12 and 24. | 24 weeks | Total of 30 participants.<br><br>Conventional home-based physical exercise programme (n = 15)<br>Predefined animated exercise programme by using Wii game console (n = 15)<br><br>These groups crossed over at 12 weeks.<br>Drop out of 3 females prior to crossover | No | Two treatment arms one arm was conventional | No | 5/10, indicating a moderate methodological rigour and subject to moderate risk of bias within the internal validity of the study. | 71% |

| | |
|---|---|
| Limitations | The inclusion criteria required compliance with the ACR criteria (Arnett *et al.* 1988; Aletaha *et al.* 2010) in addition to patients achieving a self-reported disease activity <30 mm and biological DMARD therapy (excluding conventional DMARD (cDMARD)). As a result, there was great variance among the patients with regard to sex, age, disease duration, type of biologic DMARDs medication used and history of biologics used and therefore outcomes could not solely be attributed to the intervention. In addition, the lack of possible homogeneity in terms of the disease processes means that the ability to generalise the outcomes to all RA patients is limited (Wells *et al.* 2012). This is further compounded in that the recruitment and allocation of patients based on a convenience method, which did not randomise / stratify or conceal the allocation processes, increasing the risk of researcher bias and compounding group homogeneity concerns. A resultant sample of 30 participants were enrolled into one of two groups, resulting in 15 participants in both groups. Group one started with the conventional home-based physical exercise programme, while group two started with the predefined animated exercise programme by using Wii game console. These groups crossed over during the study.<br><br>This study was a single-centre, crossover trial with two treatment arms. The crossover of patients within the design, provides an effective control. However, it limits the washout periods for treatment effects that may carry over. Thus, outcomes in the second phase of the study may result from a compound effect of the manner in which the two periods of intervention either synergise or antagonise. To counter this effect, a mono-centre standardized instruction for the animated exercise programme was utilised; however, the fact that all participants were from the Department of Rheumatology at Charité University Hospital, Berlin, limits the generalisability of the outcomes and also increases the likelihood of the patients having had previous exposure to the intervention types which would compromise patient naivety to the treatment and therefore their actual versus perceived response (Mouton 2006; Wells *et al.* 2012).<br><br>The intervention plan required participants to attend two to three supervised sessions to rehearse the exercises in the conventional group or to become acquainted with the device in the Wii fit plus group. All patients then received an exercise manual that outlined what participants were required to do, which was to exercise three times a week for 30 minutes at each session. Possible concerns around the home exercise include compliance with the required regimes, the variance in time of the regime, the intensity of the regime and the recording of compliance with the regime. In addition, because there were four groups of exercises (yoga, strengthening, balance, aerobics) the homogeneity of the training programmes between the different groups is unclear.<br><br>The outcomes related to qualitative data collection, where patients were requested to explain their experiences over the 12-week course of the study. This should preferably have been done by means of an exercise diary as memory recall bias (Mouton 2006) is possible in terms of the post-exercise interview process. With regard to the quantitative data, the |

| | |
|---|---|
| | codes as well as the categories were discussed and revised continuously by the study group, which comprised two physicians and one sports scientist in order to increase inter-assessor comprehension and to avoid counterproductive coding by multiple assessors. The measures included HAQ, VAS (patients' global assessment) and SF36 (Busija *et al.* 2011; Maska, Anderson and Michaud 2011), in addition to the use of isometric measurement tools (Hengstman *et al.* 2008) for assessment of neck extensors, neck flexors, shoulder abductors, elbow extensors, elbow flexors, 3-point grip strength, hip flexors, knee extensors and knee flexors. A respiratory function test (peak expiratory flow by peak flow meter) (Lebowitz *et al.* 1982) and six-minute walk test (used to assess the submaximal level of functional performance with a MCID of 20 meters) (ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories 2002; Perera *et al.* 2006; Bennell, Dobson and Hinman 2011) were used as quantitative measures. The baseline measures for many of these outcome measures were not included in the article and therefore the homogeneity of the groups cannot be commented on. Furthermore, these outcomes when reported did show differences between the groups. Therefore, the lack of homogeneity in terms of the RA in addition to the clinical differences among participants at entry into the study further complicated the ability of the study to compare the interventions between the two groups.<br><br>The study fails to make use of power-analysis, and with this would not have the capacity to make provisions for drop-out. With this, the study did not make use of any clear forms of missing data analysis (such as intention-to-treat) given the study did in fact suffer dropouts as previously noted. As a result, the study prevents the reader from contextualizing the study in terms of risk of underpowered calculations, control for both type II error and allowing for outcomes to draw conclusions better than chance (Kendall 2003; Mouton 2006; Christley 2010). |
| Outcome | The authors reported that the animated home-based exercise programme using a Wii game console was feasible and beneficial for RA patients when compared to standard physical home exercises. Similar effects were observed between the groups, thus indicating that such an animated Wii game programme might be an alternative supportive option for RA patients (Zernicke *et al.* 2016). |
| Discussion | The authors of this study indicated that the outcomes of this study were favourable even though the sample size was very small and would have required the use of non-parametric statistics, thus only allowing for trends to be observed. This limitation, along with the vague inclusion criteria, suggests that the outcomes may have been affected by factors such as patient demographics, RA dynamic characterises, medication use / adherence, RA natural history or simply chance. This negatively affects the study's ability to accurately and fairly compare the two intervention groups, thus resulting in a biased outcome. Furthermore, the type of intervention programmes and the cross over procedure also provided a base for the introduction of patient effects (novelty, naivety and prior experience with the intervention (Mouton 2006; Moustgaard *et al.* 2014)) that are likely to have affected patient compliance and thus responses to the outcome measures. |
| Conclusion: | The limitations noted in the table above, together with the ambiguity reflected in the review of the study by the reviewers, indicate that the study presentation lacked the clarity required to allow for accurate and clear understanding of the methodology. This alone introduces bias as the understanding by various readers is thus problematic and the ability to repeat the study and attain the same results becomes compromised. Thus, when reviewing the outcomes, which were not statistically significant between the groups and only trend-based, the suggestion that one programme of invention is better than another is very limited, and it is not possible that the study could provide evidence for only one component of each programme. Therefore, the external validity is subject to high levels of bias, while the internal validity as agreed by the reviewers (71% agreement), indicates that there was average bias (5/10), with a resultant moderate bias. Given this scenario, the moderate to high levels of bias provide this study with little evidence in support of the outcomes for use in clinical practice.<br><br>In view of this limited evidence for the support of the Wii intervention programme for RA patients, together with the inability of the study to provide clear patient descriptions to facilitate field practitioners' ability to identify which RA patients would benefit most, it is recommended that this study be re-done with a larger sample size over a longer period with no crossover procedure and with clearer inclusion criteria for the purposes of limiting bias and indicating that the treatment effects are better than natural history and / or chance alone in either or both of the intervention groups. |

### 4.4.2.3 Case studies/reports and Case series.

Table 4.58: List of table numbers for case studies / series feedback and analysis.

| Tabulated feedback of reviewer data | Properties of the study, outcome, and discussion | Author/s: | Year of publication: | Title of research paper: |
|---|---|---|---|---|
| Table 4.59a | Table 4.59b | Almuhaysin | 2019 | Case Study of Treatment of a Patient after Rheumatoid Arthritis. |
| Table 4.60a | Table 4.60b | Chui, Lau and Yau | 2004 | An outcome evaluation study of the rheumatoid arthritis self-management programme in Hong Kong. |
| Table 4.61a | Table 4.61b | Chung and Mior | 2015 | Use of spinal manipulation in a rheumatoid patient presenting with acute thoracic pain: a case report. |
| Table 4.62a | Table 4.62b | Cubick, Quezada, Schumer, and Davis | 2011 | Sustained release myofascial release as treatment for a patient with complications of rheumatoid arthritis and collagenous colitis: a case report. |
| Table 4.63a | Table 4.63b | Kapale, Vardharajulu, and Warude | 2017 | Effect of Free Exercise and Rheumatoid Arthritis. |
| Table 4.64a | Table 4.64b | Kauppi, Leppänen, Heikkilä, Lahtinen and Kautiainen | 1998 | Active conservative treatment of atlantoaxial subluxation in rheumatoid arthritis. |
| Table 4.65a | Table 4.65b | Lonta | 1960 | Facilitation technics in the treatment of early rheumatoid arthritis. |
| Table 4.66a | Table 4.66b | Neuberger, Press, Lindsley, Hinton, Cagle, Carlson, Scott Dahl and Kramer | 1997 | Effects of exercise on fatigue, aerobic fitness, and disease activity measures in persons with rheumatoid arthritis. |
| Table 4.67a | Table 4.67b | Perlman, Connell, Clark, Robinson, Conlon, Gecht, Caldron and Sinacore | 1990 | Dance-based aerobic exercise for rheumatoid arthritis. |
| Table 4.68a | Table 4.68b | Williams, Brand, Hill, Hunt, and Moran | 2010 | Feasibility and outcomes of a home-based exercise program on improving balance and gait stability in women with lower-limb osteoarthritis or rheumatoid arthritis: a pilot study. |

Table 4.59a: Tabulated feedback of reviewer data

| Author/s | Almuhaysin A. A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2019 | | | | | |
| Title | Case Study of Treatment of a Patient after Rheumatoid Arthritis | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | A | A | A | A | 100% |
| | 2. What percentage of individuals refused to participate? | N/A | N/A | N/A | N/A | 100% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B2 | A | A | A | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | B1 | B1 | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | N/A | N/A | N/A | N/A | 100% |
| | 7. Is the analysis by intention-to-intervene (treat)? | N/A | A | N/A | N/A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B2 | As outlined in Table 4.59b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | A | As outlined in Table 4x.59b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | There are no practical or ethical reason why an RCT could not be conducted. | | | | |
| | 4. Include any other comments | This study was an undergraduate thesis; therefore, the study was not peer reviewed. | | | | |
| | | | | | Overall percentage agreement | 88% |

Table 4.59b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Almuhaysin A. A. |
| Year | 2019 |
| Title | Case Study of Treatment of a Patient after Rheumatoid Arthritis |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| A full examination was done including: A posture examination, palpation of pelvis, gait examination, dynamic spine examination, assessing breathing, chest expansion, two-scales standing, specific testing, anthropometric measurement, spinal distances, muscle length test, muscles strength test, goniometry, neurological examination, soft tissue technique examination, joint play. | Pre/post design | 11 days of intervention | One | No | N/A | N/A | Two A score, one B1 score and five N/A scoring. resulting in mostly fulfilled criteria. | 88% |

| | |
|---|---|
| Limitations | This case study was conducted at the Institute of Rheumatology in Prague over an 11-day period inclusive of 10 treatment sessions.<br><br>The participant was a middle-aged female, smoker (43 years old), weighing 54kg with a resultant BMI of 22.3. Her RA disease duration was 34 years (from age 10). She had had total hip replacement bilaterally with re-operations several years later. Her medical history included depression, breast reduction surgery to alleviate back pain in 2012, a Caesarean in 1999 followed by a hysterectomy in 2009. Regarding examinations, she reported previous radiographs of the chest, hips, left knee and distal joints and endoscopic examinations, both colonoscope and gastroscope, in 2017. Related to the RA she reported progressive small joint pain in her hands and feet, left knee pain and stiffness of a two-year duration, occasional low back pain and dizziness two weeks prior to intervention, left ankle pain that was chronic during walking and, in the preceding year, bilateral shoulder pain with instability. This complex presentation limits the ability of the data developed from this case report to patients of a similar background and or medical complexity and may not be applicable to RA patients with less severe RA and fewer medical co-morbidities and concurrent medical conditions.<br><br>This case study included multiple interventions such as soft tissue techniques, joint play mobilization, stretching techniques, strengthening techniques, relaxation techniques, sensomotor exercises, and general exercises as well as medications, namely, Methotrexate, Vigantolr (Vitamin D) and Acidum Folicum (Folic Acid).<br><br>In the light of the participant's medical history, it is assumed to be unlikely that she would have been naïve to the intervention used in the study. This raises the risk of preconception bias having skewed the results.<br><br>Given that this study is a case report, it means that the outcomes are limited to this patient's clinical improvement. This implies that the outcomes may not be better than chance and that further nRCTs or RCTs would need to be undertaken to confirm that the outcomes achieved are indeed reproducible in similar or different patients presenting with RA. Further research is also necessary in that this case study included not only multiple interventions as listed above and medications, which did not have dosage and duration of prescription detailed. These various factors are potentially better controlled in an RCT or nRCT where bias is excluded more systematically in terms of the agonist or antagonist interactions between the various interventions. Additionally, the patient also undertook swimming two to three times a week, participated in yoga free time and took 30-minute walks four times a week. These additional factors make it impossible to determine the effects of individual interventions and the outcomes of each of these in terms of the patient's clinical improvement is limited due to the higher levels of bias in a case study.<br><br>A strength of this study is that the sessions with the patient were detailed well. on the other hand, it was not clear why the duration of sessions was not consistent. This protocol inconsistency implies that the outcomes may be dependent on the time spent, particularly with regard to exercise as opposed to the actual interventions used. For example, in the case of exercise a measurable outcome may only be seen clinically after six to eight weeks of care; measuring clinical change in 11 days is therefore unlikely to register change resulting from exercise. As such, the impact of the various clinical interventions has not been considered specifically in this context. This, |

| | |
|---|---|
| | along with the detailed report on the patient examination and the tools used in examination (goniometer, measurements tape, neurological hammers, and plumb line) as well as the re-assessment after the 11 days assists in objectively quantifying the clinical outcomes of the patient for those interventions that have specific short-term action.<br><br>The lack of many subjective outcomes contributes to negating the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015), the observer effect and other related effects that are often carried over due to patient naivety (Mouton 2006). From a patient bias perspective, it is reported that the patient was able to perform all the exercises and that therapy had a positive effect but there was no mention that the patient had an exercise diary or a means of consistently reporting the data.<br><br>The communicability of the overall methodology is complicated by the way language is used to present the technical component of this study, which may hamper the reader's ability to fully grasp and confirm all concepts. |
| Outcome | The study concluded that 10 sessions of mixed intervention therapy (involving soft tissue techniques, joint play mobilization, stretching technique, strengthening technique, relaxation technique, sensorimotor exercises and general exercises) and medications), provided the patient improvement in ROM in the affected peripheral joints, decreased pain associated with these joints, and corrected for gait abnormalities (Almuhaysin 2019). |
| Discussion | Given that this study is a case study, its structure is more pragmatic and reflective of clinical practice. Thus, it is not able to discern the outcomes of any one modality in the clinical improvement or regression of the patient with RA. It is, however, possible to indicate that this study does suggest that there may be benefit from a programme which includes the use of soft tissue techniques, joint play mobilization, stretching technique, strengthening technique, relaxation technique, sensomotoric exercises, and general exercises) and medications.<br><br>This does require further study through a process of a randomised controlled or non-randomised trial in order to allow a better measure of the intervention effect with the use of a comparison group for clarity of the interventions effect on outcomes and a greater sample size to enlighten possible outliers in the interventions effect, enabling clinical prognostic effect of the intervention with defined barriers in generalisability (i.e. to determine whether this is greater than chance or natural history alone within the context of this case study). |
| Conclusion: | In conclusion, it is evident from the collective 88% agreement among the reviewers that this study was able either mostly or fully to fulfil the applicable criteria in the Liddle scale (good internal validity for the case study methodology). This, taken together with the limitations impacting on the external validity noted in this table, results in the study being ranked as subject to moderate risk of bias.<br><br>This implies that the outcomes are reproducible with clear methodological detail, but the effects of the outcomes are not conclusively imparted nor are the reported effects necessarily only as a result of the single modalities but rather the treatment programme. Consequently, despite its good internal validity/bias control, the study cannot be seen as free from bias because of the shortfalls in external validity/bias control. Therefore, at best, this particular study provides moderate evidence in support both of the intervention programme utilised. |

Table 4.60a: Tabulated feedback of reviewer data

| Author/s | Chui, Lau & Yau | | | | | |
|---|---|---|---|---|---|---|
| Year | 2004 | | | | | |
| Title | An outcome evaluation study of the Rheumatoid Arthritis Self-Management Programme in Hong Kong | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | B2 | B2 | B2 | B2 | 100% |
| | 2. What percentage of individuals refused to participate? | C | C | C | C | 100% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B1 | A | B1 | B1 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | C | N/A | N/A | 67% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | I | B1 | I | I | 67% |
| | 7. Is the analysis by intention-to-intervene (treat)? | N/A | A | N/A | N/A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | I | N/A | I | I | 67% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B2 | As outlined in Table 4.60b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B1 | As outlined in Table 4.60b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | There seems to be no ethical and practical reason aside from the financial limitations (given the expense of such a study) as to why an RCT could not be conducted. | | | | |
| | 4. Include any other comments | None | | | | |
| | | | | | Overall percentage agreement | 75% |

Table 4.60b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Chui, Lau & Yau |
| Year | 2004 |
| Title | An outcome evaluation study of the Rheumatoid Arthritis Self-Management Programme in Hong Kong |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Self-Management Behaviour (SMB), Arthritis Self-efficacy Scale (ASES), Arthritis Impact Measure Scale (AIMS), Arthritis Helplessness Index (AHI), VAS (Pain and Fatiuge), Use of health care | Pre/post design with additional one- and six-month follow-ups. | One- year | 84 participants started the RASMP intervention with unknown size of grouping for Questionnaire A and B.<br><br>14 dropouts (when and which grouping was not reported)<br><br>Completers (n=70):<br>Questionnaire A (n=38)<br>Questionnaire B (n=32) | No | No | No | The score derived from the Liddle scale include one B1, one B2, one C, two I and three N/A scoring.<br>These rankings result in predominantly unfulfilled or inadequately described to offer scoring for the applicable criteria. | 75% |

| | |
|---|---|
| Limitations | The study was undertaken by "the Hong Kong Society for Rehabilitation". The participants appear to be carried over from participants of a recently conducted RASMP evaluation carried out by Chui *et al.* (1998). The limited detail offered in this publication prevents the reader gaining clarity as to the related recruitment context variables such as specialist care seeking behaviour relating to the type of participant recruited limiting the capacity of the reader to create a perception from the outset exists that the sample consists of patients that have previously sought specialist care, implying that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). The sample is derived from the previously conducted RASMP study by Chui *et al.* (1998). The study describes the demographic of the sample (as a whole), the sample was predominantly female (84.3%), averaging six years disease duration, 67.2% between the ages of 35 and 54 years, 61.4% obtaining an education of secondary school or above, and 93% regularly received medical check-ups. Despite this poorly described group with no indication of the distribution of the participants as two separate groups given that they received different outcomes. The demographic description includes the following 78,6% of the participants in the study (as a whole) were not currently involved in self-help groups. Not only does this imply that outside of the intervention received in this programme 21.4% had additional current and ongoing similar interventions at the same time resulting in absolute confounding. We cannot say that any of the participants were naive to any degree. The baseline values for the outcome measures are reported on, however in several the outcomes the two groups are reported together as a single group (where the outcomes are shares), while others are reported separately (unshared outcomes) preventing the comparison (homogeneity) of this "subgroups". The data is also only shown for the 70 completers of the study. Questionnaire A subgroup had 38 completers and Questionnaire B subgroup had 32 completers, both lacking clarity of the number of participants in each subgroup at the start of the study and how many dropped out.<br><br>The intervention used in this study Rheumatoid Arthritis Self-Management Programme (RASMP) was derived from the Arthritis Self-Management Programme (ASMP) (Arthritis Foundation 1992). The ASMP is a programme of interventions including relaxation exercise, pain and stress coping strategies, joint protection, management of cognitive symptoms, and self-management skills (focused on the effects of the disease). However, the RASMP programme adds to this the management of pain, assistive device use, and techniques tailored for RA needs for joint protection. A previous study by Chui *et al.* (1998) revealed that the RASMP evaluated and showed effect on and the capacity to enhance disease-related knowledge, and the participant's skills. It appears the cited study was utilised in the formative evaluation of the programme and provided data for improvement of the programme.<br><br>The study fails to detail the RASMP intervention, aside from the general components addressed by the patient education programme, the study makes note the nine intervention programmes used over the year. The assumption that the nine programmes were run in parallel is based on the 12-month follow-up and the year study duration, but this is not clearly stated. We do not have a clearly detailed programme (including duration of intervention, environment, supervision, how components of the patient education were implemented, what material was covered in the intervention, among others (Fraser 2008)), in a way that it could be replicated. If we looked up the Chui *et al.* (1998) publication we could possibly comment on the intervention replication ability but not with-in this paper. The lack of this detail raises several questions of potential bias including protocol consistency (Kendall 2003).<br><br>The collection of outcome measures are all subjective self-administered questionnaires. The lack of objective outcomes raising the concerns relating to the data obtained (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Furthermore, the question is raised around performance bias (Moustgaard *et al.* 2014), Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015), Placebo effect (Kaptchuk 1998; Bialosky *et al.* 2011), memory recall and decay (Mouton 2006; Ricker, Vergauwe and Cowan 2016), emotional and personality |

| | |
|---|---|
| | variability, and learning/practice effect (Bordens and Abbott 2002). The outcome measure tools are however aligned with the objectives of the study and able to report a variety of aspects related to RA (Coster 2013). The two questionnaires utilised in this study form the primary outcome measure tools and were constructed from the following: (1) ASES (pain and other symptoms – (Lorig *et al.* 1996)); (2) SMB (exercise, participant-physician communication, management of cognitive symptoms – (Lorig *et al.* 1996)); (3) AIMS (mood and anxiety – (Meenan *et al.* 1992)); (4) AHI (Stein, Wallston and Nicassio 1988); (5) VAS (pain –(Huskisson 1983)); (6) VAS (fatigue – (Huskisson 1983)). Previous studies show reliability and validity of these outcomes (Lorig *et al.* 1985; Goeppinger *et al.* 1989; Lorig and Holman 1993; Lorig *et al.* 1996; Barlow, Turner and Wright 2000), but these studies are on arthritis not RA (thus the use in RA is dubious). This is important as the validity, reliability and suitability of the tools utilised impart influence on the clinical interpretation (Coster 2013; Bahreini *et al.* 2020). From these six outcome tools the study created two questionnaire sets to reduce the length of the assessment outcome. They were then used in assessment periods in different groups of people (i.e., the 84 participants were in essence divided into 2 groups – questionnaire A and questionnaire B (baseline group participant numbers are not detailed)). Questionnaire package A included the following outcome measures (or aspects of outcome tools): ASES, SMB AIM (mood), AHI, VAS (pain). Questionnaire package B included the following outcome measures (or aspects of outcome tools): ASES, SMB, AIM (anxiety and pain), VAS (fatigue). In addition, both "groups" reported on their utilization of health care over the two months preceding the assessment. Given the construction and lack of previous evaluation of the constructed outcome tool for RA patients we cannot offer conclusive commentary on validity, reliability, suitability, sensitivity, and specificity (Coster 2013; Bahreini *et al.* 2020). The splitting of the outcome tools raises questions around the value of the tool when not carried out as a complete tool. An additional concern develops in that splitting the outcome tool into two questionnaires utilised on sperate subgroups of the sample effectively creates two groups that are not identical from which the power of the group size may drop to underpowered levels (Mouton 2006; Christley 2010).<br><br>The nine programmes run for this study fails to report the details of these programmes' environment, timing (related to the intervention-to-assessment), protocol followed in assessment periods, therapist's involvement in the assessment and if they were the outcome assessor as well, are not detailed. Regarding the timing factor, were the nine ran at different times and the data collection staggered or parallel, this uncertainty introduces several risks into the study. The study did not apply a power analysis for any of the outcome measures preventing the reader from determining whether the study was underpowered, particularly for the subgroups and suffering the drop-out (Kendall 2003; Mouton 2006; Christley 2010). The data was reported on only completers of the study. The loss of 14 participants (16.67%) lacks clarity as to which group suffered the largest drop-out. The study failed to report whether they all participated and dropped-out prior to post-intervention assessment or during the follow-ups is unclear. Furthermore, the reason for the dropouts is not stated and the homogeneity of the dropouts with that of the completers is unknown. Furthermore, given the reporting of only completers implies the lack of intention-to-treat analysis utilised in the study. The follow-up periods allowed for the study to determine longevity of the effect. As an additional analysis, the study stratified the data by disease duration (less than a year and more than three years) |
| Outcome | The authors concluded that following the RASMP intervention (community education programme for the RA patient) offers benefit to the RA patient in perceived self-efficacy relating to managing arthritis, establishing positive behaviour traits for health, and stabilizing of the effects over the follow-up periods. At the post-test assessment the participants showed statistically significant improvement in SMB, ASES, AIMS, AHI, and VAS (pain and fatigue) with a ($p < 0.01$). these improvements were maintained at six-months follow-up. The score of ASES was the only outcome measure shown to report significant difference between the post-intervention assessment to the one-month follow-up, and again between the one-month and six-month follow-ups. The lack of significant change in the remaining outcomes shows the effect of the intervention archives stability over the post-intervention period. In addition, the programme showed effective in reducing participants medical care appointments. The results were statistically significant for both baseline to post-intervention follow-up and baseline to six-month follow-up (both $p < 0.05$). Disease duration stratification of the data showed participants with less than a year's disease responded significantly more so than the participants with more than three years disease, over the period pre/post assessment period for AIMS (mood and pain), SMB (exercise and management of cognitive symptoms), VAS (pain) with a ($p < 0.05$) (Chui, Lau and Yau 2004). |
| Discussion | The structure of this case series structure is pragmatic and reflective of clinical practice. Thus, it is not able to discern the outcomes of any one modality in the clinical improvement or regression of the patient with RA in this case study. It is however possible to indicate that this study does seem to suggest that there may be benefit from an education-based programme which included the use of relaxation exercise, pain and stress coping strategies, joint protection, management of cognitive symptoms, self-management skills, and management of pain, assistive device use, and techniques tailored for RA needs for joint protection. This, as the author rightly points out does require further study through a process of a randomised controlled to allow a better measure of the intervention effect (i.e., to determine whether this is greater than chance or natural history alone within the context of this case study). |
| Conclusion | Therefore, in conclusion for this study, it is evident from the reviewers' responses (collective 75% agreement) that this study was predominantly unable to fulfil or inadequately described any of the applicable criteria in the Liddle scale. This in addition to the limitations noted in this table that result in the study being ranked as subject to as high bias. This implies that the outcomes are not necessarily reproducible and that the outcomes are not only necessarily due to the interventions alone. As a result, the study cannot conclusively be seen as not being influenced by bias and the outcomes of the study cannot be determined to be better than chance or natural history alone. Therefore, at best, this particular study provides limited evidence in support of the intervention programme utilised and the reviewers are in agreement that with the high degree of bias, that further studies (should the intervention be reproducible based on other publications descriptions) in this particular area are warranted to validate the outcomes and then also to determine the contribution of each of the intervention programme components in order to allow for more specific recommendations to be made to practitioners in the field. |

Table 4.61a: Tabulated feedback of reviewer data

| Author/s | Chung C.L.R. and Mior S.A. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2015 | | | | | |
| Title | Use of spinal manipulation in a rheumatoid patient presenting with acute thoracic pain: a case report. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | B2 | B2 | B2 | B2 | 100% |
| | 2. What percentage of individuals refused to participate? | N/A | N/A | N/A | N/A | 100% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B2 | B1 | B2 | B2 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | N/A | N/A | N/A | N/A | 100% |
| | 7. Is the analysis by intention-to-intervene (treat)? | N/A | A | N/A | N/A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | C | As outlined in Table 4.61b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B2 | As outlined in Table 4.61b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | There is no ethical reason why an RCT could not be conducted, however cost and limited sample could be a limiting factor in undertaking an RCT design. | | | | |
| | 4. Include any other comments | It could be argued that there is no evidence that the treatment was targeting a joint affected by RA. The intervention may therefore not relate to the disease but rather to the RA patient's back pain that is not necessarily related to the patient's RA. | | | | |
| | | | | | Overall percentage agreement | 88% |

Table 4.61b: Properties of the study, outcome, and discussion

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Author/s | Chung C.L.R. and Mior S.A. | | | | | | | |
| Year | 2015 | | | | | | | |
| Title | Use of spinal manipulation in a rheumatoid patient presenting with acute thoracic pain: a case report. | | | | | | | |
| Study properties: | | | | | | | | |
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Pain (no outcome tool specifically stated in use) | At each appointment | One week (seven days and four treatment sessions) | One | No | N/A | N/A | Two B2 and six N/A scoring. resulting in a mostly unfulfilled criterion. | 88% |
| Limitations | The patient's context was well described through a history and assessment findings. Thus, the hypertensive 66 year old patient, diagnosed with RA had a history of being on methotrexate and use of gold injections, chloroquine, and prednisone. NSAIDS were documented for the presenting compliant. Pain relieving factors were through chiropractic manipulation and mobilisation, physical therapy modalities and soft tissue interventions. Thus, the unique patient presentation and examination findings, as are expected in this study design, limit the generalisability of the findings to the RA population. In terms of the outcome measures, the case report does not cover outcome measurement tools and their validity, nor does it detail the assessment and quantification. Furthermore, the study fails to state MCID values as a benchmark for improvement (Bahreini *et al.* 2020). The intervention was individualised and modified during the four treatments. Firstly, two chiropractors offered two consecutive treatments each and adjunctive therapy modalities were modified according to the patient's response. The first two sessions included soft tissue and interferential current along with manipulation, while the patient was advised to continue daily walks with rest according to her needs, and to use cryotherapy for control over general pain. The second intervention was modified according to the patient's feedback that cryotherapy aggravated her coughing, which increased the thoracic pain she experienced (the cryotherapy was not stated to have been excluded for the therapeutic plan). In the second two treatments, a second chiropractor provided the intervention, in treatment session three focussing on manipulation and advising the patient to avoid lifting but otherwise to continue her daily activities. There was no mention of the use of adjunctive therapies. During the fourth session, manipulation was utilised in the same manner as in session three, but this session incorporated myofascial release therapy. Although the description of this is good and provides a basis for understanding the author's pragmatic approach, it nevertheless shows significant limitations with regard to repeating this intervention in clinical practice and applying it to the correct patient to achieve optimal clinical outcomes. On discharge the patient received ergonomic advice, exercise prescription and collaborative management as needed. This programme of care and multimodality intervention with modification introduces complexity and limits the causal effect between intervention and outcome, which is even worse with regard to specific interventions in the programme (Robins, Hernán and Siebert 2004). In addition, the fact that two chiropractors were involved in providing the therapeutic intervention introduces variability in the intervention and treatment approach. This is further complicated by the fact that the patient required co-management by a physician for abdominal pain. Given the significant bias, it could be argued that the placebo effect (Kaptchuk 1998; Bialosky *et al.* 2011), the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015) and touch therapy (Melzack 1983; Moayedi and Davis 2013; Senderovich *et al.* 2016) are not controlled for in this study, and, additionally, that the patient was not naive to the modalities used in the intervention as she had sought chiropractic care because it had offered relief in the past. | | | | | | | |
| Outcome | A patient diagnosed with costovertebral joint dysfunction and a myofascial strain of the surrounding musculature was unresponsive to initial treatment involving a generalized manipulative technique. According to the case report, her condition (thoracic pain) improved following a specifically applied manipulation protocol with modified forces. Thus, the authors state that the patient reported to have experienced significant and immediate pain relief from the procedure performed in the third and fourth visit. Manipulative therapy was limited to pain relief, while the patient also received ergonomic education and active exercises" (Chung and Mior 2015). | | | | | | | |
| Discussion | With this being a case study, its structure is more pragmatic and reflective of clinical practice. Consequently, it is not able to discern the outcomes of any one modality in the clinical improvement or regression of the patient with RA. Nevertheless, this study does seem to suggest that there may be benefit from a programme, which includes the use of manipulation at specific thoracic segments. As is pointed out this field requires further study through the process of a randomised controlled or non-randomised trial in order to allow a better measure of the intervention effect so as to determine whether this is greater than chance or natural history alone within the context of this case study. | | | | | | | |
| Conclusion: | The reviewers' responses with 88% agreement indicated that this study was unable to completely fulfil the applicable criteria in the Liddle scale. This along with the limitations affecting external validity result in the study being ranked as subject to a moderate to high bias. This implies that the outcomes are not reproducible and also that they are not due to the interventions alone. At best, therefore, this particular study provides no evidence in support of the intervention programme. | | | | | | | |

Table 4.62a: Tabulated feedback of reviewer data

| Author/s | Cubick E.E., Quezada V.Y., Schumer A.D. and Davis C.M. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2011 | | | | | |
| Title | Sustained release myofascial release as treatment for a patient with complications of rheumatoid arthritis and collagenous colitis: a case report | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Researcher | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | B1 | B1 | B1 | B1 | 100% |
| | 2. What percentage of individuals refused to participate? | N/A | N/A | N/A | N/A | 100% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B1 | A | B1 | B1 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | N/A | N/A | N/A | N/A | 100% |
| | 7. Is the analysis by intention-to-intervene (treat)? | N/A | A | N/A | N/A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B2/C | As outlined in Table 4.62b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B2/B1 | As outlined in Table 4.62b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | There is no apparent ethical reason that an RCT could not be conducted. However, limitations are noted in feasibility and population size from which the sample is drawn may be limiting factors. | | | | |
| | 4. Include any other comments | The study makes use of a limited set of objective information, and that there are reported objective differences that could be within error of devices. Additionally, there appears to be no direct evidence that treatment was targeting for symptoms indirectly related to and not directly part of the RA. | | | | |
| | | | | | Overall percentage agreement | 88% |

Table 4.62b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Cubick E.E., Quezada V.Y., Schumer A.D. and Davis C.M. |
| Year | 2011 |
| Title | Sustained release myofascial release as treatment for a patient with complications of rheumatoid arthritis and collagenous colitis: a case report |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Cervical pain VAS<br>Systemic pain VAS<br>Fatigue P4 instrument (morning, afternoon, evening, during activity)<br>GIT function<br>QoL (AIMS2)<br>ROM (cervical; flexion, extension, rotation bilaterally, lateral flexion bilaterally) (goniometer) | Pre and post evaluation were conducted at each treatment and during the interim between initial and final treatment series.<br>The measures were individually reported for each outcome. Each outcome measure was applied differently during the treatment | Two weeks (sustained release)<br>eight weeks brake and two more treatments | One | No | N/A | N/A | Two B1 and six N/A scoring. resulting in mostly fulfilled criteria. | 88% |

| | |
|---|---|
| Limitations | This case study seems to have been based at the University of Miami's Department of Physical Therapy, Miami, USA. The patient was described as 54-year-old white female presenting with both cervical (neck) and systemic pain (without clarity of described), fatigue and gastrointestinal symptoms of nausea, bowel urgency and explosive diarrhoea (the latter diagnosed as collagenous colitis a form of irritable bowel). Her medical history included medication (infliximab) and previous treatment exposure (chiropractic) reporting no pain relief. The authors attempted to control for confounding factors, by restricting chiropractic care, continued infliximab therapy throughout. The initial use of pain medication was utilised as an outcome measure (the patient decreased use during the study). At the initial assessment the physician recorded findings of the postural assessment, fascial mobility, ROM of the cervical spine, and imagining (radiological) findings. These provided a fairly comprehensive context for the intervention and allows for the extrapolation of outcomes to patients with a similar demographic and symptom profile.<br><br>Given that this was a case study, the criterion of homogeneity is not applicable, but also does infer that there is an increase in bias and thus outcomes generalisability.<br><br>In terms of rigour, the study makes use of a mix of outcome (subjective and objective) thereby enabling control of patient perception influences (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). However, the study lacked details of validity and reliability of the outcome measures. There is a lack of clarity with regards to the relationship between intervention administration and outcome measurement, which reduces the ability of the study to clearly delineate the outcomes as a result of the intervention. No rationale was provided for the variable time points for the different measures and how these related to the intervention and / or what impact the measures were expected to have reflected at the different time points. These concerns introduce measurement bias (negative or positive). Additionally, blinding of the assessor was not apparent and thus the influence of assessor bias is not controlled for. All these interactions between the interventions and measurement in addition to the lack of statistical rigour increases the chances that the outcomes are as a result of natural history and / or chance as compared to actual interventional changes.<br><br>From the report, the patient was not naive to interventions utilised in this study. Some interventions are well described (viz., techniques, duration, application, and description) which allows for their replication in clinical practice / future studies, however there are others that are poorly described and thus difficult to replicate. As a result, the overall programme of care is not clinically reproducible (Barnes 1990). To further complicate this the intervention was administered by three different therapists / practitioners, which increases the likelihood of differential application (viz.no mention standardised training programme) and thus inter-therapist variability bias. Additionally, these concerns are compounded by lack of blinding of the patient and therapists, which further decreases the ability of the case study to directly correlate outcomes to the intervention programme.<br><br>Results suggested improvement, but with the lack of published MCID comparisons, there is limited ability of the reader to understand the clinical impact of the outcomes obtained (Bahreini *et al.* 2020). Then also, in the context of a case study, the ability to determine statistically significant outcomes is limited, as this statistical significance cannot be taken as being better than natural history and /or chance. These outcomes are further biased / obscured in that the authors use of subjective quotes providing a perceived depth to the data; however, the use seems to skew the data towards a positive intervention effect. This is then negated in the case study where the authors later acknowledge the return of symptoms as a result of a stressful period. |

| | |
|---|---|
| Outcome | The authors claim that the study indicates improvement in outcome measures (pain and ROM of the cervical spine, fatigue, systemic pain, and QoL) following a six-intervention session programme, showed positive effects at five-weeks, after which the patient's pain and IBS signs and symptoms regressed to near baseline values. The two additional intervention sessions post-interlude provided the positive gains needed to regain these effects. The authors concluded that MFRT-sustained release technique would effectively offer patients with RA and collagenous colitis positive outcomes, as an adjunct to medical intervention, resulting in short- and long-term improvement of QoL and comorbid symptomatology (Cubick *et al.* 2011). |
| Discussion | Given that this study is a case study, its structure is one that is more pragmatic and thus more reflective of clinical practice. This implies that the study is more prone to bias resulting from 1. lack of blinding (patient and assessors), 2. varied interventions that are tailored to the patient (and often not fully detailed) resulting in a unique programme of care, 3. outcome measures that are not necessarily meant to measure outcomes of RA specifically (but rather symptomatic change) and lastly 4. the lack of statistical rigour that can imply change that is greater than chance or natural history.<br><br>Thus, the case study is also not able to discern the outcomes of any one modality (as the intervention is a programme) in the clinical improvement or regression of the patient with RA. It may be possible that there was benefit from a MFRT, but as the authors point out, this requires further study. The most appropriate a structured randomised controlled or non-randomised trial which manages the limitations of a case study and attempts to elucidate the actual clinical change attributable to the MFRT. This will require that the future study addresses all the limitation of this case study as noted under limitations in order to determine whether this clinical change is positive or negative and then also greater than chance or natural history. |
| Conclusion: | Therefore, in conclusion for this particular study, it is evident from the reviewers' responses (collective 88% agreement) that this study was able to fulfil the applicable criteria in the Liddle scale resulting in moderate to high degree of bias. This, in addition to the limitations noted in this table suggest that the study attains a moderate to high bias ranking. This implies that the outcomes are not necessarily reproducible, better than chance or natural history and that the outcomes are not only necessarily due to the interventions alone.<br><br>Therefore, at best, this particular study provides limited evidence in support of the intervention programme being clinically effective for RA. As a result, the study conclusions have to be seen with caution regarding the influence of bias and confounding. Thus, it is suggested that further studies in this particular area are warranted to validate the outcomes using specific, valid and reliable outcome tools and then also to determine the clinical effectiveness of the allowing for more specific recommendations to be made to practitioners in the field. |

Table 4.63a: Tabulated feedback of reviewer data

| | | | | | | |
|---|---|---|---|---|---|---|
| Author/s | Kapale P., Vardharajulu G., and Warude T. | | | | | |
| Year | 2017 | | | | | |
| Title | Effect of free exercise and rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | C | A | C | C | 67% |
| | 2. What percentage of individuals refused to participate? | C | A | C | C | 67% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | C | B1 | C | C | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B2 | C | - | 0% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | N/A | A | C | - | 0% |
| | 7. Is the analysis by intention-to-intervene (treat)? | N/A | A | N/A | N/A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B1 | As outlined in Table 4.63b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | A | As outlined in Table 4.63b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | No practical or ethical reason | | | | |
| | 4. Include any other comments | As the main outcome measure is not defined and there is no control group, any changes cannot be assessed. | | | | |
| | | | | | Overall percentage agreement | 59% |

Table 4.63b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Kapale P., Vardharajulu G. and Warude T. |
| Year | 2017 |
| Title | Effect of free exercise and rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Functional status scale | Pre/Post | Five-weeks | 16 | No | N/A | N/A | Three C, three N/A, and three failures to agree scoring, resulting in mostly unfilled criteria, of applicable criteria. | 59% |

| | |
|---|---|
| Limitations | The study is based at the Krishna Institute of Medical Sciences Deemed Institute in Karad, District of Satara, Maharashtra, India. The 16 participants that made up the cases series were derived from the hospital rheumatology department and / or clinic, or the districts the university medical school services. Without clear evidence as to the background from which the participants are derived. If the participants are solely taken from a specialist care environment, it would imply that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012). limiting the generalisability of the study's findings into clinical practice. The converse would be true if the participants were drawn from the general community.<br><br>The 16 participants were recruited by means of convenience sampling (raising risk of sampling bias), were clinically diagnosed with RA (diagnostic criteria not cited). The participants were comprised of both sexes (with disproportionately more females (5:11)) and of a range of "age symptomatology". The participants were subject to exclusion if they had structural or functional extremities impairments from birth, any absenteeism of the first or last assessment period and / or the concurrent the use of electro-modalities whilst on the study. Furthermore, the sample size was small given this broad recruitment criteria. No further demographic information is offered negating the ability of the reader in identifying confounding factors (for example, age, medications, and the prescription of them), as these will (if widely distributed in the group) result in varied outcomes between the participants and thus a regression to the mean in terms of outcomes. This implies that the improvements noted may not be better than chance alone.<br><br>The intervention was minimally described and provided a significant barrier to the reader in replicating the intervention. It consisted of "free exercises" three times a week over a five-week duration. These exercises incorporated both upper and lower limb exercises in three sets of 10 repetitions each. Despite a long explanation of the free exercise mechanism of action and the involvement of gravity as the resistance source little is detailed in terms of the implemented intervention. In terms of the participant naivety, it is not reported in the study limiting any degree of analysis, and raising questions around patient perception bias, learning bias, etc. (Heine *et al.* 2012).<br><br>The Assessor is assumed to lack blinding of the participants interventions (given the study design = case series, despite the possibility of blinding still being somewhat achievable) introducing assessor bias (Kendall 2003; Moustgaard *et al.* 2014). The nature, timing, environment, and number of assessors is not velar and for this reason all these factors raise concerns of bias introduced into the study (Pater *et al.* 1998; Richardson 2007). The assessment followed a pre/post design with the functional status scale as the outcome measure utilised. The exact details of this outcome tool (lacking citation), its validity and reliability are all unknown and thus raise the concerns of RA population appropriateness.<br><br>The outcome measure had baseline values (pre-test assessment) are reported as a mean with standard deviation. However, the study failed to take this opportunity to control for a number of other bias and confounding factors (i.e., age, disease duration and activity, medication, comorbidities) (see Section 4.4.1.1).<br><br>The study states (in the discussion) that the effect of free exercise on pain and functional disability were both tested, and both returned significantly reduced post intervention. However, this is not clearly described nor understood as the outcome tool in use (functional status scale) was unclear as to what it measured and if assessing only functional capacity where the authors derive data on pain remains unclear. When the conclusion is read the study stats "From the results of this study it is concluded that free exercises improved the range of motion, relief of pain, and reduction in functional disability in rheumatoid arthritis patient". The addition of ROM |

| | |
|---|---|
| | and pain are incorporated into the conclusion while now clear outcome tool is utilised to provide such data nor whether this data produced is valid and reliable (Coster 2013; Bahreini *et al.* 2020).<br><br>The study makes use of statistical analysis (including means, standard deviations and p-values based on the Wilcoxon match pair test and t test) an aspect of a study seemingly commonly excluded in case reports and series. Given the absenteeism as an exclusion criterion and the pre/post assessment design the use of intention-to-treat analysis is not viable, power analysis was not conducted to determine the necessary sample size required to detect intervention influence on what appears to be two outcomes (pain and functional disability – seemingly derived from a single outcome tool – functional capacity scale). This threatens the studies of underpowered analysis and compromises the strength of the results, the generalisability and replicability of the study in both clinical and research environments (Gupta 2011; Hoskin 2012; Bahreini *et al.* 2020).<br><br>It should be noted that an inherent weakness of the case report and case series design is the lack of control or comparison group limits the value of statistical analysis and the results should be read with such caution. |
| Outcome | The authors reported that significant benefit can be gained through free exercise over a five-week period. The outcome measures demonstrated that the participants experiences, significant improvement in pain ($p$<0.001) as well as, in ROM ($p$<0.0001). Both these findings illustrate that functional status is improved in the RA sample (Kapale, Vardharajulu and Warude 2017). |
| Discussion | Given that this study is a case series structure is one that is more pragmatic and reflective of clinical practice. Thus, it is not able to discern the outcomes of any one modality in the clinical improvement or regression of the patient with RA in this case study. It is however possible to indicate that this study does seem to suggest that there may be benefit from a “free exercise” intervention employed in this study. This does require further study through a process of a randomised controlled or non-randomised trial in order to allow a better measure of the intervention effect (i.e., to determine whether this is greater than chance or natural history alone within the context of this case study). |
| Conclusion: | Therefore, in conclusion for this particular study, it is evident from the reviewers’ responses (collective 59% agreement) that this study failed to fulfil at all, the applicable criteria in the Liddle scale. This in addition to the limitations noted in this table that result in the study being ranked as subject to as high risk of bias. This implies that the poorly described outcome/s are not reproducible, and that the outcome/s are not only necessarily due to the interventions alone. As a result, the study cannot conclusively be seen as not influenced by bias and the outcome/s of the study cannot be determined to be better than chance or natural history alone. Having given this outcome, the reviewers are aware that the article seems to be a translation, or it was written in a second language for the authors, thus it is also necessary to indicate that language bias may have had an impact on the review of this article.<br><br>Therefore, excluding the consideration or language bias, this study provides little to no evidence in support of the intervention programme utilised. As a result, the study conclusions must be seen with caution regarding the influence of bias and confounding. Further studies in this particular area are required to validate the outcome/s and to determine the contribution of such an intervention in order to allow for more specific recommendations to be made to practitioners in the field. |

Table 4.64a: Tabulated feedback of reviewer data

| Author/s | Kauppi M., Leppanen L., Heikkila S., Lahtinen T. and Kautiainen H. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1998 | | | | | |
| Title | Active conservative treatment of atlantoaxial subluxation in rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | B1 | A | B1 | B1 | 67% |
| | 2. What percentage of individuals refused to participate? | B1 | A | A | A | 67% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B2 | B1 | B1 | B1 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | A | N/A | A | A | 67% |
| | 7. Is the analysis by intention-to-intervene (treat)? | C | A | A | A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B1 | As outlined in Table 4.64b there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B1 | As outlined in Table 4.64b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an R CT cannot be done. | There are no practical or ethical reasons that a RCT could not be conducted. | | | | |
| | 4. Include any other comments | The intervention programme was far to extensive, multidisciplinary, and complicated. Limiting the study from creating any causal link for a single aspect of the intervention to an outcome. The study fails to create a protocol consistency and in such should have provided stratified data, for factors like adherence and capacity to implement the intervention in the home-environment. Furthermore, the lack of a clear and well-defined set of inclusion and exclusion criteria missed an opportunity to control for some of the confounding and bias factors in the study or RA and offer the reader a clear picture of the population to which the studies could be generalised to. | | | | |
| | | | | | Overall percentage agreement | 75% |

Table 4.64b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Kauppi M., Leppanen L., Heikkila S., Lahtinen T. and Kautiainen H. |
| Year | 1998 |
| Title | Active conservative treatment of atlantoaxial subluxation in rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Cervical and / or occipital pain (VAS) for previous day, last month, morning, afternoon, and night.<br>- atlantoaxial distance<br>- atlantoaxial instability<br>- atlantoaxial impaction<br>- VAS for programme implementation | For pain baseline, after 2 weeks, 6 months and then 12 months. And the sessions following the 12-month follow-up)<br><br>Radiology (baseline, month 12) | 54 weeks (14 days intervention, and two "control and motivational visits" (follow-ups) at six months and 12 months later). | 20 | No | N/A | N/A | Two B1, three A and three N/A scoring, resulting in mostly completely filled criteria, with the remaining applicable criteria mostly filled. | 75% |

| | |
|---|---|
| Limitations | This study was conducted at the Rheumatism Foundation Hospital in Finland, with participants recruited via Finnish rheumatologist referral of patients with atlantoaxial subluxation (AAS). Despite the prevalence of AAS in RA patients ranging between 29 and 70% (Kauppi *et al.* 1998) the perception from the outset exists that the sample consists of patients that have previously sought specialist care (given the referral recruitment), implying that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012).<br><br>The first 20 participants were admitted into the study with none of these participants being excluded. The study did not report any inclusion or exclusion criteria and it is unclear whether specific instructions were given to referring rheumatologists who would then have had the ability to screen the patients. From the provided information, all participants had clinically diagnosed RA (diagnostic unstated), the participants all had AAS (validated in the baseline radiology), and a lack of cervical symptomatology was not required, nor did it result in exclusion. The participant description post inclusion revealed a mean age of 47.9 years (range 26 to 67 years and SD of 9.7 years) with female to male ration (19:1) and a disease duration averaging 19.9 years (range 3 to 32). The disease activity status was recorded by means of ESR (average 32.8 mm/h (SD of 21.6 mm/h), indicating very variable disease activity. All participants had deformity of the peripheral joints in addition to "most" of the participants reporting other complaints outside of the cervical spine region. And lastly, the authors report that there was a mean latency period from referral to entry into the study and start of the treatment programme (8.9 weeks; SD 4.0 weeks). This may be the cause of variance between the participants but may additionally result in further variance at the time of intervention initiation if the descriptive data was taken 8.9 weeks before treatment started. This is also complicated by the fact that six participants had reported co-morbidity secondary amyloidosis (with variance over latency period). In terms of medication control and prescription, 19 were prescribed DMARDs, 19 were prescribed low-dose steroids, all 20 of the participants were prescribed NSAIDs. The study fails to report on the types of medication and whether they were all the same in regards frequency, duration of the therapy, and dosage of the medication. It is thus clear that the patient profile was variable, lacking homogeneity and does not allow the reader to extrapolate the outcomes of this study to patients in clinical practice as it is difficult to identify which patients would benefit from the intervention.<br><br>The intervention was a programme provided by an in-hospital multidisciplinary for the initial part of the study (14 days) followed by a home-base exercise programme and hospital re-admission (three days) after the six-month follow-up (Robins, Hernán and Siebert 2004). The programme consists of education of the patient, physiotherapy, aids (including collars and practical aids) and pharmaceutical therapy both symptomatic and DMARDs (Kauppi 1996; Kauppi and Anttila 1996). As an adjunctive, participants incorporated general rehabilitation exercises (i.e., hydro-gymnastics). Based on the range of interventions, the programmes were individualised, which introduces a protocol inconsistency and reduces the capacity of the reader to replicate the intervention in clinical practice. It does however mean that participants physical limitations could be taken into account, but since these are unreported in the results (i.e., as different groupings of patients, it is difficult to assume that the outcomes for all patient types were similar – thus the outcomes of the study as aggregated results, means they actually represent potentially vastly different scenarios. |

| | |
|---|---|
| | Participant naivety was reported in that participants had no exposure to "special neck exercise programmes for rheumatoid cervical spine". However, this contradicts the recruitment method where referrals came from rheumatologists that may have previously exposed patients to similar interventions. The lack of clear nativity in the sample raises concerns of bias introduced by participant preference and prior experiences, which may colour their responses to subjective outcome measures (Mouton 2006). The impact of naivety also has implications for programme adherence and the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015) and how these two concepts counter each other. The adherence was partially monitored by supervision and secondarily by an exercise diary, a VAS of participant capacity to carry out the programme during the first six months and a questionnaire on compliance. It is however, reported that some participants increased frequency of exercise when neck pain was experienced. Thus, at best, there was some moderate control for bias in terms of exercise adherence and supervision.<br><br>A strength of the study was that it study made use of both subjective and objective outcomes controlling bias from subjective input. The lack of validity and reliability for the outcome measures prevents a clear understanding as to whether these tools were appropriate outcomes measures for the interventions used (Coster 2013; Bahreini *et al.* 2020). The range in variance of patient presentation also suggests that the outcome measures may have been subject to floor or ceiling effects in terms of the range of improvement available to each participant. These effects would have resulted in the outcomes aggregating to a mean value, as opposed to actual clear clinical changes. Thus, the use of subgroup reporting may have allowed better reporting of the outcomes. In terms of the statistics, the study reported a power analysis (*p* value of 0.05 – two-tailed) for determining the possibility of rejecting the null hypothesis (i.e., outcomes were better than chance alone), with 20 participants. Given that 20 patients were accepted onto the study by convenience sampling it assumes no dropouts. The limitation of the statistics is the use of the Bonferroni-adjusted paired sample confidence interval, which indicates that there were outliers in the study that needed to be accounted for. The statistics would have been strengthened if the authors had compared improvements in the subgroups of patients (dependent on activity limitation) against known MCID values for the measurement outcomes (Bahreini *et al.* 2020). |
| Outcome | The authors concluded that of the 20 participants 17 complied with the intervention well and participated in the exercise component of the intervention at least one day a week. Furthermore, this high rate of adherence to the programme continued for the entire 12-month study duration. The cervical spinal pain had reflected statistically significant reductions during the first intervention period (14-day in-patient) at the week 2 assessment, with a *p*-value of 0.001. These favourable results were maintained for 12 months. positive trends were noted in the changes in cervical radiology outcomes. Given these findings RA patients with AAS could possibly benefit from education and motivation to participate in neck active care, gaining significant relief from chronic pain in the cervical spine, and potentially reducing AA instability during "active conservative treatment". The authors state "thus, the treatment has to be multidisciplinary and based on a combination of "patient education, but collars, ergonomics, practical aids, muscular exercises, massage, anti-inflammatory and analgesic drugs, local injections and active disease-modifying medication together form the combination, which may be able to fulfil the aims of active conservative treatment"; in order to achieve the desired outcomes. |
| Discussion | With this study being structured as a case series, it falls prey to a lack of homogeneity (if the inclusion is not clearly defined), patient naivety, treatment novelty and patient prior experiences and expectations influencing the outcomes if not managed appropriately. Additionally, the statistical control of the outcome's limits reflecting the true clinical outcomes of all participants and represents averagely what occurred. This regression to the mean may influence the ability of the outcomes measures to detect change over time and / or due to treatment. Finally, the variability of the outcome measures is also affected (ceiling and floor effects) if participants are very varied, in addition to being affected by lack of assessor blinding and use of measures that have no shown validity or reliability as outcome measures – particularly when measuring RA outcomes. However, if the outcomes were reported by subgroup or through individual analysis of outcomes, the structure and reporting would be pragmatic and reflective of clinical practice. This does not however allow for appropriate assessment of any one modality in the clinical improvement or regression of the patient with RA. The study does however seem to suggest that there may be benefit from a programme it applied. Thus, as the authors rightly point out does require further study through more rigours RCT or n-RCT processes, addressing concerns expressed in the limitations (viz., larger sample size; longer follow-ups; long term outcomes; effect of patient variables on the outcomes |
| Conclusion | Therefore, in conclusion for this particular study, it is evident from the reviewers' responses (collective 75% agreement) that this study was able to fulfil the majority of the applicable criteria in the Liddle scale indicating a high degree of bias. This combined with the limitations noted in this table suggest a high bias ranking. Given this ranking, it implies that the outcomes are not necessarily clinically reproducible and that the outcomes are not only necessarily due to the interventions alone.<br><br>Therefore, and at best, this particular study provides limited evidence in support of the intervention programme utilised for RA patients. As a result, the study conclusions have to be seen with caution regarding the influence of bias and confounding. Further studies in this particular area are warranted to validate the outcomes and then also to determine the contribution of each of the intervention programme components in order to allow for more specific recommendations to be made to practitioners in the field. |

Table 4.65a: Tabulated feedback of reviewer data

| Author/s | Lonta M.K. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1960 | | | | | |
| Title | Facilitation technics in the treatment of early rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | B1 | A | B1 | B1 | 67% |
| | 2. What percentage of individuals refused to participate? | N/A | N/A | N/A | N/A | 100% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | C | A | C | C | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | N/A | N/A | N/A | N/A | 100% |
| | 7. Is the analysis by intention-to-intervene (treat)? | N/A | A | N/A | N/A | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | C | As outlined in Table 4.65b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | C | As outlined in Table 4.65b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | No practical or ethical reason why an RCT could not be conducted. | | | | |
| | 4. Include any other comments | The study participant appears to be in acute RA disease status. This would limit the generalisability of the finding as chronic RA differs in many ways and are frequently used in trials. | | | | |
| | | | | | Overall percentage agreement | 84% |

Table 4.65b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Lonta M.K. |
| Year | 1960 |
| Title | Facilitation technics in the treatment of early rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| ROM and strength (right wrist, hand, and knee and both shoulders)<br>Sedimentation rate<br>ADL<br>Radiology (no site stated)<br>Gait | Physical exam at admission (baseline) with no recorded measurements | From the 24th of March to 21st of April, 29 days (the dates in reporting do not match up with the reported duration. | One | No | N/A | N/A | One B1, one C scores with six N/A scoring resulting in reviewers finding that of the two applicable criteria; one criterion was mostly fulfilled while the other unfulfilled. | 84% |
| Active joint ROM:<br>- Right knee (extension and flex to work out range).<br>- Right wrist (extension and flex to work out range.<br>- Right index finger (flexion at the metacarpophalangeal joint, proximal, and distal interphalangeal joint.<br>- Bilateral shoulder (flexion, internal and externally rotate) | Day two, four, ten, 12, and 29. | | | | | | | |

| | |
|---|---|
| Limitations | It was assumed that the study was carried out at the Massachusetts General Hospital, Boston (author location). The patient was admitted for seeking therapy. This implies that the patient suffered from a more severe or debilitating form or disease activity of RA prompting the care seeking behaviour. If the participant was solely taken from a specialist care environment, it would imply that RA disease activity and severity may be more severe than the average RA patient (Wells *et al.* 2012) further limiting the generalisability of the study's findings into clinical practice. However, the patient presentation is fairly well described, with the patient being 40 years of age, with a social and medical history including both previous and current work environment, parity, surgical history and home-life. The chief complaint was pain and swelling of the joints (of the right fingers, wrist, and knee, along with bilateral shoulders), fatigue induced by physical exertion and loss of ROM in the right hand accompanied by morning stiffness, wrist pain on the right and ipsilateral knee swelling and stiffness particularly in the mornings, as well as generalised joint ache. Reliving factors reported that improvement was seen with active movement of the joints. Given that this was a case study, homogeneity between groups is not applicable. This compounds bias by both a small sample (being that it is a case report of a single participant) and its reduced generalisability. At baseline, the patient presented with tense apprehension, weakness, and limited ROM of the right knee as well as the shoulder joints. Additionally, the examination revealed limitations in the ADL (right hand) and ambulation limitations (based on observation). No formal recording and reporting measures were utilised at the baseline assessment of the patient (e.g., a goniometer utilised in the ROM findings). Reported special investigations included a raised sedimentation rate indicating active state of inflammation, and radiological features. No clear diagnostic criteria were used as the study predates the 1987 (Arnett *et al.* 1988) and 2010 ARA/ACR criteria (Aletaha *et al.* 2010), thus the comparability of this patient in terms of diagnosis, disease progression and stage of disease to others in this systematic review is problematic and cannot be assumed to follow any particular criteria. The presentation suggests that the patient was provided a diagnosis of RA after the initial evaluation and special investigations, which would make the participant naïve to standard RA care / the study intervention programme; improving the chances that the outcome is more likely to be accredited to the intervention than prior perceptions held by the patient (Mouton 2006).<br><br>The intervention programme in this study included the staged incorporation of interventions including heat, bed rest, medication, and physical modalities. On admission the patient began therapy consisting of; bed rest, hot baths (daily), application of hydrocollator (twice a day to the shoulders and knees; and the right hand and wrist), and medication (Bufferin (Aspirin), phenobarbital and Seconal). This is compounded by the subsequent incorporation of daily physical therapy (physiotherapy department) in order to strengthen all muscle groups related to the joints and maintain / improve of ROM (Robins, Hernán and Siebert 2004). Thus, interventions included active ROM exercises of all joints through proprioceptive neuromuscular facilitation whilst rhythmic stabilization (Knott and Voss 1956) was also included. Prior to each exercise session |

| | |
|---|---|
| | hydrocollator was applied to all involved joints and breathing exercises were incorporated into each physical therapy session The patient was encouraged to repeat the movement patterns within the pain free range in addition to the face-to-face sessions. The primary aim of the incorporated physical therapy was to improve the patient's upper limb function and improve self-care capacity and ambulation. After discharge the patient was reassessed every two weeks, and exercise prescription modified in accordance with the assessment results. In these subsequent sessions, the patient was given axillary crutches and instructed in the use of a four-point gait. The combined use of manual interventions in addition to medication in a staged manner that is not well described, only allows for conclusions to be drawn about the programme and not individual interventions. In addition, no comments can be made with regards to agonist or antagonist effects of interventions within the programme (Robins, Hernán and Siebert 2004). This limits the reader from determining the component's role in the outcomes. Further the study was over a short period and limits the effect induced by the intervention to manifest and provide change detectable by outcome measures and that these tools are sensitive and specific enough to detect this (Coster 2013; Bahreini *et al.* 2020).<br><br>Aside from the history and physical examination outcomes noted, the patient was re-assessed for joint ROM at each of the following four physical therapy sessions and at the single follow-up reported. The results of the joint ROM are in degrees, but the outcome tool utilised to provide consistent and reliable data is not described. In addition, the baseline data (from initial screening versus prior to the initial physiotherapy intervention is not clearly delineated (viz., computations of ROM were compared to the first physiotherapy visit and not the baseline). Home exercises were not monitored so compliance with requested home therapy may have positively (compliant) or negatively (non-compliant) affected outcomes. In terms of the ADLs, there was reported improvement in late March and early April and improved ambulation (with crutches) in early April. These are not reported subsequent to this point. This lack of accuracy of repeated measures (uniformity and reproducibility) and tracking of outcomes is questioned as the input would have been largely subjective reporting from the patient (and affected by the potential for Hawthorne and touch effects and pain tolerance/threshold (Melzack 1983; Dhondt 1999; Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015; Senderovich *et al.* 2016)) has a largely negative impact in terms of bias and thus potentially reproducibility of the outcomes in a larger study. This is compounded by the lack of assessment details around timing (between intervention and assessment), environment in which the assessment is conducted and the consistency of the assessor, the assessor not being blinded (given the study layout) introducing further bias into the study. Given that there was only one participant in the study there were not statistical analyses of the outcomes, but rather comparison at time points. It would therefore have been appropriate for the author to consider comparing changes in the patient's outcome measures to known MCID values in order to chart clinical significance (if attained or not), in order to be able to draw contextual and relevant conclusions (Bahreini *et al.* 2020). |
| Outcome | A 40-year-old female participant admitted into hospital in-patient care with indications of RA. Over 28 days a multidisciplinary intervention programme incorporating medication, thermotherapy, bed rest, a large collection physical modalities and walking aids in a staggered manner in two settings (first the in-patient and then home-based). This was aimed at improving the upper limb function, self-care capacity and ambulation. After 14 days in-patient care (with 12 days of physical therapy), it was reported that the ROM of the joints improved "significantly", and at the follow-up at two-weeks post discharge were maintained for all joints were maintained, aside from the right wrist losing five degrees of extension and gaining five degrees on flexion. The index digit lost ten degrees in the MCP, PIP and DIP joints of the right hand. The results report qualitative, subjective improvement in ADL (such as ambulation and grip strength) and pain (Lonta 1960). |
| Discussion | Given that this study is a case study its structure is one that is more pragmatic and reflective of clinical practice. However, given the lack of overt reporting of the outcomes with reliable and reproducible methods at set time intervals for specific reasons, detracts from the ability of the study to provide a structured report of findings as a result of the intervention programme used. In addition, the study cannot discern the clinical outcomes attributable to any one modality (positive or negative). Further to this the outcome measures also seem to be subject to bias from the assessor as well as the patient (with the exclusion of patient naivety to the interventions), which further highlights that the outcomes may not be only as a result of the intervention programme. The study does seem to suggest that there may be benefit from a graded programme including the use of a multidisciplinary intervention programme incorporating medication, thermotherapy, bed rest, a large collection physical modalities and walking aids. This would however require the use of a structured randomised controlled trial, controlling for several major bias and confounding factors along with a control group in order to allow a better measure of the intervention effect (i.e., to determine whether this is greater than chance or natural history alone within the context of this case study). |
| Conclusion | Therefore, in conclusion for this particular study, it is evident from the reviewers' responses (collective 84% agreement) that this study was able to mostly fulfil one criterion, another totally unfulfilled of the applicable criteria in the Liddle scale. This indicates that there is a high agreement for moderate to low rigour and moderate to high bias. In addition, the impact to the external validity by study limitations noted in this table, result in the study being ranked as subject to as high degree of bias. This implies that the outcomes are unlikely to be reproducible and are unlikely due to the interventions alone. As a result, the study is influenced by bias and the outcomes of the study cannot be determined to be better than chance or natural history alone. As a result, the study conclusions must be seen with caution regarding the influence of bias and confounding. Therefore, at best, this particular study provides no to very limited evidence in support of the intervention programme used (reviewers are in agreement regarding a high degree of bias), that further studies in this particular area are warranted to validate the outcomes and then also to determine the contribution of each of the intervention programme components in order to allow for more specific recommendations to be made to practitioners in the field. |

Table 4.66a: Tabulated feedback of reviewer data

| Author/s | Neuberger G.B., Press A.N., Lindsley H.B., Hinton R., Cagle P.E., Carlson K., Scott S., Dahl J. and Kramer B. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1997 | | | | | |
| Title | Effects of exercise on fatigue, aerobic fitness, and disease activity measures in persons with rheumatoid arthritis | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | A | B1 | B1 | B1 | 67% |
| | 2. What percentage of individuals refused to participate? | A | B2 | A | A | 67% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B1 | A | B1 | B1 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B2 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | C | B1 | B1 | B1 | 67% |
| | 7. Is the analysis by intention-to-intervene (treat)? | C | A | C | C | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B1 | As outlined in Table 4.66b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B1 | As outlined in Table 4.66b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an R CT cannot be done. | No practical or ethical reason. | | | | |
| | 4. Include any other comments | Although medication change was recorded, dosage was not mentioned either as a potential confounding variable or confirmation of efficacy of the treatment protocol. Patients change in medication as pain fluctuates, unless told categorically not to. Hence medication use can often be a mark of change in symptoms. This became a comparative trial, due to the stratification of data by attendance rate by the authors. There is no reason why a natural history control could not have been included. | | | | |
| | | | | | Overall percentage agreement | 75% |

Table 4.66b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Neuberger G.B., Press A.N., Lindsley H.B., Hinton R., Cagle P.E., Carlson K., Scott S., Dahl J. and Kramer B. |
| Year | 1997 |
| Title | Effects of exercise on fatigue, aerobic fitness, and disease activity measures in persons with rheumatoid arthritis |

Study properties:

| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
|---|---|---|---|---|---|---|---|---|
| - Multidimensional Assessment of Fatigue (MAF)<br>- Global Fatigue Index<br>- Profile of Mood States – Short Form (POMS)<br>- Arthritis Impact Measurement Scales (AIMS)<br>- Bicycle ergometer testing<br>- Joint count<br>- Grip strength (portable sphygmomanometer)<br>- 50 feet walk time<br>- Sedimentation rate | Baseline (week 0), week 6, week 12, follow-up 15 weeks.<br><br>2 baseline assessments (6 weeks apart), week 6, week 12, week 27.<br>(All outcome measures are assessed at all 5 assessments with exception of bicycle ergometer test which was at baseline, week12 and week 27) | 27 weeks | 32 recruited into study with 2 drop out before the week 6 assessment,<br>2 after week 6 assessment,<br>3 after week 12 assessment.<br><br>Thus 25 completed the study (78%) | No | N/A | N/A | One A, three B1, one C and three N/A scoring. Resulting in mostly fulfilled criteria.<br>(A fulfilled and B1 mostly fulfilled) | 75% |

| | |
|---|---|
| Limitations | The study was conducted at the University of Kansas and utilised the university database to identify potential RA participants. This form of convenience sampling introduces risk of bias, but also allowed for several avenues for recruitment and increased the likelihood of a sample with variable presentations. This protected against a sample consisting of patients that have sought out specialist care, for RA disease activity and severity that may be more severe than the average RA patient (Wells *et al.* 2012). The study recruited 32 participants, however the criteria utilised, how the diagnosis was confirmed and whether the diagnosis was validated by the study is not reported. Aside from the RA diagnosis the inclusion criteria required the participant to be English literate, have no comorbid conditions, had their rheumatologist approval to participate in the study and had no previous exposure to a regular programme of aerobic exercise (defined as half an hour three times weekly).<br><br>The patient demographic data reported a mean age of 55 years (range 30 to 70 years), mean disease duration of RA 9.8 years (range of 7 to 30 years), ethnicity (represented Caucasian, African American, Hispanic, and Asian), sex (male and female), along with education, employment, income, support structure and functional class status (class II (n=22) and class III (n=3) (Schumacher, 1993)). No previous exposure to a regular programme of aerobic exercise was permitted. At a later stage in the article, the author reports to have taken participants weight, height and blood pressure with a consistent scale and sphygmomanometer. This was taken at the initial assessment and controlled for variability in measure tool and calibration issues (regarding this measures). However, the study fails to then report the data of these. The data was analysed stratified by level of exercise participation (n=8 high (> 31/36 sessions attended); n=8 medium (between 25 to 30/36 sessions attended); n=9 low (<24/36 sessions attended)). The participants did not differ in education, mean age or sex. Aside from these groupings, there was no control to which the intervention group could be compared. Thus, although homogeneity was attained between the groups, there is no ability to compare to a control group. The study fails to provide data across the range for all recruited participants (n=32) and the drop-out of seven (representing 22% of the initial sample) (Heritier, Gebski and Keech 2003) through the study (when and reason provided).<br><br>The multidisciplinary intervention consisted of three one-hour low-aerobic classes (Robins, Hernán and Siebert 2004). Each class followed four-phases of exercise (warm-up, strengthen, low-impact aerobic, cool-down). The recorded phases of the intervention modified through the study. Initially the programme stimulated muscle strengthening, with progression to an aerobic phase dominance. The warm-up and cool-down phases shortened over time. The details of participation (average time of exercise engagement per week) were reported but not the exact exercises utilized in each phase if the intervention. The lack of these protocols raises concerns around replication and bias (Heine *et al.* 2012). By contrast the 12 weeks of exercise-based intervention is sufficient to induce measurable clinical effect (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012).<br><br>A strength of this study is the recording of assessment time to determine and thus control for the diurnal variation in fatigue (Nandgude, Hasabe and Kolsure 2018). With these, the authors report non-significant association of assessment time and fatigue rating. Furthermore, a specific procedure of outcome measures was used and the environment in which this occurred |

| | |
|---|---|
| | are detailed allowing for context and influences of bias and confounding to be controlled (Kendall 2003), along with calibration (Kimberlin and Winterstein 2008), assessment timing (Nandgude, Hasabe and Kolsure 2018), and measurement environment consistency (Richardson 2007). Lack of exposure to exercise, limited the risk of bias associated with inherent perception (Kendall 2003) and placebo effect (Kaptchuk 1998; Bialosky *et al.* 2011) and its impact on outcomes. Finally, as a weakness of the study, it does not take patient learning/practice effect on self-reported outcomes (Bordens and Abbott 2002) and the lack of reporting of MCID prevented the contextual interpretation of the outcomes (Bahreini *et al.* 2020). Assessor blinding was reported as those who "collected data were kept blinded to previous assessment data and final score calculations at subsequent assessment times to avoid bias". The authors acknowledge in their limitations, the lack of completely blinded observers and the need for future research to address this.<br><br>The study utilised ANOVA, multivariate general linear hypothesis, between group (self-selected rate of participation) continuous variable, within variable measures were categorical. Given the statistics utilised, it appears that the authors anticipated a specific outcome to which the analysis could be applied, thereby affecting the reliability of the outcomes. This may be amplified in that the study did not employ power analysis for sample size determination, underpowered the study and potentially amplifying concerns around skewed data (Kendall 2003; Christley 2010). This contrasts with the rigour with which assessment periods were interspaced with the intervention. The equidistant assessment periods during the intervention period, adds control over risk of bias related to the timing of measures in relation to the treatment procedure (Nandgude, Hasabe and Kolsure 2018), condition in which measures were taken (Richardson 2007), and time-modified confounding (Kaur, White and Bartold 2012; Walker *et al.* 2013). Despite the multiple assessment periods and the drop-out rate (Fogg and Gross 2000) of participants allowing for intention-to-treat analysis, the study did not make use of this (Gupta 2011). |
| Outcome | The study showed there was evidence that 25 participants undertaking a 12-week low-impact aerobic exercise programme, experienced a significant decrease in self-reported outcome for fatigue: particularly those participants attending moderate to high frequency of exercise group. The aerobic fitness of all participants showed improvement following participation in the planned exercise programme. Finally, the study showed no indications of the studies exercise programme exacerbating disease activity. As a result, the study concludes that findings are suggestive of statistical improvement (clinical improvement would require comparison to MCID values) as a result of exercise programme in aerobic fitness, grip strength of the hands, the 50-foot walk-time, and pain levels. The exercise programme could have contributed to both the reduction in levels of fatigue in RA patients while increased energy resources(Neuberger *et al.* 1997). |
| Discussion | With the study representing a case series, its structure inherently is more pragmatic and reflective of clinical practice, however this had left this study open to bias based on the small sample size (and subsequent group sizes), the choice of statistical analysis and control for outliers, in addition to the influencers that manifest in the form of RA variability, patient variability, exercise participation variability, the factors affecting self-reporting (Hawthorne, treatment novelty, patient naivety and placebo) and lastly impact of context and recruitment.<br><br>In addition, the study is not able to discern the outcomes of any one modality in the clinical improvement or regression of the patient with RA in this case series given that the measures measured the outcomes of the progressive programme. It is however possible to indicate that this study does suggest that there is benefit from a programme including strengthening and aerobic exercise. This, as the author points out does require further study through a process of a RCTs or n-RCTs in order to allow a better measure of the intervention effect (i.e., to determine whether this is greater than chance or natural history alone within the context of this case study). |
| Conclusion: | Therefore, in conclusion for this particular study, it is evident from the reviewers' responses (collective 75% agreement) that this study was able to fulfil a majority of the applicable criteria in the Liddle scale and that the reviewers were in agreement for the presence of a moderate degree of bias. This in addition to the limitations noted in this table imply that the results of this are subject to a moderate to high degree of bias. This implies that the outcomes are not necessarily reproducible despite the lack detail to reproduce the intervention and that the outcomes are not necessarily due to any one intervention alone. As a result, the outcomes of the study cannot be determined to be better than chance or natural history alone. As a result, the study conclusions must be seen with caution regarding the influence of bias and confounding.<br><br>Therefore, at best, this particular study provides moderate to low evidence in support of the intervention programme being used in clinical practice and that further studies in this particular area are warranted to validate the outcomes and then also to determine the contribution of each of the aspects of the exercise programme in order to allow for more specific recommendations to be made to practitioners in the field. |

Table 4.67a: Tabulated feedback of reviewer data

| Author/s | Perlman S.G., Connell K.J., Clark A., Robinson M.S., Conlon P., Gecht M.G., Caldrom P. and Sinacore J.M. | | | | | |
|---|---|---|---|---|---|---|
| Year | 1990 | | | | | |
| Title | Dance-based aerobic Exercise for rheumatoid arthritis. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | B2 | B1 | B1 | B1 | 67% |
| | 2. What percentage of individuals refused to participate? | B1 | B1 | B1 | B1 | 100% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B1 | B1 | B2 | B1 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B1 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | A | B1 | B1 | B1 | 67% |
| | 7. Is the analysis by intention-to-intervene (treat)? | C | A | C | C | 67% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B2 | As outlined in Table 4.67b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B2 | As outlined in Table 4.67b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | There is not specific ethical reason a RCT could not be carried out. | | | | |
| | 4. Include any other comments | Many of the post tests were not completed by all subjects. This impacted on the outcomes as no rigorous assessment of the data was performed to determine if those that did were an accurate reflection of the total group. | | | | |
| | | | | | Overall percentage agreement | 79,375% |

Table 4.67b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Perlman S.G., Connell K.J., Clark A., Robinson M.S., Conlon P., Gecht M.G., Caldrom P. and Sinacore J.M. |
| Year | 1990 |
| Title | Dance-based aerobic Exercise for rheumatoid arthritis. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| - Joint count (swelling and pain on motion or tenderness)<br>- AIMS<br>- 50-foot walk time<br>- POMS | Pre/Post for each programme run | Five intervention programmes over two years, with each intervention programme lasting 16 weeks. | 53 total starting sample<br>Completers 43 (81%)<br>Loss to follow-up of 10 (only 2 participants had reasons given) | No | N/A | N/A | The study scored four B1, three N/A and one C. This outcome represents that most criteria applicable to this study in the Liddle scale are mostly fulfilled. | 79% |
| VAS (patients perceived impact of arthritis on QoL) | Pre/Post used for the last two-programmes | | | | | | | |

| | |
|---|---|
| Limitations | The participants for this study were drawn from the Chicago metropolitan area via media and referral from health care professionals, family and / or friends (Wells *et al.* 2012). Recruitment required that patients met the following criterion: completion of a self-report of physical medical history, a letter of approval from the primary physician, confirmation of RA diagnosis (criteria not stated), no history of surgery six-months prior to study enrolment, ambulatory (with/out walking aids), no impairment of hearing and willingness to participate in outcome assessments. To supplement this a review of medical records was undertaken by the research staff, noting precautionary notes from the primary physician and medication history/prescription. Many of the participants had several medical conditions. Despite the variability in the participants constituting the sample, most were deemed suitable for the study, the only exceptions were participants with heart rates above 100 and blood pressure of 160/95 or greater.<br><br>The demographic information of the participants who completed the study (n=43) were recorded as a single collective. The data included sex distribution, ethnicity, age, ARE functional class and education status. The non-homogeneity between participants in respect of the RA dynamic characteristics influences the generalisability of the study results. Given, the sample comprised of 72% Caucasian females, 84% were above 40 and 33% above 60, “most” classified as ARE class I or II and 77% achieved post school level education; indicates that the outcomes of this study would be applicable to this group only. The variance of the RA dynamic characteristics and variable presentation of comorbidities, further obscures that outcomes and regresses improvement noted to the mean. The pre-test evaluation sample included 53 participants of which only 43 (81%) completed the intervention with a post-test evaluation.<br><br>The programme comprised five interventions conducted over two years, each intervention programme was 16 weeks in duration (16 weeks x 5 = 80 weeks), which does not seem probable given the time available for the entire study. The intervention, devised by a multidisciplinary team was a supervised dance-based exercise programme, including exercise and education/discussion sessions referred to as EDUCIZE (warm-up, followed by aerobic exercise and then mat work)) (Robins, Hernán and Siebert 2004). The discussion was used to enhance integration of the skills and concepts introduced via the programme. EDUCIZE was group based conducted in two hours sessions twice a week over 16 weeks. The sessions comprise of one hour of exercise and an hour discussion period, to improve flexibility, strength, posture, balance and coordination(Heine *et al.* 2012).<br><br>Outcome measures were assessed at baseline and at the end of the 16-week intervention and included a good combination of objective and subjective outcomes, thus limiting the effects of bias related to patient naivety, intervention novelty, Hawthorne effects as well as placebo effects (Mouton 2006; Richardson 2007; Moustgaard *et al.* 2014; Sedgwick and Greenwood 2015; Snibsøer *et al.* 2018). However, given that this study only had one group that was defined by pre-post measures and the lack of a power analysis means that the study may be under powered in being able to draw firm conclusions in addition to the analysis being limited to only student’s *t* tests without comparison to a partner group or comparator (Fogg and Gross 2000; Kendall 2003; Christley 2010). This is also identified by the authors in that they indicate that RCTs are needed to confirm their findings whilst also limiting the confounding factors present in this study. |
| Outcome | The analysis of the data suggested no deleterious effects on RA disease activity. The reports indicate that physician-assessed articular pain, swelling, 50-foot walk time, pain, and depression decreased significantly. There was significant improvement in lower extremity function, vigour, and fatigue with perceptions of general health improved |

| | |
|---|---|
| | significantly. Thus, the authors concluded that there was synergism between dance-based aerobic exercise and problem-solving discussion in the treatment of RA patients (Perlman *et al.* 1990). |
| Discussion | With the structure of a case series being more pragmatic and reflective of clinical practice, it is more likely to suffer from bias. In this study, the lack of comparability between participants at baseline as well as limited control over the dynamic RA factors and demographic variables, prevents clear analysis of confounding derived from these variables, but more importantly the present variability prevents the reader excluding natural history improvement and chance acting in this study. The disproportionate representation with in the already small sample size, may compromise the external validity, as factors like ethnicity and sex potentially influences response to the intervention, and thus the patient to whom the intervention can be applied is limited. The study reports change in outcome from a potentially underpowered baseline that suffers dropout during the intervention. The design prevents intention-to-treat analysis. These, risk mis-reporting the effectiveness of the intervention, aside from the fact that the study has no control group to offer a point of reference for the outcome. The intervention is implemented in a programme; supervised (dance) exercise (was modified to meet the participants limitations – changing the protocol of intervention into tailored therapy with a common broad framework), while psychological support and education supplemented the exercise within this programme. This programme design preventing the reader capacity to define component contribution to outcomes noted, and whether synergistic or antagonistic relationships exist between these components. Despite the above, the study reports to have achieve significant outcomes through this programme. However, this requires further investigation, particularly as the authors acknowledges aspects of bias and confounding highlighted in the analysis this review carried out. Synergistic effect is one of them (Robins, Hernán and Siebert 2004). It is recommended that future investigations should be considered and make effort to control for the confounders and bias risks, by starting with a structured RCTs with carefully considered controls to negate pitfalls. |
| Conclusion: | Therefore, in conclusion for this particular study, it is evident from the reviewers' responses (collective 79% agreement) that this study was mostly fulfil the applicable criteria in the Liddle scale. The reviewers agree that it is at low to moderate risk of bias.<br><br>When we consider this with the limitations noted in this table, this results in the study being ranked as subject to as low-to-moderate risk of bias. This implies that the outcomes are likely to be reproducible and that the outcomes are likely to be due to the interventions programme alone despite the lack of clarity around the synergistic/detractive effect of each part of the programme. As a result, the study conclusions have to be seen with caution regarding the influence of bias and confounding. Therefore, this study's intervention programme collectively could be considered to provide low to limited evidence of its efficacy. and that an RCT would be necessary with a more inclusive sample to determine the true effect of the intervention (in its parts as well as, as a programme) in RA, provide validity to the outcomes, and allow for specific recommendations to practitioners in the clinically setting. |

Table 4.68a: Tabulated feedback of reviewer data

| Author/s | Williams, S. B., Brand, C. A., Hill, K. D., Hunt, S. B. and Moran, H. | | | | | |
|---|---|---|---|---|---|---|
| Year | 2010 | | | | | |
| Title | Feasibility and Outcomes of a Home-Based Exercise Program on Improving Balance and Gait Stability in Women With Lower-Limb Osteoarthritis or Rheumatoid Arthritis: A Pilot Study. | | | | | |
| Criteria | | Reviewer 1 | Reviewer 2 | Reviewer 3 | Ranking | Percentage Agreement (%) |
| EVALUATION CRITERIA FOR THE STUDY: | 1. Are the study participants well-defined in terms of time, place, and person? | A | A | B1 | A | 67% |
| | 2. What percentage of individuals refused to participate? | B1 | N/A | B1 | B1 | 67% |
| | 3. Are outcomes measured in a standard, valid and reliable way? | B1 | A | B1 | B1 | 67% |
| | 4. Are outcomes measured in the same way for both intervention and control groups? | N/A | N/A | N/A | N/A | 100% |
| | 5. Are factors other than the intervention e.g., confounding factors, comparable between intervention and control groups and if not comparable, are they adjusted for in the analysis? | N/A | B2 | N/A | N/A | 67% |
| | 6. What percentage of individuals recruited into the study are not included in the analysis? (loss to follow-up). | A | A | A | A | 100% |
| | 7. Is the analysis by intention-to-intervene (treat)? | C | A | N/A | - | 0% |
| | 8. Are results homogeneous between sites? (multicentre/multisite studies only). | N/A | N/A | N/A | N/A | 100% |
| OVERALL ASSESSMENT OF THE STUDY: | 1. How well was the study done to minimise bias? IF coded as B1, B2 or C, what is the likely direction in which bias might affect the study results? | B1 | As outlined in Table 4.68b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 2. Is the overall effect of the study due to the study intervention? | B1 | As outlined in Table 4.68b, there are multiple points of bias as a result of confounding factors that play a role in obscuring the results, collectively moving the reported outcomes away from the true intervention effect. | | | |
| | 3. Explain if there is any practical/ethical reason why an RCT cannot be done. | There is no practical or ethical reasoning that an RCT could not be conducted. The study proves feasibility of the intervention. | | | | |
| | 4. Include any other comments | This study sample was made up of two conditions OA and RA, with the minority of the participants made up of RA. The study failed to report a subgroup analysis for the RA specific participants. | | | | |
| | | | | | Overall percentage agreement | 71% |

Table 4.68b: Properties of the study, outcome, and discussion

| | |
|---|---|
| Author/s | Williams, S. B., Brand, C. A., Hill, K. D., Hunt, S. B. and Moran, H. |
| Year | 2010 |
| Title | Feasibility and Outcomes of a Home-Based Exercise Program on Improving Balance and Gait Stability in Women With Lower-Limb Osteoarthritis or Rheumatoid Arthritis: A Pilot Study. |

| Study properties: | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Form of measurement | Frequency of measurement | Duration of study | Number of participants | Assessors blinded | Control used | Randomisation of participants | Ranking | Total Percentage Agreement |
| Fall risk: FROP-Com assessment tool – modified.<br>Balance measures: the Clinical Test of Sensory Interaction on Balance – modified, the Functional Reach Test, the Step Test, the NeuroCom Balance Master (long plate), the Chattecx Balance System.<br>Leg muscle power measures: the sit to stand test (on the NeroCom Balance Master (long plate).<br>Gait measures: the NeuroCom Balance Master (long plate)<br>Clinical measures of gait: gait velocity, timed up and go test.<br>Related measures: BMI, site of lower limb pain and most affected joints, number of falls in the past 12 months (baseline) and past 4 months (4th month assessment), lower-limb pain (VAS), activity levels (Human Activity Profile), Falls self-efficacy (the Modified Falls Efficacy Scale, Average Modifed Falls Efficacy Scale). | Pre/post-test design. (Baseline and four months) | Four-month intervention | 49 total only 39 were eligible and completed. | No | No | No | Two A's and two B1's with three N/A's and one undefined scoring.<br>The applicable criterion to the study ranked entirely to mostly fulfilled. | 71% |

| | |
|---|---|
| Limitations | Recruitment was carried out by means of a multi-centred, convenience sampling method, carried out in private rheumatology clinics and public hospitals, in Australia. This raises the perception that the sample consists of patients who have previously sought specialist care, implying that their RA disease activity and severity may be greater than in the case of the average RA patient (Wells *et al.* 2012). And in this raises the naivety of the sample into question.<br>When the inclusion criteria were reviewed, it was noted that all patients were female, limiting generalisability to female RA patients. Generalizability was further limited by the use of mixed groups (OA and RA) with regard to the lower limbs, as well as by the use of the old 1987 criteria (Arnett *et al.* 1988). By contrast, the exclusion criteria were more stringent and covered the following: not having lower-limb arthritis, being bed-bound, having Parkinson's disease, a stroke, multiple sclerosis, a history of cardiac syncope or epilepsy, having undergone lower-limb surgery within prior year, or having intra-articular viscosupplementation or a corticosteroid injection in the prior six months. Nevertheless, the criteria failed to specify the confounders such as co-morbidities, patient demographics and RA factors such as age, disease duration, medications, serology type and any previous therapeutic effort. Lastly, the recruitment included patients from both public hospitals and private rheumatology clinics, but it was not specified that the distribution of these patients was equitable between the groups. Additionally, patients with no medical support and utilising public resources are generally known to have more severe symptoms owing to the delay in receiving treatment, which may have decreased the homogeneity between the groups.<br>The fact that recruitment took the form of convenience sampling may have inadvertently created bias, thus not allowing the study to measure the intervention effectively and reducing the generalisability of the outcomes. Furthermore, the paucity of baseline assessment reporting specifically for OA or RA patients in terms of the outcome measures prior to the start of the study negates the ability to detect whether the groups were representative of the RA population. This does not make for confidence as to whether the concerns raised in the previous paragraph were controlled for. Credit must be given in that an a priori power revealed that 37 participants per group would have had the power of 0.8 to detect an effect size of 0.5 at the 0.05 level on two of the outcome measures (Step Test and Timed Up & Go); however, this choice of outcomes negates the possibility that the sample size is also applicable to other outcome measures. In terms of the repeated outcome measures, the pre/post design only provided for one post- assessment over four months, thus negating the ability to control dropout effect and to use intention-to-treat analysis; moreover, the design does not document changes in clinical outcomes over the time period, which are unlikely all to occur at the same time (Gupta 2011). In the study's favour, assessment was conducted in a controlled laboratory setting but to the study's disadvantage medication changes, among them pain medication, DMARDs and others commonly utilised in treating RA, and their effects were not documented, as was the case with the influences of changes in the home environment in response to the study. Finally, the FROP-Com tool and Clinical Test of Sensory Interaction on Balance (CTSIB) were not utilised in a manner that is suggested in the literature (Shumway-Cook and Horak |

| | |
|---|---|
| | 1986; Russell *et al.* 2008), therefore its ability to reflect the outcomes accurately in its modified format is debatable. This was compounded by the CTSIB outcomes for some patients reportedly being affected by the ceiling effect (Coster, Ludlow and Mancini 1999; Liang 2000; Coster 2013) By contrast; the NeuroCom Balance Master did not suffer the same limitations as the previous tools but was excluded from the a priori power analysis (Kendall 2003; Christley 2010). The gait velocity was tested on six meters of the 10-meter walkway which as a test is pragmatic, but has not shown validity, particularly when an indoor gait aid was allowed (Coster 2013; Bahreini *et al.* 2020). Furthermore, the sit to stand test, for which a standard-sized wooden box was used, was good in terms of the study but was of variable proportion to the patients, potentially affecting their outcomes on the basis not only of vectors of movement but also the effect of body weight. Thus, although the study described assessments well the authors' ability to reflect the outcomes is problematic. In terms of the subjective outcome measures, pain measures were implemented well, although it was unclear as to whether the reporting was a composite or for specific joints. The fall measures in the past 12 months overlapped in terms of the time period considered, thus potentially under-reporting changes with the intervention; they may also have been biased by memory recall and issues around the Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015). Little detail is provided for in terms of the activity levels (Human Activity Profile), limiting the assessment of this tool. Lastly, it was also noted that some outcomes were affected by equipment availability or malfunctioning, which queries whether these units were calibrated before the study and whether the outcome measures were reliable. This led to missing data in some instances, decreasing the computed numbers of sample readings which may as a result have been less the 37 required by the a priori analysis.<br><br>In terms of the intervention, the exercise programmes were individualized home exercise programmes, which implies that this was really a report on a case series of patients. Furthermore, the programmes were derived from the Otago exercise programme and the Visual Health Information Exercise Prescription Kits which are cited as programmes relevant to falls in uncompromised patients (in other words, having no disease) and not specifically RA or OA patients (Campbell *et al.* 1997). This implies that the length of time provided for these programmes should perhaps have been longer than the noted four months to attain better measurable outcomes for patients with disease. It is also suggested that the programmes were based on pre-study assessment findings, comorbid conditions, safety, and the endurance of participants and that they included some recommended exercises that may not have been considered for all patients or that some patients would have considered additional. Although it makes practical sense that the assessor of the patients at entry designed the intervention programmes, it does not exclude a level of bias as the intake assessor also determined the group allocation. Additionally, the same person also reviewed the intervention programme at two home visits (four and eight week) and modified, ceased, or progressed the prescription according to progress with modification of exercise to patient challenge in order to ensure appropriate exercise completion. This complexity in the programmes, in addition to the participants retaining previous exercise activities, provides further implications that this is a case series reporting. The contacts to ensure adherence and continued support (at patient request) also mean further variability in terms of access to the physiotherapist which should have been documented in an adherence diary to better keep track of variability. Given the variability of the inputs into the study, there would have been a need for Bonferroni adjustments to account for outliers in the data. This is particularly the case as the reported 67% exercise adherence rate, although higher than most studies, would have been a significant confounder in this already variable patient group. |
| Outcome | At baseline, 64% of the sampled participants reported having suffered falling during the 12 months preceding the study. The sample averaged a fall risk score of 14.5 according to the Falls Risk of Older People (Community Setting), 42% of the participants rated with a moderate risk for falls (16–23). Following the four-month intervention participants achieved improvement in performance of most balance outcomes and related measures. These included improvements in falls risk with a $p$=0.01, an improvement in activity levels ($p$=0.015), a reduction in fear of falling ($p$=0.022), an improved functional reach test ($p$=0.001), improved index for sit to stand ($p$=0.001), increased step width during walking ($p$=0.001) and improved BMI ($p$=0.006). According to these findings, the individualized home-based balance and exercise programme was shown to be feasible in implementing for the older female patients with OA or RA. They results indicate the intervention may improve the stability during walking and "other functional activities" (Williams *et al.* 2010). |
| Discussion | This being a case study, its structure is one that is more pragmatic and reflective of clinical practice. Thus, it is not able to discern the outcomes of any one modality in the clinical improvement or regression of the RA patient. It is, however, possible to indicate that this study does seem to suggest that there may be benefits from an individualised home-based exercise programme that includes the use of balance, walking and strengthening exercises customised to the individual needs and capacity of the patient. The programme incorporated exercise undertaken by the participant prior to joining the study with physiotherapist support available telephonically.<br><br>This intervention, as the author points out, requires further investigation (via RCT or nRCT) in order to allow for a better measure of the intervention effect, particularly when taking limitations into consideration (i.e., a focused investigation on RA/OA with a larger sample size including both male and female participants to enhance generalisability). thus it is difficult to ascertain whether the comparisons to MCID changes in this context are actually clinically meaningful to the RA / OA patient (Bahreini *et al.* 2020). |
| Conclusion | It is evident from the reviewers' responses (collective 71% agreement) that this study was able to mostly fulfil a majority of the applicable criteria in the Liddle scale, resulting in the reviewers finding for a low to moderate degree of bias. This, in addition to the limitations noted in this table, indicates that the study is at a moderate to high risk of bias. Therefore, the study provides limited evidence in support of an intervention programme that could be clinically beneficial. Future studies are warranted to validate the outcomes and determine the contribution of the intervention programme to RA or OA specifically in order to allow for more specific recommendations to be made to practitioners in the field. |

## 4.5 Conclusion:

Chapter 4 outlines the findings reveal the outcomes of the internal and external reviews conducted on the 65 articles included in this systematic review (40 RCTs, 15 nRCTs, 10 CS/CR), critically appraising the methodological rigour, both internal (viz. PEDro (PEDro 1999) / NOS (Wells, Shea and O'Connell 2003) / Liddle scales (Liddle, Williamson and Irwig 1996b)) and external, as is seen in the limitations discussion. The outcomes of each study as reported in the publication are then presented in tabular form and discussed by the researcher, in terms of the collective analysis of their data. In each Table, the conclusion summarises the evaluation of the article's contribution to evidence for the given interventions. Chapter 4 thus paves the way for Chapter 5 which overviews the evidence for the use of manual therapy, massage, and exercise as adjunctive therapies for RA patients.

# CHAPTER 5: DISCUSSION

## 5.1 Introduction

Building on Chapter 4, which presents the review of each of the 65 studies that encompass manipulation, mobilisation and exercise for the RA patient, highlighting the factors that affect the internal and external validity of the studies outcomes and allowing the level of evidence to be evaluated for or against the tested interventions, Chapter 5 presents the evidence in composite form for each of the modalities of manipulation, mobilisation and exercise in the 65 studies so that the cumulative evidence available within a domain can be understood.

For the purposes of this systematic review, the 65 articles have been grouped according to the modality tested (manipulation, mobilisation, and exercise); and then further according to whether the modality was tested on its own or within a programme of care. This latter distinction is necessary because within a programme of care, it is not possible to determine whether the exercise, manipulation, mobilisation, or any of the other interventions used, is solely or collectively responsible for the positive or negative clinical outcomes obtained in the study. As a result, the level of evidence in these instances reflects the composite intervention programme.

## 5.2 Discussion of the criteria utilised in ranking the evidence

Within the purview of the literature, there are various evidence-grading systems that may be used to rank evidence (Chapter 3 Section 3.10). In the field of manual therapies / conservative therapies, the most commonly applied grading systems include those of:

- AGREE (Dagenais and Haldeman 2011).
- GRADE (Brozek et al. 2009).
- Grading suggested by Foley et al. (2003).

When the above three systems are compared, it is evident that Foley *et al.* (2003) and Dagenais and Haldeman (2011) use a grading of evidence that is simpler and therefore more readily understandable to health care providers and others who may consult manual therapy systematic reviews. The Cochrane GRADE system is more complex, requiring a more detailed knowledge of the grading system to allow for application in clinical practice (Brożek *et al.* 2009), while the AGREE (Dagenais and Haldeman 2011) and Foley *et al.* (2003) systems are more pragmatic in application. Furthermore, both the Foley *et al.* (2003) and Dagenais and Haldeman (2011) include a category in the review of evidence that indicates whether the evidence provided in the literature accords with high, moderate, low or no evidence. Additionally, the Foley *et al.* (2003) system allows for the review to report whether this evidence agrees or conflicts in respect of a specific modality.

This means that, unlike the GRADE system, the AGREE (Dagenais and Haldeman 2011) and Foley *et al.* (2003) grading systems may collectively be able to determine whether there is good evidence for bad clinical outcomes or good evidence for good clinical outcomes with respect to a specific modality, which allows for better contextualisation of the outcomes of a systematic review.

The Dagenais and Haldeman (2011) and Foley *et al.* (2003) systems of evidence grading can best be summarised as follows:

| | |
|---|---|
| Strong Evidence: | One specific area (in the case of this systematic review one modality) represented by two or more RCTs of at least fair / moderate quality constitutes strong evidence. |
| Moderate Evidence: | One specific area (in the case of this study one modality) represented by at least one RCT of at least 'fair/moderate' quality constitutes moderate evidence. |
| Limited Evidence: | The findings reinforced by one non-experimental study (e.g., nRCT, CS and cohort study) are graded as limited evidence. |

Consensus and conflict of evidence:

| | |
|---|---|
| Consensus of Evidence: | Agreement between the data from different studies but in respect of the same modality showing that the results for a specific intervention are consistent (i.e., the outcomes are consistently in favour of or against a specific intervention) constitutes consensus of evidence. |
| Conflicting Evidence: | Incongruence found between at least two RCTs (Dagenais and Haldeman 2011) constitutes conflicting evidence. Alternatively, inconsistency within the designated study types, where the conclusion is required to be drawn from the majority results, either positive or negative, and the stronger study types are needed to determine the outcome if they themselves are not conflicted, is graded as conflicting evidence. |

No evidence

| | |
|---|---|
| No Evidence: | In a study of low/poor quality the lack of any clear evidence to support the treatment regimen is graded as no evidence. |

The following section will review the outcomes attained in Chapter 4 and present the evidence available for each of the composite areas that represent specific intervention groups.

## 5.3 A review of interventions for RA patients and evidence

On the strength of the reviews, the following single interventions were tested on the RA population: exercise (specifically strength training), soft tissue mobilisation (e.g., massage) and joint mobilisation. In the following tables a brief overview of the findings associated with each of these is presented as an encapsulated summary of the work presented in Chapter 4.

### 5.3.1 Single interventions

#### 5.3.1.1 Exercise (Strength training)

| Table 5.1: Intervention: Exercise (Strength training) | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Hoenig, Groff, Pratt, Goldberg and Franck | Hand | 1993 | RCT | 6/11 | Limited | Inconclusive | | | | | | |
| Rønningen and Kjeken | Hand | 2008 | nRCT | 6/10 | Limited to none | Positive | | | | | | |
| Brorsson, Hilliges, Sollerman and Nilsdotter | Hand | 2009 | nRCT | 5/10 | Moderate to limited | Positive | | | X | | | X |
| Cima, Barone, Porto, and de Abreu | Hand | 2013 | RCT | 6/11 | Moderate to limited | Positive | | | | | | |
| Piva, Khoja, Toledo, Chester-Wasko, Fitzgerald, Goodpaster, Smith and Delitto | Muscle function | 2019 | RCT | 8/11 | Moderate to limited | Positive | | | | | | |
| McMeeken, Stillman, Story, Kent, and Smith | General | 1999 | RCT | 10/11 | Limited | Positive | | | | | | |
| Kapale, Vardharajulu and Warude | General | 2017 | CS/CS | Three C, three N/A, and three non-agreement | Limited to none | Positive | | | | | | |

Interestingly, the testing of strength exercises only portrays a case study later in the development of this field of study within RA, where typically this occurs ahead of more rigorous nRCT and RCT studies. However, this may be because the case study in question presented evidence for free exercise, (a form of exercise therapy that relies on muscular effort with external resistance supplied by gravity alone (Leelayuwat 2017)), which is different from the other forms of strength exercise studied prior to this point and may provide the impetus for further study in this field.

Table 5.1 provides a summary for the use of strength exercises in assisting RA patients. This field of study has been dominated by RCT studies of moderate to high quality based on the internal validity review and with positive clinical outcomes that are complemented by two nRCT studies of moderate quality that also have positive clinical outcomes. This configuration of study outcomes would suggest that a strong level of evidence for strength exercises collectively exists according to the grading of evidence by Dagenais and Haldeman (2011) and Foley *et al.* (2003).

This, however, is not the case for the general application of strength exercises, which at best reflects a moderate to limited level of evidence according to the same grading parameters (Foley *et al.* 2003; Dagenais and Haldeman 2011). Similarly, the combination of studies for hand strength exercises exclusively only finds a moderate level of evidence available to support the use of this modality in addressing the hand symptoms of RA patients (Foley *et al.* 2003; Dagenais and Haldeman 2011), which

is not dissimilar to the single RCT that focusses on the improvement of muscle function (Foley *et al.* 2003; Dagenais and Haldeman 2011).

The above conclusions are only based on the internal validity scores as reflecting the methodological rigour of the articles that were reviewed; they exclude the impacts of factors not measured by the scales but having a potential bearing on the outcomes of the studies. Given the external validity limitations, many of these studies thus fall short in respect of adequately controlling for various aspects of bias not measured by the scales. Principal sources of bias not controlled for include the following: failure to accurately identify sufficient participant numbers (lack of a power analysis) (Kendall 2003; Christley 2010), failure to articulate the RA patient characteristics of the sample under study (including socio-demographic and medical variables) and compare them to the population they represent (see Section 4.4.1.1) and, in addition, failure to ensure that the groups used to test the intervention are homogenous (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015) and that the interventions are not pragmatically applied (i.e., individualised care for participants within a group) (Heine *et al.* 2012). These factors open the study to external influence or bias and detract from its ability to test the intervention accurately (Bartley and Fillingim 2013; de Boer *et al.* 2015). Thus, as can be seen in the evidence contribution column in Table 5.1, the expected contribution of each article is lower than would have been expected with a review of the internal validity only. Additionally, the fact that most of the studies show positive results warrants the consideration that there may be an influence of publication bias (Callaham *et al.* 1998; Knobloch, Yoon and Vogt 2011) in that studies with unfavourable outcomes may not be reflected fairly in the literature (owing to not having been publication).

Consequentially, equal value needs be given to a study's ability to represent a research design's rigorous methodology and its inherent ability to reflect true outcomes without bias. This, together with the possibility of publication bias, means that with regard to strength exercises this systematic review errs on the side of caution, suggesting that the evidence for strength exercise collectively is limited, even though there seems to be consensus that the clinical effect thereof tends to be positive for the patients.

In short, it is not possible in terms of either the Dagenais and Haldeman (2011) and Foley *et al.* (2003) evidence scale or the conservative approach of this systematic review to state conclusively that the outcomes attained in these strength exercise studies will result in the same or a similar outcome for a patient in clinical practice. It is of concern that a number of the studies did not compare the changes attained in the clinical outcomes to their corresponding published MCIDs (Hoenig *et al.* 1993; Brorsson *et al.* 2009; Cima *et al.* 2013), so that, even though the studies attained statistical significance, it is unclear whether this resulted in clinically measurable and effectual change for the participants. Therefore, the evidence concerning strength exercise indicates that this should be used with caution and that no guarantees can be given patients in terms of expected outcomes (particularly in terms of a more generalist clinical approach as opposed to hand strength exercises alone).

In order to provide for more accurately informed patient consent, the RCT studies previously conducted in the field of strength exercise should be repeated with due consideration of limitations and points of bias as noted in the Tables 4.7, 4.21, 4.28 and 4.32, where the limitations of each of the studies' have been identified and discussed.

### 5.3.1.2 Soft tissue mobilisation

| Table 5.2: Intervention: Soft tissue mobilisation | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Field, Diego, Delgado, Garcia and Funk | Upper Limbs | 2013 | nRCT | 3/10 | No evidence | Positive | | | | | | |
| Metin and Ozdemir | Pain and fatigue | 2016 | RCT | 6/11 | Limited | Positive | | | | | | |
| Cubick, Quezada, Schumer and Davis | General | 2011 | CS/CS | Two B1 and six N/A | Limited | Positive | | | X | | | X |
| Almuhaysin | General | 2019 | CS/CS | Two A score, one B1 score and five N/A | Moderate | Positive | | | | | | |

Interestingly, further to the testing of soft tissue mobilisation, the publication of the two case studies above (both general in their approach) occurred before and after the RCT (with upper limb focus) and the nRCT (focussing on pain and fatigue).

Table 5.2 thus provides a summary of the results of tests investigating the use of soft tissue mobilisation in assisting RA patients. This field of study has been dominated by case studies of variable quality based on the internal validity review and thus achieving positive clinical outcomes. This evidence for positive outcomes is complemented by the two studies reviewed here (the nRCT of poor quality and the RCT of moderate quality), both also with positive clinical outcomes to the studies. This configuration of study outcomes would suggest that, in terms of the Dagenais and Haldeman (2011) and Foley *et al.* (2003) evidence-grading system, at best a moderate level of evidence for soft tissue mobilisation collectively exists.

Taken individually by focus area, the relevant studies in Table 5.2 indicate that this is, however, not the case for soft tissue mobilisation, which at best reflects a moderate evidence with the focus on pain and fatigue according to the same grading parameters (Foley *et al.* 2003; Dagenais and Haldeman 2011). By contrast, the use of soft tissue mobilisation for the upper limb finds no more than a limited level of evidence available for the use of this modality (Foley *et al.* 2003; Dagenais and Haldeman 2011), which is not dissimilar to the case studies which applied soft tissue mobilisation more generally (Foley *et al.* 2003; Dagenais and Haldeman 2011).

It must be noted that the above conclusions are only based on the internal validity scores as reflective of the methodological rigour of the articles that were reviewed, thus excluding the impacts of factors not measured by the scales but having a potential bearing on the outcomes of the studies. Given the external validity limitations, many of these studies therefore fall short in respect of adequately controlling for various aspects of bias not measured by the scales. These principally include the failure to accurately identify a sufficient sample size (Kendall 2003; Christley 2010), to articulate the RA patient characteristics (see Section 4.4.1.1) of the population/sample under study, as well as to ensure that the groups used to test the intervention are homogenous (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015) and that the interventions are not pragmatically applied (i.e., individualised care for participants within a group) (Heine *et al.* 2012). These potential sources of bias in the studies detract from the studies' capacity to test the intervention without undue external influence from these sources and therefore accurately (particularly with regard to the case studies). Thus, as can be seen in the evidence contribution column in Table 5.2, the expected contribution of each article is lower than would have been expected when reviewed by the internal validity alone. Additionally, with most of the studies showing positive clinical results, it needs to be considered that publication bias may favourably influence the conclusions drawn in this systematic review. For this reason, research efforts should balance and equally value both the structure of a strong internal validity and the inherent ability of the outcome measures to accurately report true intervention effects without bias (external validity).

The above, together with the possibility of publication bias; means that in the case of soft tissue mobilisation this systematic review errs on the side of caution, suggesting that the evidence for soft tissue mobilisation collectively is limited, even though there seems to be consensus that the clinical effect tends to be positive for the patients.

From a clinical vantage point, however, whether the outcomes attained in these soft tissue mobilisation studies are drawn according to the AGREE and Foley *et al.* (2003) evidence scale or the conservative approach of this systematic review, it is not possible to state conclusively that these outcomes will result in the same or a similar outcome for a patient in clinical practice. This is of particular concern in that some of the studies did not compare the changes attained in the clinical outcomes to published MCID values, so that, even though the studies attained statistical significance, it remained unclear whether this resulted in clinically measurable and effectual change for the participant/patient (Bahreini *et al.* 2020). Therefore, the evidence with regard to soft tissue mobilisation indicates that this modality should be used with caution and that no guarantees can be presented to patients in terms of expected outcomes (particularly in the context of a more generalist clinical approach and therapy focussing on the upper limbs, by contrast with which pain- and fatigue-focussed therapy seems to have stronger evidence in support of the use of soft tissue mobilisation).

In order to provide for more accurately informed patient consent, the RCT and nRCT studies previously conducted should be repeated with due consideration given to controlling for the limitations and points of bias as noted in the Tables 4.30 (RCT) and 4.44 (nRCT), where each study's limitations have been identified and discussed. Additionally, the CS/CR should have their outcomes confirmed or refuted through a formally structured trial.

### 5.3.1.3 Joint mobilisation

| Table 5.3: Intervention: Joint mobilisation | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Dhondt, Verbruggen, Oostendorp and Duquet | Spinal Pain | 1999 | RCT | 8/11 | Moderate | Positive | | | | | | |
| Levitsky, Kisten, Lind, Nordström, Hultholm, Lyander, Hammelin, Gentline, Giannakou and Faustini | Hands | 2019 | RCT | 9/11 | Limited to no evidence | Positive | | | X | | | X |

Whereas Table 5.2, summarising the studies on soft tissue mobilisation, has an RCT with one specific area of focus and one nRCT and two case studies dealing with another, Table 5.3 has two RCTs on joint mobilisation, but each of the two RCT's have a specific focus (one in spinal pain and the other in the hand).

Thus, Table 5.3 provides a summary for the use of joint mobilisation in the treatment of RA patients. The studies are solely RCTs of a high quality in respect of their internal validity review and both report positive clinical outcomes. Together, these studies suggest that, according to the grading of evidence by AGREE and Foley *et al.* (2003), a high level of consensus exists collectively in the evidence for joint mobilisation.

However, bearing in mind that each RCT study has a specific focus area, the RCT studies considered separately can only attain at best a moderate level of evidence for each area, given their high internal validity ranking (Foley *et al.* 2003; Dagenais and Haldeman 2011). When the different focus areas are viewed separately. According to the grading system of Dagenais and Haldeman (2011) and Foley *et al.* (2003) the final CS/CR would only contribute limited evidence.

The above conclusions are, however, only based on the methodological rigour scores from the respective scales used to review the articles. These results are drawn, while overlooking external factors not explicitly measured by these scales that may introduce bias and thus potentially influence the clinical outcomes of the studies (viz., these factors provide external validity limitations). Tables 4.11 and 4.22 demonstrate that consistent themes develop, among them the following: not accurately identifying the number of participants required to achieve valid outcomes for a specific intervention

(Kendall 2003; Christley 2010), not clearly articulating the RA patient characteristics of the population under study (see Section 4.4.1.1), and not ensuring that test groups are homogenous (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015) and that interventions are pragmatically applied (i.e., individualised care for participants within a group) (Heine *et al.* 2012) are evident. These concerns lead to areas of bias that detract from the study's ability to accurately test and measure the intervention without undue influences from these sources (Bartley and Fillingim 2013; de Boer *et al.* 2015). As can be seen in the evidence contribution column of Table 5.3, the expected contribution of each article is in fact lower than would have been expected from the review of the internal validity only. Additionally, with most studies showing positive clinical results, publication bias, which can influence favourable conclusions in this systematic review, needs to be guarded against. For this reason, research efforts should balance and equally value the structure of a strong internal validity and the inherent ability of the outcome measures to accurately report true intervention effects without bias (external validity). Thus, with due regard for the risk of publication bias, this systematic review suggests that the evidence supporting joint mobilisation in both applications (whether for use in the specific treatment of the RA patient's spine or hands).

From a practical and clinical vantage point, therefore, whether the conclusions are drawn on the Dagenais and Haldeman (2011) and Foley *et al.* (2003) evidence scale or via the conservative approach of this systematic review, it is not possible to state conclusively that the outcomes attained in these joint mobilisation studies will result in the same or a similar outcome for a patient in clinical practice. This conclusion is specific for joint mobilisation of the spine and hands.

The concerns articulated in the above paragraph are amplified by the fact that these studies did not compare reported changes attained in the clinical outcomes to published MCID's. Thus, even though they reached statistical significance, it remains unclear whether this resulted in clinically measurable and effectual change for the patient/s. This may be illustrated by the fact that the inconclusive outcomes for the joint mobilisation of the knee are based on statistical results. This, however, does not imply that there was no clinical benefit to the patient, nor is this determinable. Similarly, the statistically significant findings of the other studies in this systematic review do not imply that their reported positive outcomes reflect clinical benefit. Thus, this conundrum is not answerable for the current studies.

In order to provide for more accurately informed patient consent, the RCT studies previously conducted should be repeated with due consideration given to the limitations and points of bias as noted in the Tables 4.11 and 4.22, where each has been identified and discussed.

### 5.3.1 Multi-Modal Interventions:

From the articles found within the systematic review of the literature and the subsequent reviews thereof, the following multimodal programmes were identified: spinal manipulation as part of a programme of care, joint mobilisation as part of a programme of care, exercise and medication, exercise as a programme of care and, lastly, exercise as a set of varied exercises (viz., an exercise programme).

#### 5.3.1.1 Spinal manipulation (part of a programme)

| Table 5.4: Intervention: Spinal manipulation (part of a programme) | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Chung and Mior | Thoracic pain | 2015 | CS/CS | Two B2 and six N/A | Limited | Positive | | | X | | | X |

Table 5.4 provides a reflection of only one article representing spinal manipulation as part of a programme. The research took the form of a case study, the latter rating as offering a limited level of evidence for manipulation as part of an intervention programme according to the grading of evidence by Dagenais and Haldeman (2011) and Foley *et al.* (2003).

The conclusions set out in Table 5.4 are, however, only based on the internal validity scores as reflective of the methodological rigour of the article reviewed, excluding the impact of factors not measured by the scales, but having a potential bearing on the outcomes of the study. Given the external validity limitations, the study done by Chung and Mior (2015) thus falls short in terms of adequately controlling for various aspects of bias not measured by the scales. These principally include but are not limited to the following: not effectively articulating the RA patient characteristics of the population under study (see Section 4.4.1.1) and, additionally, not ensuring that the interventions are spelt out clearly (Heine *et al.* 2012) and that measurements are appropriate and applied at correct time intervals (Kendall 2003; Coster 2013; Bahreini *et al.* 2020). These shortfalls detract from the study's ability to accurately measure the intervention outcomes without undue bias. Thus, as indicated in the evidence contribution column in Table 5.4, the study provides limited evidence in support of the intervention, which is not incongruent with the internal validity rating achieved through the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system.

Although possible, it is difficult to determine whether the positive outcomes articulated here may have been influenced in this systematic review by publication bias (Callaham *et al.* 1998; Knobloch, Yoon and Vogt 2011). Given that only one such article was sourced, the study's favourable outcomes may not fairly reflect the literature. Even with this consideration, both the systematic review outcomes and the outcomes are per the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system tend towards a limited degree of evidence for the use of spinal manipulation in a programme.

In view of the limited evidence and the fact of having only one publication from which to draw conclusions, it is not possible to state conclusively that the outcomes attained in this treatment programme incorporating spinal manipulation would result in the same or similar outcomes for a patient in clinical practice. This is further complicated in that case studies do not generally compare the changes attained in the clinical outcomes of the case they published to MCID (Hoenig *et al.* 1993; Brorsson *et al.* 2009; Cima *et al.* 2013). Thus, even if statistical significances were computed (again unusual in a case study), it is unclear whether this would translate into clinically measurable and effectual change for the patient. Therefore, the evidence regarding manipulation as part of a treatment plan should be considered with caution the patient should be offered no guarantees of expected clinical outcomes.

This clear deficiency of research in this specific area highlights that for accurate informed consent to be provided, the case study should be appraised as forming the foundation for a future RCT or pragmatic nRCT. For this recommended future study to provide a rigorous methodological, high quality (in terms of internal and external validity) and clinically significant contribution to the repository of available knowledge. Due consideration should be given to the identification and addressing of this case study's limitations. This would result in stronger evidence able to confirm or refute the evidence provided by this preliminary case study.

### 5.3.1.2 Mobilisation (part of a programme)

Table 5.5: Intervention: Mobilisation (part of a programme)

| | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Romanowski, Špiritović, Romanowski and Straburzyńska-Lupa | Knee pain | 2020 | RCT | 9/11 | Limited to no evidence | Inconclusive | | | | | | |
| Almuhaysin | General | 2019 | CS/CS | Two A score, one B1 score and five N/A | Moderate | Positive | | | X | | X | |

When comparing Table 5.4 on manipulation as part of a treatment protocol with Table 5.5 on mobilisation as part of a treatment protocol, it becomes evident that there is one RCT more in the collective evidence bundle to be considered for mobilisation.

Thus, Table 5.5 provides a summary for the use of joint mobilisation as part of an intervention programme in the treatment of RA patients. The studies reflect one RCTs of high quality in respect of methodological rigour but are inconclusive clinical outcomes. This is complemented by one case study with positive clinical outcomes. This collection of studies suggests that there is limited conflicting evidence for joint mobilisation, according to the grading of evidence by Dagenais and Haldeman (2011) and Foley *et al.* (2003).

Additionally, considering that the RCT study has a specific focus area anatomically (the knee), it would be more appropriate to consider these studies separately in order for the evidence for each focal area to be fairly considered. In the light of the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system, it would seem that mobilisation of the knee within a programme could only attain limited to no evidence, with strong reliance on the specific study's high internal validity scores. The CS/CR only contributes a limited amount of evidence for the general application of mobilisation in a programme according to the grading system by Dagenais and Haldeman (2011) and Foley *et al.* (2003).

The above conclusions are, however, only based on a review of the internal validity / methodological rigour scores from the respective scales used to review of the articles. This results in the exclusion of factors that influence the study but are not explicitly measured by the scales. It is, however, reasonable to consider these factors (i.e., factors that provide external validity limitations), given their potential to influence the clinical outcomes of the studies. Table 4.22 show that the following themes are consistently present: the lack of an *a priori* analysis (Kendall 2003; Christley 2010) and the lack of clear descriptive outlines of the RA patient characteristics under study (see Section 4.4.1.1) as well as the lack of clear homogeneity between the test groups (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015) and defined intervention protocols (Heine *et al.* 2012) are present. These themes are shared by the RCT studies and open them to external confounder/s and / or bias, which detracts from the ability of these studies to accurately test and measure the intervention. Thus, as can be seen in Table 5.5, under the evidence contribution column, the expected contribution of each article is in reality lower than expected from the review scores of internal validity alone.

A further confounder is to be seen in the outcome's column of Table 5.5, which denotes the clinical outcomes of the study according to the author/s. Only the CS/CR showed positive clinical outcomes, as compared to the RCT. These outcomes may indicate that publication bias is limited in this grouping of articles reflecting mobilisation treatment as part of an intervention programme. However, it may also be reflective of the fact that other studies may not have been published due to negative clinical outcomes. Collectively, this form of publication bias tends to influence favourable conclusions in systematic reviews and is something that needs to be guarded against. Therefore, research efforts should balance and equally value the structure of a strong internal validity, and the inherent ability of the outcome measures to accurately report true intervention effects without bias (external validity). Consequently, taking this approach into account, the present systematic review suggests that joint mobilisation as part of an intervention programme [in all applications (for use in the general RA treatment or treatment of the RA patient's knees)] is limited, and specifically inconclusive for use in knee RA.

From a practitioner's vantage point, the clinical and practical value, whether the conclusions are drawn on the Dagenais and Haldeman (2011) and Foley *et al.* (2003) evidence scale or via the conservative approach of this systematic review; is limited. It is not possible to state conclusively that the outcomes attained by joint mobilisation as part of an intervention programme in these studies will result in the same or similar outcomes for a patient in a clinical setting (specifically for the general application of joint

mobilisation). By contrast, the use of joint mobilisation for the knees as part of a larger intervention programme is at this point very limited and without clinical benefit. Therefore, patients cannot be given any expectation regarding improvement for this type of intervention.

These clinical and practical limitations articulated in the above paragraph are amplified in that the studies in Table 5.5 did not provide comparison of the clinical changes that they reported to published minimally clinically important differences. Thus, even though they attained statistical significance in some instances, it is unclear whether this resulted in clinically measurable and effectual change for the patient. By way of illustration, the inconclusive outcomes for the joint mobilisation of the knee as part of an intervention programme are based on statistical analyses but this does not imply that there was no clinical benefit to the patient (in terms of the changes recorded and reported). This is, however, not determinable in the context provided by the study. By contrast, the statistically significant changes in clinical outcomes for the other studies may not convert to clinical benefit. Resolution for this conundrum is not possible for these studies and thus it is recommended that the evidence in regard of joint mobilisation as part of a greater intervention programme be used with caution and that no guarantees be presented to patients in terms of expected outcomes.

In order to provide more accurately informed patient consent, the RCT studies previously conducted should be repeated with due diligence to limitations and points of bias as noted in the Tables 4.22. Additionally, the CS/CR should have their outcomes confirmed or refuted through a formally structured trial, with attention given to the study's limitations identified and discussed in Table 4.59.

### 5.3.1.3 Exercise (aerobic) with acetaminophen/placebo

| Table 5.6: Intervention: Exercise (aerobic) + acetaminophen/placebo | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Meeus, Hermans, Ickmans, Struyf, Van Cauwenbergh, Bronckaerts, De Clerck, Moorken, Hans and Grosemans | Pain | 2015 | RCT | 8/11 | Moderate | Positive | | **X** | | | | **X** |

Table 5.6 provides the solitary article representing the effects of aerobic exercise and medication use (acetaminophen) in the treatment of pain in RA patients. The study took the form of a RCT that rates as a moderate level of evidence for aerobic exercise and medication use (acetaminophen) according to the grading of evidence according to Dagenais and Haldeman (2011) and Foley *et al.* (2003). From the above evidence, a conclusion based on the internal validity score provided by the PEDro scale was drawn in reporting the review of the study's methodological rigour. It must be considered that this review excluded factors that have a potential influence on the clinical outcomes of the study.

RA presents a significant challenge for researchers in that it has various permutations in its clinical presentation, which affects the outcome/s of studies when inadequately controlled for. This RCT fell short in adequately controlling for this and the RA participant characteristics (see Section 4.4.1.1), as well as RA dynamic characteristics (Brahee, Pierre-Jerome and Kettner 2003; So *et al.* 2003; Rindfleisch and Muller 2005; Longmore *et al.* 2014; Ridgley, Anderson and Pratt 2018; Sharma *et al.* 2018) and, additionally, did not use an *a priori* analysis to determine an appropriate sample size (Kendall 2003; Christley 2010) or ensure that the interventions were spelt out clearly (Heine *et al.* 2012) and that measurements were valid, reliable and contextually appropriate (Coster 2013; Bahreini *et al.* 2020) and applied at correct time intervals (Pater *et al.* 1998; Platt, Schisterman and Cole 2009). These shortfalls detract from the study's ability to accurately measure the intervention outcomes without undue bias (Bartley and Fillingim 2013; de Boer *et al.* 2015). Thus, the evidence contribution column in Table 5.6 presents that the study provides moderate evidence, which is congruent with the internal validity rating achieved through the Dagenais and Haldeman (2011) and Foley *et al.* (2003) rating.

Although possible, it is difficult to determine whether the positive outcomes articulated here may influence this systematic review in terms of publication bias (Callaham *et al.* 1998; Knobloch, Yoon and Vogt 2011). Given that only one such article was found, the study's favourable outcomes may not fairly reflect the literature. Even with this consideration, both the systematic review outcomes and the outcomes derived from the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system tend

towards a moderate degree of evidence for the use of aerobic exercise and medication use (acetaminophen) in the treatment of pain in RA patients.

Given the moderate evidence and the fact of only one publication from which conclusions are drawn, it is not possible to state conclusively that the outcomes attained in this treatment programme incorporating aerobic exercise and medication use (acetaminophen) in the treatment of pain in RA patients would result in the same or similar outcomes for a patient in clinical practice. This is further complicated in that this RCT did not compare the clinical outcomes attained against published minimally clinically important differences (Hoenig *et al.* 1993; Brorsson *et al.* 2009; Cima *et al.* 2013) for the respective outcomes. For this reason, even though the study attained statistical significance, it is unclear whether this resulted in clinically measurable and effectual change for the patient. Therefore, the evidence concerning the combination of aerobic exercise and medication use (acetaminophen) in the treatment of pain in RA patients should be considered with caution and no guarantees can be presented to patients in terms of expected clinical outcomes.

In order to facilitate more comprehensive and informed patient consent, the RCT should be closely evaluated, and its limitations identified and addressed (see Table 4.29) in order to allow a stronger study to be able to confirm or refute the evidence as provided by this study.

#### 5.3.1.4 Exercise (part of a programme)

In reviewing Table 5.7 below, it becomes apparent that exercise as part of a programme in the treatment of RA has been extensively researched. This is reflected in the 17 RCTs, eight nRCTs and four CS/CR included in this systematic review.

| Table 5.7: Intervention: Exercise (part of a programme) | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Lowman | Rheumatoid cripple | 1958 | nRCT | 2/10 | None | Positive | | | | | | |
| Lonta | Early RA | 1960 | CS/CS | One B1, one C and six N/A | Limited | Positive | | | | | | |
| Hawkes, Care, Dixon, Bird and Wright | Hands | 1986 | nRCT | 5/10 | Limited | Inconclusive | | | | | | |
| Kauppi, Leppänen, Heikkilä, Lahtinen, and Kautiainen | Atlanto-axial | 1998 | CS/CS | Two B1, three A and three N/A | Limited | Positive | | | | | | |
| Buljina, Taljanovic, Avdic, and Hunter | Hand | 2001 | RCT | 8/11 | Moderate | Positive | | | | | | |
| Oh and Seo | Pain and depression | 2003 | nRCT | 3/10 | None | Positive | | | | | | |
| O'Brien, Jones, Mullis, Mulherin and Dziedzic | Hand | 2006 | RCT | 8/11 | Moderate | Positive | | | | | | |
| Eversden, Maggs, Nightingale and Jobanputra | QoL | 2007 | RCT | 8/11 | None | Positive | | | | | | |
| Manning | Upper limb | 2013 | RCT | 8/11 | High to moderate | Positive | | | | | | |
| Williams, Williamson, Heine, Nichols, Glover, Dritsaki, Adams, Dosanjh, Underwood and Rahman | Hand | 2015 | RCT | 9/11 | High | Positive | | | | | | |
| Hamilton, Bywaters and Please | General | 1959 | RCT | 6/11 | Limited to none | Positive | | | X | | | X |
| Spiegel, Spiegel, Ward, Paulus, Leake and Kane | General | 1986 | nRCT | 4/10 | Limited to none | Inconclusive | | | | | | |
| Perlman, Connell, Clark, Robinson, Conlon, Gecht, Caldron and Sinacore | General | 1990 | CS/CS | Four B1, three N/A and one C | Moderate | Positive | | | | | | |
| Noreau, Martineau, Roy and Belzile | General | 1995 | nRCT | 4/10 | No evidence | Inconclusive | | | | | | |
| Vljeland, Zwinderman, Vandenbroucke, Breedveld and Hazes | General | 1996 | RCT | 8/11 | Moderate to limited | Positive | | | | | | |
| Bell, Lineker, Wilkins, Goldsmith and Badley | General | 1998 | RCT | 11/11 | Moderate | Positive | | | | | | |
| Scholten, Brodowicz, Graninger, Gardavsky, Pils, Pesau, Eggl-Tyl, Wanivenhaus and Zielinski | General | 1999 | RCT | 7/11 | Limited to none | Positive | | | | | | |
| Scholten, Brodowicz, Graninger, Gardavsky, Pils, Pesau, Eggl-Tyl, Wanivenhaus and Zielinski | General | 1999 | nRCT | 6/10 | Limited to none | Positive | | | | | | |
| Westby, Wade, Rangno and Berkowitz | General | 2000 | RCT | 7/11 | Moderate to limited | Inconclusive | | | | | | |
| Van den Ende, Breedveld, Le Cessie, Dijkmans, De Mug and Hazes | General | 2000 | RCT | 8/11 | Moderate | Positive | | | | | | |
| Lineker, Bell, Wilkins and Badley | General | 2001 | RCT | 6/11 | Limited | Positive | | | | | | |
| Chui, Lau and Yau | General | 2004 | CS/CS | One B1, one B2, one C, two I and three N/A | Limited | Positive | | | | | | |
| Brodin, Eurenius, Jensen, Nisell, Opava and Group | General | 2008 | RCT | 7/11 | Moderate | Inconclusive | | | | | | |
| Baillet, Payraud, Niderprim, Nissen, Allenet, Francois, Grange, Casez, Juvin and Gaudin | General | 2009 | RCT | 9/11 | Moderate | Inconclusive | | | | | | |
| Hurkmans, Van den Berg, Ronday, Peeters, Le Cessie and Vliet Vlieland | General | 2010 | nRCT | 4/10 | Limited to none | Inconclusive | | | | | | |
| Breedland, van Scheppingen, Leijsma, Verheij-Jansen and van Weert | General | 2011 | RCT | 8/11 | Moderate to limited | Inconclusive | | | | | | |
| Gizińska, Rutkowski, Romanowski, Lewandowski and Straburzyńska-Lupa | General | 2015 | nRCT | 3/10 | No evidence | Positive | | | | | | |
| Shinde and Varadharajulu | General | 2017 | RCT | 6/11 | Moderate to limited | Positive | | | | | | |
| Garner, Fenton, Martin, Creaser, Johns and Barnabe | General | 2018 | RCT | 6/11 | Limited to none | Positive | | | | | | |

The studies are dominated by nine RCTs of a moderate to high quality in respect of their methodological rigour but with variable clinical outcomes. This is complemented by eight RCTs of moderate methodological rigour and several nRCTs and CS/CR of variable methodological rigour quality. A collective pooling of the studies using the grading of evidence by Dagenais and Haldeman (2011) and Foley *et al.* (2003) suggests that a high level of conflicting evidence for exercise as part of a programme exists.

However, considering that ten of these studies focus on specific anatomical regions or clinical outcomes, it may be inappropriate to consider the collective evidence as opposed to the evidence per region. Thus, with view to the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading, it would seem that:

- Treatment of the hands with exercise as part of a programme is supported by at least three RCTs of moderate to high quality, indicating that in this regard exercise as part of a programme has a high / strong level of evidence in support of its use. There is, however, one nRCT which has inconclusive outcomes. Given that it performed averagely on the NOS scale, this may be based on external influences on the outcomes. Therefore, although strong evidence exists, it is necessary to consider the limitations of the studies to determine if any of them are affected by external bias concerns.
- The remaining specific areas have, according to the Dagenais and Haldeman (2011) and Foley *et al.* (2003), limited evidence in support of the use of exercise as part of a programme for either upper limb RA presentation, addressing the quality of life, pain and depression in a RA patient, or addressing clinical outcomes in early RA and /or the RA cripple.

By contrast, in the case of the application of exercise as part of a programme for the RA patient more generally, a total of 12 RCTs, five nRCTs and two CS/CR of variable internal rigour and variable clinical outcomes contribute to the evidence available. According to the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading, this repository of evidence provides strong evidence in favour of the use of exercise as part of a programme with clinically beneficial outcomes. This is based on the requirement that at least three RCTs have a moderate to high reviewer rating, with positive outcomes.

However, the above conclusions are only based on a review of the internal validity / methodological rigour scores from the respective scales used for the reviewed articles. This process results in the exclusion of factors that may influence the study but are not explicitly measured by the scales. It is therefore reasonable to consider these factors as they have the potential to influence the clinical outcomes of the studies (e.g., these factors provide external validity limitations).

With reference to Tables 4.2, 4.3, 4.4, 4.5, 4.14, 4.18, 4.23, 4.34, 4.35, 4.37, 4.39, 4.40 (RCTs), Tables 4.45, 4.48, 4.52, 4.56 (nRCTs), and 4.60, 4.67 (CS/CR), it is evident that themes of relating to a lack of an *a priori* analysis (Kendall 2003; Christley 2010), a lack of clear descriptive outlines of the RA patient characteristics under study (see Section 4.4.1.1), lack of clear homogeneity between the test groups (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015) and a lack of definition

of intervention protocols (Heine *et al.* 2012) are present. These concerns detract from the studies' ability to accurately test the intervention without undue influences from these sources (Bartley and Fillingim 2013; de Boer *et al.* 2015). Thus, as can be seen in the evidence contribution column in Table 5.7, the expected contribution of each article is lower than would have been expected from the review of the internal validity alone.

A further confounder in reviewing these articles is seen in the outcome's column of Table 5.7, which denotes the clinical outcomes of the study according to the authors. Most of the studies (21 of 29 (72,4%)) show positive clinical outcomes, which may reflect to a greater degree the presence of publication bias as opposed to actual evidence in favour of exercise as part of an intervention programme. This form of publication bias tends to influence favourable conclusions in systematic reviews and is something that systematic reviews need to guard against (Callaham *et al.* 1998; Knobloch, Yoon and Vogt 2011). The fact that eight articles are inconclusive should be of concern to the reader, particularly with regard to clinical expectations, as approximately 27,6% of the outcomes (8 of 29) are inconclusive and do not provide evidence in favour for or against the changes in clinical outcomes.

There may be an argument that these results stem from external factors that place influence on the study (Bartley and Fillingim 2013; de Boer *et al.* 2015). Therefore, it is necessary to balance against each other and place an equal value on a study's ability to present a good internal validity and its inherent ability to measure the outcomes without bias (external validity). With this approach in mind, this systematic review suggests that evidence for exercise as part of an intervention programme for RA patients is limited and seemingly conflicted.

From a practitioner's vantage point, the clinical and practical value of exercise as part of an intervention programme is limited. It is not possible to state conclusively that the outcomes attained in these exercises as part of an intervention programme will result in the same or similar outcomes for a patient in clinical practice (specifically with regard to the general application of exercise as part of a programme and with specific regard to upper limb RA presentation, addressing the QoL, pain and depression in a RA patient as well as clinical outcomes in early RA, and / or the RA cripple). The only area that seems to have strong evidence is exercise as part of an intervention programme for the hands of RA patients. Therefore, patients cannot be given any expectation of clinical improvement for this type of intervention, other than when it is applied to patients in addressing RA symptoms of the hands.

Notwithstanding the above, it must be remembered that the clinical and practical limitations articulated in the above paragraph are based on statistical analyses and outcomes only and that the studies did not provide comparison of the clinical changes to reported MCID. Thus, even though they attained statistical significance in some instances, it is unclear whether this actually resulted in clinically measurable and effectual change for the patient (Bahreini *et al.* 2020).

In order to provide for more accurately informed patient consent, the RCT studies previously conducted, should be repeated with due diligence to limitations and points of bias as noted in the Tables 4.2, 4.3,

4.4, 4.5, 4.14, 4.18, 4.23, 4.34, 4.35, 4.37, 4.39, 4.40 (RCTs), Tables 4.45, 4.48, 4.52, 4.56 (nRCTs), and 4.60, 4.67(CS/CR), where each study's limitations have been identified and discussed. For the anatomically or clinically specific studies for which there was limited evidence, these too should be repeated with due diligence to their noted limitations as outlined in Chapter 4 (Tables 4.6, 4.13, 4.26, 4.31, 4.41 (RCTs), Tables 4.47, 4.51, 4.53 (nRCTs), and Tables 4.64, 4.65 (CS/CR)).

#### 5.3.1.5 Exercise (mixed)

In reviewing Table 5.8 (on following page), it becomes apparent that exercise as part of a programme in the treatment of RA has been extensively researched. This is reflected in the fact that 14 RCTs, five nRCTs and two CS/CR on the subject were identified for this systematic review.

| Table 5.8: Intervention: Exercise (mixed) | | | | | | | COLLECTIVE EVIDENCE | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | STRONG | MODERATE | LIMITED | NO EVIEDNCE | CONFLICTED | CONSENSUS |
| **Authors** | **Region treated** | **Year** | **Study type** | **Internal validity** | **Evidence contribution** | **Outcomes** | | | | | | |
| Dellhag, Wollersjö and Bjelle | Hands | 1992 | RCT | 7/11 | Moderate | Positive | | | | | | |
| Mannerkorpi and Bjelle | Shoulder | 1994 | RCT | 5/11 | Limited to none | Positive | | | | | | |
| Häkkinen, Sokka, Kotaniemi and Hannonen | Strength, disease activity, functional capacity, BMD | 2001 | RCT | 6/11 | Moderate to limited | Positive for clinical Negative for BMD | | | | | | |
| de Jong, Munneke, Lems, Zwinderman, Kroon, Pauwels, Jansen, Ronday, Dijkmans and Breedveld | BMD | 2004 | RCT | 9/11 | Moderate to limited | Positive | | | | | | |
| Häkkinen, Sokka, Kautiainen, Kotaniemi and Hannonen | BMD / strength | 2004 | RCT | 7/11 | Moderate | Positive | | | | | | |
| Williams, Brand, Hill, Hunt and Moran | Lower limb | 2010 | CS / CS | Two A and B1, one C and three N/A | Limited | Positive | | | | | | |
| Durcan, Wilson, and Cunnane | Sleep / fatigue | 2014 | RCT | 7/11 | Moderate to limited | Positive | | | | | | |
| Romanowski, Straburzyńska-Lupa, Romanowska and Lorenc | Knee pain | 2014 | nRCT | 2/10 | None | Positive | | | | | | |
| McManus, Visker and Cox | Sleep | 2015 | RCT | 4/11 | Limited to none | Inconclusive | | X | X | | X | |
| Law, Saynor, Gabbitas, Jones, Kraus, Breslin, Maddison and Thom | Large joints | 2015 | nRCT | 6/10 | Moderate | Negative | | | | | | |
| Harkcom, Lampman, Banwell and Castor | General | 1985 | RCT | 9/11 | Limited | Positive | | | | | | |
| Hansen, Hansen, Langgaard and Rasmussen | General | 1993 | RCT | 7/11 | Moderate to limited | Inconclusive | | | | | | |
| Lyngberg, Harreby, Bentzen, Frost and Danneskiold-Samsøe | General | 1994 | RCT | 7/11 | Limited | Positive | | | | | | |
| Van den Ende, Hazes, le Cessie, Mulder, Belfor, Breedveld and Dijkmans | General | 1996 | RCT | 7/11 | Moderate | Positive | | | | | | |
| Häkkinen, Mälkiä, Häkkinen, Jäppinen, Laitinen and Hannonen | General | 1997 | RCT | 5/11 | Moderate to limited | Positive | | | | | | |
| Neuberger, Press, Lindsley, Hinton, Cagle, Carlson, Scott, Dahl and Kramer | General | 1997 | CS / CS | One A, three B1, one C and three N/A | Moderate to limited | Positive | | | | | | |
| de Jong, Munneke, Zwinderman, Kroon, Jansen, Ronday, van Schaardenburg, Dijkmans, Van den Ende and Breedveld | General | 2003 | RCT | 9/11 | Moderate to limited | Positive | | | | | | |
| Strasser, Leeb, Strehblow, Schobersberger, Haber and Cauza | General | 2011 | RCT | 7/11 | Moderate | Positive | | | | | | |
| Häkkinen, Hannonen, Nyman, Lyyski and Häkkinen | General | 2003 | nRCT | 6/10 | Moderate | Inconclusive | | | | | | |
| Karatepe, Günaydin, Türkmen and Kaya | General | 2011 | nRCT | 1/10 | Limited to none | Positive | | | | | | |
| Zernicke, Kedor, Müller, Burmester, Reißhauer and Feist | General | 2016 | nRCT | 5/10 | Limited | Positive | | | | | | |

The studies in Table 5.8 are dominated by eight RCTs of a moderate to high quality in respect of their methodological rigour but with variable clinical outcomes (although predominantly positive). These are complemented by six RCTs of moderate to limited methodological rigour and a group of seven nRCTs and CS/CR of variable methodological rigour quality and variable clinical outcomes. This collection of studies suggests that a high degree of conflicting evidence exists for mixed exercise protocols in the treatment of RA patients, based on the grading of evidence by Dagenais and Haldeman (2011) and Foley *et al.* (2003).

However, considering that ten of these studies focus on specific anatomical regions or clinical outcomes, it may be inappropriate to consider the collective evidence as opposed to the evidence per region. Thus, with view to the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system, it would seem that the following can be concluded:

- The studies focussing on bone mineral density comprise the only area boasting two RCT studies of moderate to high methodological rigour, thus supporting a high level of evidence. However, this outcome is not unanimous as the studies have conflicting outcomes, with two of their outcomes indicating change and one no changes. Thus, there is a high level of conflicting evidence that requires further study
- In terms of strength, both studies were rated as moderate in terms of methodological rigour, and both found for improvement in strength. Thus, there is a high level of evidence for positive clinical outcomes with mixed exercise protocols in the treatment of RA patients.
- The remaining areas of focus in the studies in Table 5.8, including hands (positive clinical outcomes) and impact on large joints (negative clinical outcomes), all had moderate evidence in support of their respective intervention/s according to the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system.
- The clinical outcomes of the two RCTs dealing with sleep/fatigue, one well-structured and the other poorly structured, are variable, thus this area requires further study to allow appropriate evidence grading.
- The outcomes of the RCTs, case series and nRCTs dealing with the regions of the shoulder, lower limb and knee respectively were poorly rated and therefore provided at best limited evidence in support of the use of mixed exercise protocols in the treatment of RA patients. These areas require further studies to confirm or refute the findings currently available.

With regard to the application of mixed exercise for the RA patient more generally, Table 5.8 indicates that seven RCTs, three nRCTs and one CS/CR of variable internal rigour and variable clinical outcomes contribute to the evidence available. According to the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system, this repository of available knowledge would support strong evidence in favour of the use of mixed exercise protocols in the treatment of RA patients. This is based on the requirement that at least three RCTs have a moderate to high reviewer rating, with positive outcomes.

However, the above conclusions, derived solely from the Dagenais and Haldeman (2011) and Foley *et al.* (2003) grading system, are only based on a review of the internal validity / methodological rigour

scores from the respective scales used to review the articles and as a result factors that influence the study but are not explicitly measured by the scales are excluded. It is therefore reasonable to consider these factors, as they have the potential of influencing the clinical outcomes of the studies (e.g., these factors influence the external validity).

With reference to Tables 4.9, 4.15, 4.19, 4.20, 4.24, 4.36, 4.38 (RCTs), Tables 4.46, 4.49, 4.57 (nRCTs), and Table 4.66 (CS/CR), it is evident that themes relating to a lack of an *a priori* analysis (Kendall 2003; Christley 2010), a lack of clear descriptive outlines of the RA patient characteristics under study (see Section 4.4.1.1), a lack of clear homogeneity between the test groups (Pocock *et al.* 2002; Kendall 2003; Bland and Altman 2011; de Boer *et al.* 2015) and a lack of definition of intervention protocols (Heine *et al.* 2012) are present. These themes are areas in each of the studies that open the study to external influence or bias, which detract from the study's ability to test and measure the intervention accurately. In Table 5.8 the evidence contribution column reflects the level of evidence for each of the included studies; however, their contribution as a result of the above external validity limitations is in reality lower than would have been expected with a review of the internal validity only.

A further confounder in reviewing these articles is seen in the outcome's column of Table 5.8, which denotes the clinical outcomes of the study according to the authors. The majority of studies (17 of 21 (80,9%)) show positive clinical outcomes, which to may reflect the presence of publication bias as opposed to actual evidence in favour of mixed exercise protocols in the treatment of RA patients. This form of publication bias tends to influence favourable conclusions in systematic reviews and is something that systematic reviews need to guard against (Callaham *et al.* 1998; Knobloch, Yoon and Vogt 2011). The fact that four articles are inconclusive should sound a warning bell with regard to the clinical expectations as three of 21 articles in this category (14,2%) report outcomes that are inconclusive and do not provide evidence in favour of or against the changes in clinical outcomes while the remaining two articles report negative outcomes (one reporting some negative outcomes as well as certain positive outcomes).

It may be argued that these results stem from external factors that place influence on the study (Bartley and Fillingim 2013; de Boer *et al.* 2015). For this reason, research efforts should balance and equally value the structure of a strong internal validity and the inherent ability of the outcome measures to accurately report true intervention effects without bias (external validity). In the light of this approach, this systematic review suggests that evidence for mixed exercise protocols in the treatment of RA patients is moderate to limited and seemingly conflicted.

From a practitioner's vantage point, therefore, the clinical and practical value of exercise as part of an intervention programme is limited. It is not possible to state conclusively that the outcomes attained in the studies investigating these mixed exercise protocols for the treatment of RA patients will result in the same or similar outcomes for a patient in clinical practice, especially with regard to the general application of mixed exercise protocols and the specific areas of sleep / fatigue and the anatomically specific domains of the hands, large joints, shoulder, lower limb and knee. The only areas that seem to have strong evidence are mixed exercise protocols in the treatment of RA patients for BMD and strength

development. Therefore, patients cannot be given any expectation with regard to clinical improvement for this type of intervention, other than when it is applied to patients in addressing RA symptoms of the hands.

Notwithstanding the above, it must be remembered that the clinical and practical limitations articulated in the above paragraph are based on statistical analyses and outcomes only and that the studies did not provide comparison of the clinical changes to reported MCID. Thus, even though they attained statistical significance in some instances, it is unclear whether this resulted in clinically measurable and effectual change for the patient.

In order to provide more informed patient consent, the studies of non-specific / general regions, should be repeated with due emphases on addressing limitations and points of bias as noted in Tables 4.9, 4.15, 4.19, 4.20, 4.24, 4.36, 4.38 (RCTs), 4.46, 4.49, 4.57 (nRCTs) and 4.66 (CS/CR), where each of the study's limitations has been identified and discussed. For those studies with specific focus areas (anatomical or clinical) where there was limited evidence, these too should be repeated with due diligence to their noted limitations as outlined in Chapter 4 (Tables 4.10, 4.8, 4.12, 4.14, 14.6, 4.25, 4.27 (RCTs), 4.54, 4.50 (nRCTs) and 4.68 (CS/CR)).

## 5.4 Conclusion of Chapter 5

With reference to the preceding discussion of each of the intervention types included in this systematic review, Table 5.9 provides a summary.

| Table 5.9: Evidence for intervention summary | Level of Evidence of Intervention | | | | | |
|---|---|---|---|---|---|---|
| | **Strong** | **Moderate** | **Limited** | **Consensus** | **Conflicting** | **No evidence** |
| **Single interventions** | | | | | | |
| Exercise (strength) | | | X | X | | |
| Soft tissue mobilisation | | | X | X | | |
| Joint mobilisation | | | X | X | | |
| | | | | | | |
| **Multiple interventions** | | | | | | |
| Spinal manipulation (part of a programme) | | | X | X | | |
| Mobilisation (part of a programme) | | | X | | X | |
| Exercise (aerobic) + acetaminophen | | X | | X | | |
| Exercise (part of a programme) | | | X | X | | |
| Intervention: Exercise (mixed) | | | X | | X | |

Table 5.9 demonstrates that there is limited evidence available for manipulation, mobilisation, massage, and exercise in the treatment of the RA patient. The notable exceptions to these modalities are the use of aerobic exercise with acetaminophen, which has moderate support for the use of the combination to effect positive clinical results for the RA patient.

Additionally, and hidden within the results of the exercise as part of an intervention programme, the treatment of the hands with exercise as part of a programme is supported by at least three RCTs of moderate to high quality, indicating that exercise as part of a programme has a high / strong level of evidence in support of its use.

Furthermore, the use of a mixed exercise programme in the treatment of the RA patient found evidence to support its improvement of bone mineral density (two RCT studies of moderate to high methodological rigour), thus a high level of evidence. Additionally, strength gains in terms of the RA patient's muscle function had a high level of evidence for positive clinical outcomes.

The remaining studies provided very variable methodological rigour owing to the impact of external influencers and variable outcomes that did not provide consistency in clinical outcome to the level required to support them for use in clinical practice other than with caution. Although none of the studies directly reported adverse events related to the treatment interventions, it must be taken into account that several studies fell prey to dropout (with these participants reporting an increased severity in their RA and / or an inability to continue with the intervention as a result of their clinical presentation as their reasons for dropping out).

Given this outcome, particularly with regard to the larger portion of the study, it is perhaps pertinent to consider an excerpt from Oh and Seo (2003), who state:

> *"Relieving pain has been reported to be the most important objective in health education programs for people with RA (Tucker and Kirwan 1991). Nevertheless, whether pain relief is a proper outcome variable in evaluating the effects of an RA intervention programme is debated (Lindroth et al. 1989). Investigators have proposed that improvement in physical ability might be a more proper outcome variable than is pain because pain is an impediment to normal physical activities (Donovan, Blake and Fleming 1989; Lindroth et al. 1989)".*

As suggested by Oh and Seo (2003) above, the variability may be related to the pain and subjective outcomes measures used in the many studies completed in the domain covered by the present systematic review. This reliance on pain as an outcome measure may not be the correct proxy to show patient improvement, RA status and / or disease progression. All these factors were identified as external validity influencers in the study and thus the present systematic review concurs with Oh and Seo (2003) that future studies need to be more circumspect in their choice of outcomes measure not only to attain consistency and allow for later meta-analyses, but also to chart patient progress more accurately.

# CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

## 6.1 Introduction

Concluding this systematic review, Chapter 6 draws together the findings of Chapters 4 and 5 to present recommendations not only for future research and systematic reviews in the field of alternative / adjunctive therapies for treating RA, but also in order to assist with the application of these findings in clinical practice.

## 6.2 Conclusion

As stated in Chapter 1, this researcher set out to determine the level of current evidence for the conservative treatment of RA with particular focus on mobilisation, manipulation, and exercise modalities, commencing with a systematic literature search through which 40 RCTs, 15 nRCTs and 10 CS/CR articles were identified as complying with the inclusion criteria for the systematic review under the domains of mobilisation, manipulation, and exercise respectively. Next, reviewers were appointed to assist the researcher by way of reviewing all articles independently as well as contributing to the hand search if they were aware of any other articles that needed to be included in this systematic review. The purpose of these reviews was to assess the internal validity (methodological rigour) of the articles via the appropriate review scale for the study type presented in each article.

Using these reviews, the researcher in Chapter 4 reported on the internal validity (reviewer assessment) and external validity evaluation of each of the studies, both of which then enabled the researcher to assess the impact of the biases reported through these mechanisms in order to determine their impact on the clinical outcomes reported by the authors. For example, if a study proved to have low internal and external validity, this would indicate a high degree of bias in this study, which in turn implies that the study's clinical outcomes are not necessarily solely due to the intervention and that the outcomes attained in the study cannot be extrapolated into clinical practice as the likelihood for the practitioner to achieve the same outcomes is negligible (viz., the evidence available in the published literature is compromised and thus may not provide appropriate outcomes for the patient).

Based on the above evaluations, the aggregate evidence available within the body of knowledge was then determined using the criteria outlined by both Foley et al. (2003) according to the Foley evidence ranking scale, and Dagenais and Haldeman (2011) via the AGREE scale. These evaluations found that there is limited evidence available for manipulation, mobilisation, massage and exercise in the treatment of the RA patient. These studies demonstrate variable methodological rigour, poor control of external bias, and variable clinical outcomes, thus providing an arena of conflicting studies. As mentioned in Chapter 5, the notable exceptions to the above are the use of aerobic exercise with acetaminophen (moderate evidence supporting positive clinical outcomes), exercise as part of an intervention programme in the treatment of the hands (high level of evidence in support of its use for positive clinical outcomes) and a mixed exercise programme to improve bone mineral density (a high level of evidence

for positive outcomes) and attain muscular strength gains (high level of evidence for positive clinical outcomes).

Therefore, with exception of those indicated in the previous paragraph as having requisite support from the literature, the use of these interventions should be implemented cautiously in practice, and patients should not be provided with expectations that are at odds with the literature currently available. In order to improve this status, the next section of this chapter provides recommendations for improving a systematic review such as this, also for improving the research within the domain of conservative treatment (mobilisation, manipulation, massage and exercise) for the treatment of RA patients and, lastly, for what practitioners could do in practice.

## 6.3 Recommendations

### 6.3.1 Recommendations to Improve This Systematic Review

For systematic reviews, particularly the larger systematic reviews, there is always a challenge to ensure that all reviewers are familiar with the appropriate assessment scales for reviewing the articles (Belur *et al.* 2018). Thus, it is always beneficial to familiarise the reviewers with the appropriate measurement tools. Use of a mock review process (on an article that is not part of the formal research process) would be ideal, thus allowing reviewers unfamiliar with the tools to apply them, to ask questions and to become more familiar with its application. This would aid the process of review further than was done in this study insofar as questions could be asked of the researcher if clarity were required; in addition, the reviewers should receive a document outlining the requirements along with each of the scales (Appendix G; H; and I), as was the case with this systematic review.

Alternatively, the scales could be improved to allow for a simpler and more efficient review process, although it is acknowledged that the use of only one scale would be inappropriate given that this would not fairly evaluate all study types; the biases in each study design are unique and may influence the study design to different degrees in terms of compromising their validity (i.e., the evaluation need to be relative to the study type).

There is an argument in the literature that the outcomes of a systematic review may be biased by language owing to the exclusion of, for example, publications in languages other than English. Although this argument holds value, the counterargument is that articles may lose precision and quality in translation (Scollon, Scollon and Jones 2011) (i.e., even if the research budget were sufficient to allow for article translations). Therefore, other options for the review of non-English articles need to be considered, such as multi-lingual reviewers. This, however, may create other concerns in that potential reviewers within a particular domain may be limited within a given language / cultural context, which then decreases the pool of reviewers since authors tend to be excluded as reviewers because of potential bias towards their own work and potential conflict-of-interest bias towards a peer's publication). This complex conundrum of concerns requires further consideration, particularly in fields where there are few noted authorities who also publish in that domain.

As articulated in Chapter 5, publication bias (Moher *et al.* 2009b; Stern and Kleijnen 2020) may have had an impact on the outcomes of this study, given that most of the studies reviewed for this systematic review indicated positive outcomes with very few (less than 30%) indicating inconclusive results and even fewer indicating negative outcomes to the interventions. For this reason, the researcher has chosen to take a conservative approach in the study and its analyses, preferring to underestimate overall outcomes and advising caution with regard to the use in clinical practice of the interventions under investigation in the articles and studies reviewed here rather than overlooking the concerns relating the various interventions that have emerged in the course of this research.

Additionally, although it is possible that this systematic review may have been subject to inclusion criteria-based bias, nevertheless, in order to limit this risk, attempts were purposefully made to seek out and not exclude grey literature from this systematic review (Hemingway and Brereton 2009; Liberati *et al.* 2009; Moher *et al.* 2009b), as evidenced by the inclusion of Almuhaysin (2019) and Manning (2013), whose dissertations fall within the interest area.

### 6.3.2 Recommendations for Future Studies

Future studies should use the prescribed outlines and / or checklists (e.g., the CONSORT for RCTs) (Schulz, Altman and Moher 2010), STROBE for nRCTs (von Elm *et al.* 2007), and CARE for CS/CR (Gagnier *et al.* 2013) when developing study protocols in order to ensure that the internal validity requirements of methodological rigour are improved. These outlines/checklists include the following considerations (but may not be limited to these, dependent on study type):

- the number of participants per specific study (e.g., an RCT and nRCT should have an *a priori* sample calculation completed to determine study power) (Fogg and Gross 2000; Kendall 2003; Mouton 2006; Rothwell 2006; Christley 2010; Bahreini *et al.* 2020),
- sample recruitment strategies (viz., location and site (Wells *et al.* 2012), sample inclusion / exclusion criteria (serology, functional class, disease duration and disease activity imaging studies (England *et al.* 2019; Myasoedova *et al.* 2020)),
- sample homogeneity (with general RA population and between groups), concealed allocation and randomisation procedures (Fogg and Gross 2000; Pocock *et al.* 2002; Kendall 2003; Mouton 2006; Bland and Altman 2011; de Boer *et al.* 2015; Mansournia *et al.* 2017; Nair 2019),
- a matching outcome measure protocol (Coster 2013),
- and outcome measures (appropriateness, validity, reliability, suitability, sensitivity, and specificity) appropriate to the mechanism or proposed mechanism of the intervention/s (Coster 2013; Bahreini *et al.* 2020).

It should be noted that many of the RA studies reviewed in this systematic review neglected to consider one or more of the following factors in terms of group homogeneity and / or the impact on intervention outcomes, thus significantly hampering external validity:

- The patient factors related to ethnicity, personality, social / cultural background, age, sex and severity of the condition. These were variably reported in the studies, but their impact on pain perception, pain thresholds, and pain sensitisation in turn impacts the outcomes of the studies, particularly those with only subjective outcomes measures that are directly or indirectly linked to pain (Bartley and Fillingim 2013; Bahreini *et al.* 2020). Attention should also be paid to other patient variables such as lifestyle factors (viz., smoking, alcohol consumption, caffeine consumption, parity; menopausal status, pregnancy and one-year post-natal status) that may influence response to treatment and thus impact the outcomes (Pedersen *et al.* 2006a; McInnes and Schett 2011; Gizińska *et al.* 2015; Alpízar-Rodríguez *et al.* 2017; Yang *et al.* 2017; Borba, Zandman-Goddard and Shoenfeld 2018; Okada *et al.* 2019; Myasoedova *et al.* 2020).
- The patient's occupation/work. This may affect outcomes dependent on the treated site. For example, studies on RA of the hand should consider not only whether the patient's occupation is manual or non-manual occupation (Naqvi, Hassali and Aftab 2019) but also the handedness of the patient (Lee, Gandevia and Carroll 2009; Dragert and Zehr 2013).
- The effects of RA medication. Patients' RA medication use can influence their inclusion or exclusion as potential participants in a study as this is one of the factors for determining a level of homogeneity between test groups. A patient's medication also allows for identification of specific RA subpopulations with regard to specific treatments (Smolen, Aletaha and McInnes 2016). In addition, medication factors (such as adherence to and duration of prescription, changes to the prescription prior to or during the study, types of medication, dosage and administration methods, pharmacokinetic and pharmacodynamics characteristics, drug interactions and drug wash-out periods) can also have an impact on the outcomes attained in a study, and thus medication requires monitoring (Zhao *et al.* 2017; Heidari *et al.* 2019). Alternatively, patient stratification within studies needs to be considered (Zhao *et al.* 2017; Heidari *et al.* 2019).
- The monitoring or exclusion of other interventions used simultaneously by patients during their participation in a trial. This improves the measurement of the tested intervention (Mouton 2006). Similarly, surgical intervention (during a period prior to the study and / or planned during the study period), needs clear documentation and consideration as it can potentially negate the intervention effect and influence the analysis of outcome measures for the study, which then impact on the outcomes attributed to the tested intervention (Naqvi, Hassali and Aftab 2019).
- Monitoring the effects of RA comorbidities (e.g., osteopenia, osteoporosis, asthma, depression (Kłodziński and Wisłowska 2018)) and RA flare-ups (seldomly reported on, controlled for or considered for their confounding capacity; of the 65 articles in this review only eight took this into consideration in one form or another (Harkcom *et al.* 1985; Hawkes *et al.* 1986; Hoenig *et al.* 1993; Van den Ende *et al.* 1996; McMeeken *et al.* 1999; Karatepe *et al.* 2011; Cima *et al.* 2013; Romanowski *et al.* 2020)) to control their influence on outcomes attributed to the tested intervention. One manner of achieving this is to stratify participants according to these comorbidities (England *et al.* 2019; Levitsky *et al.* 2019; Myasoedova *et al.* 2020). A similar approach can be taken for concomitantly diagnosed medical conditions that may influence

response to treatment (e.g., diabetes-induced peripheral neuropathy (Bergmann and Peterson 2010; Byfield and Chapman-Smith 2011; Dagenais and Haldeman 2011)).

- Defining exclusion criteria for specific RA patients. This allows the identification of subpopulations that cannot benefit from the tested intervention (Durcan, Wilson and Cunnane 2014).
- Control of patient-related effects, which include the placebo effect (Kaptchuk 1998; Bialosky *et al.* 2011), Hawthorne effect (Mouton 2006; Richardson 2007; Sedgwick and Greenwood 2015), patient naivety and treatment novelty (Mouton 2006), patient blinding (Moustgaard *et al.* 2014), memory recall and memory decay (particularly in subjective outcome studies only) (Mouton 2006; Ricker, Vergauwe and Cowan 2016) and touch therapy (Melzack 1983; Moayedi and Davis 2013; Senderovich *et al.* 2016). Additionally specific to RA the concepts of pain threshold versus pain tolerance and pain sensitisation also need to be considered (Dhondt 1999; Melia *et al.* 2019).

By contrast, researcher effects related to impacting on a study's outcome (which were variably considered in the reviewed studies in Chapter 4) include patient screener / assessor blinding and concealed allocation (Moustgaard *et al.* 2014), blinding of the outcome/s measurer / assessor (Kendall 2003; Moustgaard *et al.* 2014).Also in this category is protocol consistency for both the intervention and the assessment protocols, a crucial factor both in the case of multiple researchers / research assistants and a single researcher because it avoids the use of patient-specific interventions, which detract from the intervention under investigation) (Kendall 2003).

Finally, intervention and assessment protocols need to be consistently re-evaluated to ensure that the most appropriate outcome measures, in other words, measures that are valid, reliable, sensitive and specific for RA patients, are utilised (Coster 2013; Bahreini *et al.* 2020). There also needs to be alignment between the protocols and outcomes measures and the objectives of the study (Coster 2013). With regard to patient factors and patient effects, there should be a balance of objective and subjective outcome measures (Moustgaard *et al.* 2014; Snibsøer *et al.* 2018). Choice of outcome measures should also be dictated by the type of RA patient as different measures will suffer ceiling and floor effects (Liang 2000; Coster 2013) based on patient capability (Liang 2000) because this will impact on the degree of responsiveness to the intervention (Coster, Ludlow and Mancini 1999).

Furthermore, the study's duration needs to be designed around the hypothesised clinical effect of the intervention. If the intervention is suggested to cause immediate changes, then a pre-post study design is appropriate, whereas if a study requires to have changes in fascial or collagen structures (tendons) that are measured, the study then requires at least three months before changes can be clinically detected (Fraser 2008).

With regard to the treatment protocols, these need to be clearly delineated (Heine *et al.* 2012). This detail is critically important in manual therapy as variances in these applications can be possible reasons for the outcomes attained and may influence the application of the exercise protocols for RA patients in practice. For this reason, specific protocol weaknesses need to be identifiable (Heine *et al.* 2012).

The relationship between the intervention and measurement / data collection points needs to be clearly delineated as this impacts on the immediate, short-, medium-, and long-term treatment effect of the single intervention, the protocol, or the programme (Pater *et al.* 1998). The environment in which the outcomes are measured, and the interventions are conducted (whether home-based or in a clinical setting) and the time needed for the intervention to occur are both significant in this respect. In typical exercise interventions, the usual period for clinical change is six to eight weeks; in RA patients, pain inhibition may play a role in extending this minimum time frame. Thus, studies with exercise programmes of four weeks may show no effect for certain modalities, not because of the lack of effect but because insufficient time has been allocated to their exercise programmes, thus insufficient exposure to the interventions will appear to show no effect and the conclusions may then suggest that the exercise programme is not efficacious for RA. While this may be so in terms of the limited time allocated for that intervention, this factor needs to be stated as a limitation in the study (Bell 1998; Breedland *et al.* 2011; Heine *et al.* 2012; Piva *et al.* 2019). This line of thinking then also affects the choice of control as using the patients as their own controls has both biomechanical and neurological influences on the outcomes (Dragert and Zehr 2013; Villafañe, Cleland and Fernandez-de-Las-Peñas 2013). Controls of waitlisted or usual care groups need to be homogenous with the intervention group (as discussed above), but the weigh-up is required because losing the comparative group at the end of the study functionally changes the study's capacity for longevity analysis.

Finally, with regard to the statistical component, which is most influenced by the power of the study (the majority of RA studies are underpowered), the use of *post hoc* analysis to control bias is not recommended (Fogg and Gross 2000; Kendall 2003). Rather, the study should strive by means of inclusion criteria, stratification and / or other methodological means to decrease the need for statistical bias control (de Boer *et al.* 2015). The intention-to-treat analysis should be used as a last resort to generate the missing data / participant numbers needed to attain an adequately powered study (e.g., by making use of dropout data), as this process of artificially calculating expected outcomes is a bias all on its own (Gupta 2011). All these statistical manipulations also affect assumptions of "normal" data distribution in the RA patients that have been included in the study. Therefore, the use of non-parametric statistical procedures that rely on no or few assumptions about the shape or parameters of the population distribution from which the sample was drawn" (Hoskin 2012; Bahreini *et al.* 2020) incurs further bias (Bordens and Abbott 2002).

Despite the complexity of RA, due consideration should be taken to address and attempt to control for all the internal and external validity risks noted above. Efforts need to be directed at creating a vigorous study design to define the efficacy of the modalities covered in this review. In this way a baseline comparative reference point can be formalised to provide context for studies utilising non-natural history control groups. Further to this point, insight can also be gained through an efficacy study into the imparted effects of components within intervention programme designs.

An overview of the available evidence sourced and analysed in this systematic review (Chapter 4 and 5) reveals its sparsity. If this sparsity of evidence contributing to validate or discredit the therapeutic potential of massage, mobilisation, and manipulation in the treatment of RA patients is considered together with the therapeutic potential of these modalities (discussed in Chapter 2, section 2.8), a hypothetical basis of effect is clear. Consequently, this research has opted, as indicated throughout, to err on the side of caution and safety, which is the reason for the single citation of moderate evidence at best, with support from low-grade citations. To validate the modalities, further research is required to provide moderate to high grade evidence to allow for more accurate and confident conclusions to be drawn. Further research efforts are thus strongly recommended in order to demystify these modalities by building a repository of evidence.

Finally, studies in the field of this systematic review should also consider investigating outcome tools and MCID values specific to the general RA population. This would allow for more accurate reporting of the impact the interventions considered here may offer, which would avoid the kind of situation cited by Williams et al (2015) where, despite their application of MCID values for clinical effect interpretation, they referenced MCID values determined for post-surgical cohorts. This argues that it is plausible that the MCID values are not truly applicable in the conservative therapy paradigm (Bahreini et al. 2019) and as a result the findings of this study would detract from clinical context and force at best application with caution in clinical practice. It is therefore worth evaluating MCID reference values in the general RA population in order to enhance conversion of evidence into clinical practice.

### 6.3.3 Recommendations for Practitioners

There are limited recommendations for practitioners stemming from this systematic review because many of the articles did not report outcomes as related to MCID (Bahreini *et al.* 2020), the latter being associated with actual clinical improvement. Consequently, most of the studies conclude their outcomes based on the statistical significance attained, which does not necessarily imply that clinical benefit was perceived by the patient or quantifiably evident.

To conclude, this systematic review must recommend that all the interventions discussed herein be applied with caution in the RA population. Even though some areas of the review provide a greater level of evidence, the compromised external validity of these “stronger” studies (particularly in respect of the RA population they utilised) limits this systematic review and therefore also the practitioner from adequately identifying the patients who would most benefit if there were more substantial and more reliable evidence on these interventions.

# BIBLIOGRAPHY


Alaaeddine, N., Okais, J., Ballane, L. and Baddoura, R. M. 2012. Use of complementary and alternative therapy among patients with rheumatoid arthritis and osteoarthritis. *J Clin Nurs*, 21 (21-22): 3198-3204.

Alcantara, J. and Leach, M. J. 2015. Chiropractic attitudes and utilization of evidence-based practice: the use of the EBASE questionnaire. *Explore*, 11 (5): 367-376.

Aletaha, D., Neogi, T., Silman, A. J., Funovits, J., Felson, D. T., Bingham III, C. O., Birnbaum, N. S., Burmester, G. R., Bykerk, V. P. and Cohen, M. D. 2010. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis & Rheumatism*, 62 (9): 2569-2581.

Allied Health Professions of South Africa. Africa, A. H. P. o. S. 2015. *Code of ethics in terms of section 54(9) of Regulations NO.R.127 of 12 February 2001 to the Allied Health Professions Act, Act 63 of 1982, as amended including guidelines for good practice and guidelines for making professional services known.* Pretoria, South Africa: Government gazette. Available: http://ahpcsa.co.za/wp-content/uploads/2015/10/39531_18-12_NationalGovernment-1.pdf (Accessed

Almuhaysin, A. A. 2019. Case Study of Treatment of a Patient after Rheumatoid Arthritis.

Almutairi, K. B., Nossent, J. C., Preen, D. B., Keen, H. I. and Inderjeeth, C. A. 2021. The prevalence of Rheumatoid Arthritis: A systematic review of population-based studies. *J Rheumatol*, 48 (5): 669-676.

Alpízar-Rodríguez, D., Pluchino, N., Canny, G., Gabay, C. and Finckh, A. 2017. The role of female hormonal factors in the development of rheumatoid arthritis. *Rheumatology*, 56 (8): 1254-1263.

Alvarez-Nemegyei, J., Bautista-Botello, A. and Dávila-Velázquez, J. 2009. Association of complementary or alternative medicine use with quality of life, functional status or cumulated damage in chronic rheumatic diseases. *Clinical rheumatology*, 28 (5): 547-551.

American Academy of Orthopaedic Surgeons. 1965. *Joint motion: method of measuring and recording*. Churchill Livingstone.

Anderson, D. L. 2001. Development of an instrument to measure pain in rheumatoid arthritis: Rheumatoid Arthritis Pain Scale (RAPS). *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 45 (4): 317-323.

Anwer, S. and Alghadir, A. 2014. Effect of isometric quadriceps exercise on muscle strength, pain, and function in patients with knee osteoarthritis: a randomized controlled study. *Journal of physical therapy science*, 26 (5): 745-748.

Arnett, F. C., Edworthy, S. M., Bloch, D. A., Mcshane, D. J., Fries, J. F., Cooper, N. S., Healey, L. A., Kaplan, S. R., Liang, M. H. and Luthra, H. S. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 31 (3): 315-324.

Aromataris, E. 2020. Furthering the science of evidence synthesis with a mix of methods. *JBI Evidence Synthesis*, 18 (10): 2106-2107.

Arthritis Foundation. 1992. *Safe and effective: Outcomes summary.* Available: http://www.arthritis.org/media/programs/Evaluation_Summary_Exercise_ HIGH.pdf (Accessed

Astrand, I. 1960. Aerobic capacity in men and women with special reference to age. *Acta physiol scand*, 49 (169): 1-89.

Astrand PO, R. K. 1986. *Textbook of work physiology. Physiological bases of exercise*. McGraw-Hill.

Atchison, J. W. 2000. Manipulation efficacy: upper body. *Journal of back and musculoskeletal rehabilitation*, 15 (1): 3-15.

ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. 2002. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*, 166: 111-117.

Avina-Zubieta, J. A., Thomas, J., Sadatsafavi, M., Lehman, A. J. and Lacaille, D. 2012. Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. *Annals of the rheumatic diseases*, 71 (9): 1524-1529.

Bahreini, M., Safaie, A., Mirfazaelian, H. and Jalili, M. 2020. How much change in pain score does really matter to patients? *The American journal of emergency medicine*, 38 (8): 1641-1646.

Baillet, A., Payraud, E., Niderprim, V.-A., Nissen, M. J., Allenet, B., Francois, P., Grange, L., Casez, P., Juvin, R. and Gaudin, P. 2009. A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. *Rheumatology*, 48 (4): 410-415.

Baillet, A., Vaillant, M., Guinot, M., Juvin, R. and Gaudin, P. 2012. Efficacy of resistance exercises in rheumatoid arthritis: meta-analysis of randomized controlled trials. *Rheumatology*, 51 (3): 519-527.

Barlow, J. H., Turner, A. P. and Wright, C. C. 2000. A randomized controlled study of the Arthritis Self-Management Programme in the UK. *Health education research*, 15 (6): 665-680.

Barnes, J. F. 1990. *Myofascial Release: The search for excellence: a comprehensive evaluatory and treatment approach*. JF Barnes.

Bartley, E. J. and Fillingim, R. B. 2013. Sex differences in pain: a brief review of clinical and experimental findings. *British journal of anaesthesia*, 111 (1): 52-58.

Basaran, A. 2009. Doğal aromaterapötik bitkiler ve uçucu yağlar. *Türkiye Klinikleri Tıp Bilimleri Dergisi*, 29 (5): 86-94.

Bearne, L. M., Coomer, A. F. and Hurley, M. V. 2007. Upper limb sensorimotor function and functional performance in patients with rheumatoid arthritis. *Disability and rehabilitation*, 29 (13): 1035-1039.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. and Erbaugh, J. 1961. An inventory for measuring depression. *Archives of general psychiatry*, 4 (6): 561-571.

Beck, M. F. 2012. *Theory and Practice of Therapeutic Massage*. Cengage Learning.

Bedi, G. S., Gupta, N., Handa, R., Pal, H. and Pandey, R. 2005. Quality of life in Indian patients with rheumatoid arthritis. *Quality of life Research*, 14 (8): 1953-1958.

Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., Pitkin, R., Rennie, D., Schulz, K. F. and Simel, D. 1996. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA*, 276 (8): 637-639.

Bell, M., Lineker, S., Badley, E. M., Wilkins, A. and Maguire, L. 1992. ARTHRITIS COMMUNITY RESEARCH & EVALUATION UNIT (ACREU).

Bell, M., Lineker, S., Wilkins, A., Maguire, L. and Badley, E. 1994. The value of pilot study data when designing randomized controlled trials of arthritis therapy in the community. *Arthritis & Rheumatism 37 (9 SUPPL)*: 365.

Bell, M. J., Lineker, S.C., Wilkins, A.L., Goldsmith, C.H. and Badley E.M,. 1998. A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis.

Belur, J., Tompson, L., Thornton, A. and Simon, M. 2018. Interrater Reliability in Systematic Review Methodology. *Sociological Methods and Research*: 1-29.

Benjamin, M. and Ralphs, J. 1996. Tendons in health and disease. *Manual therapy*, 1 (4): 186-191.

Benjamin, M. and Ralphs, J. 1997. Tendons and ligaments-an overview. *Histology and histopathology*, 12 (4): 1135-1144.

Bennell, K., Dobson, F. and Hinman, R. 2011. Measures of physical performance assessments: Self-Paced walk test (SPWT), stair climb test (SCT), Six-Minute walk test (6MWT), chair stand test (CST), timed up & go (TUG), sock test, lift and carry test (LCT), and car task. *Arthritis Care & Research*, 63 (S11): S350-S370.

Bergmann, T. F. and Peterson, D. H. 2010. *Chiropractic Technique: Principles and Procedures*. Elsevier/Mosby.

Bhalerao, M. S., Bolshete, P. M., Swar, B. D., Bangera, T. A., Kolhe, V. R., Tambe, M. J., Wade, M. P., Bhowate, S. D., Sonje, U. B., Gogtay, N. J. and Thatte, U. M. 2013. Use of and satisfaction with complementary and alternative medicine in four chronic diseases: a cross-sectional study from India. *Natl Med J India*, 26 (2): 75-78.

Bhasin, S. and Cheung, P. P. 2015. The role of power Doppler ultrasonography as disease activity marker in rheumatoid arthritis. *Disease markers*, 2015

Bhogal, S. K., Teasell, R. W., Foley, N. C. and Speechley, M. R. 2005. The PEDro scale provides a more comprehensive measure of methodological quality than the Jadad scale in stroke rehabilitation literature. *Journal of clinical epidemiology*, 58 (7): 668-673.

Bialosky, J. E., Bishop, M. D., George, S. Z. and Robinson, M. E. 2011. Placebo response to manual therapy: something out of nothing? *Journal of Manual & Manipulative Therapy*, 19 (1): 11-19.

Bland, J. H. 1988. Joint, muscle and cartilage physiology as related to exercise application and practical management in rheumatic disease. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 1 (2): 99-108.

Bland, J. M. and Altman, D. G. 2011. Comparisons against baseline within randomised groups are often used and can be highly misleading. *Trials*, 12 (1): 1-7.

Blunt, C. 2015. Hierarchies of evidence in evidence-based medicine. The London School of Economics and Political Science (LSE).

Bodkhe, R., Balakrishnan, B. and Taneja, V. 2019. The role of microbiome in rheumatoid arthritis treatment. *Therapeutic Advances in Musculoskeletal Disease*, 11: 1759720X19844632.

Bohannon, R. W. 1997. Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. *Archives of physical medicine and rehabilitation*, 78 (1): 26-32.

Bongartz, T., Sutton, A. J., Sweeting, M. J., Buchan, I., Matteson, E. L. and Montori, V. 2006. Anti-TNF Antibody Therapy in Rheumatoid Arthritis and the Risk of Serious Infections and MalignanciesSystematic Review and Meta-analysis of Rare Harmful Effects in Randomized Controlled Trials. *JAMA*, 295 (19): 2275-2285.

Borba, V. V., Zandman-Goddard, G. and Shoenfeld, Y. 2018. Prolactin and autoimmunity. *Frontiers in Immunology*, 9: 73.

Bordens, K. S. and Abbott, B. B. 2002. *Research design and methods: A process approach*. McGraw-Hill.

Borg, G. 1973. Perceived exertion: a note on “history” and methods. *Medicine and science in sports*, 5 (2): 90-93.

Borg, G. A. 1982. Psychophysical bases of perceived exertion. *Medicine & science in sports & exercise*,

Borkan, J., Neher, J. O., Anson, O. and Smoker, B. 1994. Referrals for alternative therapies. *Journal of Family Practice*, 39 (6): 545-550.

Boström, C. and Ljungquist, T. 1984. *Aktiv dynamisk träning-något för reumatiker med axelsmärta*. Hälso-och sjukvårdsnämnden, Stockholms läns landsting.

Botelho, M. B. and Andrade, B. B. 2012. Effect of cervical spine manipulative therapy on judo athletes' grip strength. *Journal of manipulative and physiological therapeutics*, 35 (1): 38-44.

Boucher, P. B., Brousseau, D. and Chahine, S. 2016. A jurisdictional review of the legislation governing informed consent by chiropractors across Canada. *The Journal of the Canadian Chiropractic Association*, 60 (1): 73.

Brahee, D. D., Pierre-Jerome, C. and Kettner, N. W. 2003. Clinical and radiological manifestations of the rheumatoid wrist. A comprehensive review. *Journal of manipulative and physiological therapeutics*, 26 (5): 323-329.

Braun, K.-P., Brookman-Amissah, S., Geissler, K., Ast, D., May, M. and Ernst, H. 2009. Ganzkörperkryotherapie bei Patienten mit entzündlich-rheumatischen Erkrankungen. *Medizinische Klinik*, 104 (3): 192-196.

Breedland, I., van Scheppingen, C., Leijsma, M., Verheij-Jansen, N. P. and van Weert, E. 2011. Effects of a group-based exercise and educational program on physical performance and disease self-management in rheumatoid arthritis: a randomized controlled study. *Physical therapy*, 91 (6): 879-893.

Breedveld, F. C. and Kalden, J. R. 2004. Appropriate and effective management of rheumatoid arthritis. *Annals of the rheumatic diseases*, 63 (6): 627.

Brentano, F., Schorr, O., Gay, R. E., Gay, S. and Kyburz, D. 2005. RNA released from necrotic synovial fluid cells activates rheumatoid arthritis synovial fibroblasts via Toll-like receptor 3. *Arthritis & Rheumatism*, 52 (9): 2656-2665.

Breuer, G. S., Orbach, H., Elkayam, O., Berkun, Y., Paran, D., Mates, M. and Nesher, G. 2006. Use of complementary and alternative medicine among patients attending rheumatology clinics in Israel. *IMAJ-RAMAT GAN-*, 8 (3): 184.

Briggs, A. M., Fary, R. E., Slater, H., Ranelli, S. and Chan, M. 2013. Physiotherapy co-management of rheumatoid arthritis: Identification of red flags, significance to clinical practice and management pathways. *Manual therapy*, 18 (6): 583-587.

Brodin, N., Eurenius, E., Jensen, I., Nisell, R., Opava, C. H. and Group, P. S. 2008. Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. *Arthritis Care & Research*, 59 (3): 325-331.

Brorsson, S., Hilliges, M., Sollerman, C. and Nilsdotter, A. 2009. A six-week hand exercise programme improves strength and hand function in patients with rheumatoid arthritis. *Journal of rehabilitation medicine*, 41 (5): 338-342.

Brosseau, L., Rahman, P., Poitras, S., Toupin-April, K., Paterson, G., Smith, C., King, J., Casimiro, L., De Angelis, G. and Loew, L. 2014. A systematic critical appraisal of non-pharmacological management of rheumatoid arthritis with appraisal of guidelines for research and evaluation II. *PLOS ONE*, 9 (5): e95369.

Brownfield, A. 1998. Aromatherapy in arthritis: a study. *Nursing Standard (through 2013)*, 13 (5): 34.

Brożek, J., Akl, E. A., Alonso-Coello, P., Lang, D., Jaeschke, R., Williams, J. W., Phillips, B., Lelgemann, M., Lethaby, A. and Bousquet, J. 2009. Grading quality of evidence and strength of recommendations in clinical practice guidelines: part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions. *Allergy*, 64 (5): 669-677.

Bruce, B. and Fries, J. F. 2005. The health assessment questionnaire (HAQ). *Clinical and experimental rheumatology*, 23 (5): S14.

Brydon, L., Wright, C., O'Donnell, K., Zachary, I., Wardle, J. and Steptoe, A. 2008. Stress-induced cytokine responses and central adiposity in young women. *International Journal of Obesity*, 32 (3): 443-450.

Brzycki, M. 1993. Strength testing—predicting a one-rep max from reps-to-fatigue. *Journal of physical education, recreation & dance*, 64 (1): 88-90.

Buchanan, C. I. and Marsh, R. L. 2002. Effects of exercise on the biomechanical, biochemical and structural properties of tendons. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 133 (4): 1101-1107.

Bucquet, D., Condon, S. and Ritchie, K. 1990. The French version of the Nottingham Health Profile. A comparison of items weights with those of the source version. *Social science & medicine*, 30 (7): 829-835.

Buljina, A. I., Taljanovic, M. S., Avdic, D. M. and Hunter, T. B. 2001. Physical and exercise therapy for treatment of the rheumatoid hand. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 45 (4): 392-397.

Busija, L., Pausenberger, E., Haines, T. P., Haymes, S., Buchbinder, R. and Osborne, R. H. 2011. Adult measures of general health and health-related quality of life: medical outcomes study short form 36-item (SF-36) and short form 12-item (SF-12) health surveys, nottingham health profile (NHP), sickness impact profile (SIP), medical outcomes study short form 6D (SF-6D), health utilities index mark 3 (HUI3), quality of well-being scale (QWB), and assessment of quality of life (AQoL). *Arthritis Care & Research*, 63 (S11): S383-S412.

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R. and Kupfer, D. J. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research*, 28 (2): 193-213.

Byfield, D. and Chapman-Smith, D. 2011. *Technique Skills in Chiropractic*. Churchill Livingstone/Elsevier.

Callaham, M. L., Wears, R. L., Weber, E. J., Barton, C. and Young, G. 1998. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA*, 280 (3): 254-257.

Callahan, L. F., Mielenz, T., Freburger, J., Shreffler, J., Hootman, J., Brady, T., Buysse, K. and Schwartz, T. 2008. A randomized controlled trial of the people with arthritis can exercise program: symptoms, function, physical activity, and psychosocial outcomes. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 59 (1): 92-101.

Campbell, A. J., Robertson, M. C., Gardner, M. M., Norton, R. N., Tilyard, M. W. and Buchner, D. M. 1997. Randomised controlled trial of a general practice programme of home based exercise to prevent falls in elderly women. *Bmj*, 315 (7115): 1065-1069.

Carbon, R., Macey, M., McCarthy, D., Pereira, F., Perry, J. and Wade, A. 1996. The effect of 30 min cycle ergometry on ankylosing spondylitis. *Rheumatology*, 35 (2): 167-177.

Carlesso, L. C., MacDermid, J. C., Gross, A. R., Walton, D. M. and Santaguida, P. L. 2014. Treatment preferences amongst physical therapists and chiropractors for the management of neck pain: results of an international survey. *Chiropractic & Manual Therapies*, 22 (1): 1-15.

Cassatella, M. A., Östberg, N. K., Tamassia, N. and Soehnlein, O. 2019. Biological roles of neutrophil-derived granule proteins and cytokines. *Trends in immunology*, 40 (7): 648-664.

Cathcart, S., Winefield, A. H., Rolan, P. and Lushington, K. 2009. Reliability of temporal summation and diffuse noxious inhibitory control. *Pain Research and Management*, 14 (6): 433-438.

Centers for Disease Control and Prevention. 2011. *Intervention programs*. Available: http://www.cdc.gov/arthritis/interventions.htm (Accessed

Chandrashekara, S. 2011. Complementary and alternative medicine in rheumatoid arthritis. *Chin J Integr Med*, 17 (10): 731-734.

Chandrupatla, D. M., Molthoff, C. F., Lammertsma, A. A., van der Laken, C. J. and Jansen, G. 2019. The folate receptor β as a macrophage-mediated imaging and therapeutic target in rheumatoid arthritis. *Drug delivery and translational research*, 9 (1): 366-378.

Chen, L. and Michalsen, A. 2017. Management of chronic pain using complementary and integrative medicine. *Bmj*, 357: j1284.

Christie, A., Jamtvedt, G., Dahm, K. T., Moe, R. H., Haavardsholm, E. A. and Hagen, K. B. 2007. Effectiveness of nonpharmacological and nonsurgical interventions for patients with rheumatoid arthritis: an overview of systematic reviews. *Physical therapy*, 87 (12): 1697-1715.

Christley, R. M. 2010. Power and error: increased risk of false positive results in underpowered studies. *The Open Epidemiology Journal*, 3 (1)

Chui, D., Lau, J. and Yau, I. 2004. An outcome evaluation study of the rheumatoid arthritis self-management programme in Hong Kong. *Psychology, health & medicine*, 9 (3): 286-292.

Chui, D. Y., Poon, P. K., Lee, E. S. and Lau, J. S. 1998. Self-management programme for rheumatoid arthritis in Hong Kong. *British Journal of Therapy and Rehabilitation*, 5 (9): 477-481.

Chung, C. L. and Mior, S. A. 2015. Use of spinal manipulation in a rheumatoid patient presenting with acute thoracic pain: a case report. *The Journal of the Canadian Chiropractic Association*, 59 (2): 143.

Cima, S. R., Barone, A., Porto, J. M. and de Abreu, D. C. C. 2013. Strengthening exercises to improve hand strength and functionality in rheumatoid arthritis with hand deformities: a randomized, controlled trial. *Rheumatology international*, 33 (3): 725-732.

Clark, H. D., Wells, G. A., Huët, C., McAlister, F. A., Salmi, L. R., Fergusson, D. and Laupacis, A. 1999. Assessing the quality of randomized trials: reliability of the Jadad scale. *Controlled clinical trials*, 20 (5): 448-452.

Coetzer, D. J. 1999. The relative effectiveness and cost effectiveness of piroxicam compared to manipulation in the treatment of acute grades 1 and 2 inversion ankle sprains.

Collins, S. L., Moore, R. A. and McQuay, H. J. 1997. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain*, 72 (1-2): 95-97.

Combe, B., Landewé, R., Lukas, C., Bolosiu, H. D., Breedveld, F., Dougados, M., Emery, P., Ferraccioli, G., Hazes, J. and Klareskog, L. 2007. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Annals of the rheumatic diseases*, 66 (1): 34-45.

Constant, C. and Murley, A. 1987. A clinical method of functional assessment of the shoulder. *Clinical orthopaedics and related research*, (214): 160-164.

Cook, C. E. and Thigpen, C. A. 2019. *Five good reasons to be disappointed with randomized trials*: Taylor & Francis.

Cooney, J. K., Law, R.-J., Matschke, V., Lemmey, A. B., Moore, J. P., Ahmad, Y., Jones, J. G., Maddison, P. and Thom, J. M. 2011. Benefits of Exercise in Rheumatoid Arthritis. *Journal of Aging Research*, 2011: 681640.

Corrao, S., Argano, C., Calvo, L. and Pistone, G. 2015. The challenge of using the rheumatoid arthritis diagnostic criteria in clinical practice. *Internal and Emergency Medicine*, 10 (3): 271-275.

Coster, W., Ludlow, L. and Mancini, M. 1999. Using IRT variable maps to enrich understanding of rehabilitation data. *Journal of outcome measurement*, 3 (2): 123-133.

Coster, W. J. 2013. Making the best match: selecting outcome measures for clinical trials and outcome studies. *American Journal of Occupational Therapy*, 67 (2): 162-170.

Cramer, G. D., Tuck Jr, N. R., Knudsen, J. T., Fonda, S. D., Schliesser, J. S., Fournier, J. T. and Patel, P. 2000. Effects of side-posture positioning and side-posture adjusting on the lumbar zygapophysial joints as evaluated by magnetic resonance imaging: a before and after study with randomization. *Journal of manipulative and physiological therapeutics*, 23 (6): 380-394.

Cross, M., Smith, E., Hoy, D., Carmona, L., Wolfe, F., Vos, T., Williams, B., Gabriel, S., Lassere, M. and Johns, N. 2014. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Annals of the rheumatic diseases*, 73 (7): 1316-1322.

Cubick, E. E., Quezada, V. Y., Schumer, A. D. and Davis, C. M. 2011. Sustained release myofascial release as treatment for a patient with complications of rheumatoid arthritis and collagenous colitis: a case report. *International journal of therapeutic massage & bodywork*, 4 (3): 1.

Cuppusamy, D. 2014. A systematic review to determine the evidence to support the use of flexion distraction chiropractic technique.

Dagenais, S. and Haldeman, S. 2011. *Evidence-based management of low back pain*. Elsevier Health Sciences.

Damjanović, V., Vasilj, I., Vlak, T. and Zelenika, D. 2009. Prevalence and risk factors of the Rheumatoid Arthritis in Herzegovina Region in 2003–2005. *Collegium antropologicum*, 33 (2): 73-77.

Dandona, P., Aljada, A. and Bandyopadhyay, A. 2004. Inflammation: the link between insulin resistance, obesity and diabetes. *Trends in immunology*, 25 (1): 4-7.

Danneskiold-Samsøe, B. and Grimby, G. 1986. Isokinetic and isometric muscle strength in patients with rheumatoid arthritis. The relationship to clinical parameters and the influence of corticosteroid. *Clinical rheumatology*, 5 (4): 459-467.

Davis, M. A., Mackenzie, T. A., Coulter, I. D., Whedon, J. M. and Weeks, W. B. 2012. The United States Chiropractic Workforce: An alternative or complement to primary care? *Chiropractic & Manual Therapies*, 20 (1): 1-8.

de Boer, M. R., Waterlander, W. E., Kuijper, L. D., Steenhuis, I. H. and Twisk, J. W. 2015. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. *International Journal of Behavioral Nutrition and Physical Activity*, 12 (1): 1-8.

de Jong, Z., Munneke, M., Kroon, H. M., van Schaardenburg, D., Dijkmans, B. A., Hazes, J. M. and Vlieland, T. P. V. 2009. Long-term follow-up of a high-intensity exercise program in patients with rheumatoid arthritis. *Clinical rheumatology*, 28 (6): 663-671.

de Jong, Z., Munneke, M., Lems, W. F., Zwinderman, A. H., Kroon, H. M., Pauwels, E. K., Jansen, A., Ronday, K. H., Dijkmans, B. A. and Breedveld, F. C. 2004. Slowing of bone loss in patients with rheumatoid arthritis by long-term high-intensity exercise: results of a randomized, controlled trial. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 50 (4): 1066-1076.

de Jong, Z., Munneke, M., Vilim, V., Zwinderman, A., Kroon, H., Ronday, H., Lems, W., Dijkmans, B., Breedveld, F. and Vliet Vlieland, T. 2008. Value of serum cartilage oligomeric matrix protein as a prognostic marker of large-joint damage in rheumatoid arthritis—data from the RAPIT study. *Rheumatology*, 47 (6): 868-871.

de Jong, Z., Munneke, M., Zwinderman, A. H., Kroon, H. M., Jansen, A., Ronday, K. H., van Schaardenburg, D., Dijkmans, B. A., Van den Ende, C. H. and Breedveld, F. C. 2003. Is a long-term high-intensity exercise program effective and safe in patients with rheumatoid arthritis?: Results of a randomized controlled trial. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 48 (9): 2415-2424.

de Jong, Z., Van der Heijde, D., McKenna, S. and Whalley, D. 1997. The reliability and construct validity of the RAQoL: a rheumatoid arthritis-specific quality of life instrument. *British journal of rheumatology*, 36 (8): 878-883.

de Jong, Z. and Vlieland, T. P. V. 2005. Safety of exercise in patients with rheumatoid arthritis. *Current opinion in rheumatology*, 17 (2): 177-182.

Deeks, J. J., Dinnes, J., D'Amico, R., Sowden, A. J., Sakarovitch, C., Song, F., Petticrew, M. and Altman, D. 2003. Evaluating non-randomised intervention studies. *Health technology assessment (Winchester, England)*, 7 (27): iii-173.

Del Rincón, I., Williams, K., Stern, M. P., Freeman, G. L. and Escalante, A. 2001. High incidence of cardiovascular events in a rheumatoid arthritis cohort not explained by traditional cardiac risk factors. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 44 (12): 2737-2745.

Della-Peruta, J. K. 2019. Improving Chronic Pain Management to Reduce the Opioid Epidemic: A Literature Review.

Dellhag, B. and Bjelle, A. 1995. A Grip Ability Test for use in rheumatology practice. *J Rheumatol*, 22 (8): 1559-1565.

Dellhag, B., Wollersjö, I. and Bjelle, A. 1992. Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 5 (2): 87-92.

Department of Health of South Africa. 2020. *The Patients' Rights Charter.* Pretoria: Available: http://www.doh.gov.za/doc/legislation/patientsright/chartere.html (Accessed 11 November 2021).

Derksen, V., Huizinga, T. and van der Woude, D. 2017. The role of autoantibodies in the pathophysiology of rheumatoid arthritis. In: Proceedings of *Seminars in immunopathology*. Springer, 437-446. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5486798/pdf/281_2017_Article_627.pdf (Accessed

DeVocht, J. W., Pickar, J. G. and Wilder, D. G. 2005. Spinal manipulation alters electromyographic activity of paraspinal muscles: a descriptive study. *Journal of manipulative and physiological therapeutics*, 28 (7): 465-471.

Dhondt, T. W., L. A Verbruggen, RA B Oostendorp, W Duquet, W. 1999. Pain threshold in patients with rheumatoid arthritis and effect of manual oscillations. *Scandinavian journal of rheumatology*, 28 (2): 88-93.

Di Giuseppe, D., Alfredsson, L., Bottai, M., Askling, J. and Wolk, A. 2012. Long term alcohol intake and risk of rheumatoid arthritis in women: a population based cohort study. *Bmj*, 345: e4230.

Di Giuseppe, D., Bottai, M., Askling, J. and Wolk, A. 2015. Physical activity and risk of rheumatoid arthritis in women: a population-based prospective study. *Arthritis research & therapy*, 17 (1): 40.

Diego, M. A., Field, T., Sanders, C. and Hernandez-Reif, M. 2004. Massage therapy of moderate and light pressure and vibrator effects on EEG and heart rate. *International Journal of Neuroscience*, 114 (1): 31-44.

Dissemination, C. 2009. Systematic reviews: CRD's guidance for undertaking reviews in healthcare. *York: University of York NHS Centre for Reviews & Dissemination*,

Donovan, J. L., Blake, D. R. and Fleming, W. G. 1989. The patient is not a blank sheet: lay beliefs and their relevance to patient education. *Rheumatology*, 28 (1): 58-61.

Dragert, K. and Zehr, E. P. 2013. High-intensity unilateral dorsiflexor resistance training results in bilateral neuromuscular plasticity after stroke. *Experimental brain research*, 225 (1): 93-104.

Droz-Georget, J. 1980. High velocity thrust and pathophysiology of segmental dysfunction. *Aspects of Manipulative Therapy. Lincoln Institute of Heath Sciences*: 81-87.

Durcan, L., Wilson, F. and Cunnane, G. 2014. The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. *J Rheumatol*, 41 (10): 1966-1973.

Durnin, J. and Rahaman, M. M. 1967. The assessment of the amount of fat in the human body from measurements of skinfold thickness. *British journal of Nutrition*, 21 (3): 681-689.

Duruöz, M., Poiraudeau, S., Fermanian, J. and Menkes, C. 1996. Functional Disability Scale That Assesses Functional. *population*, 23: 1167-1172.

Ebbeling, C. B., Ward, A., Puleo, E. M., Widrick, J. and Rippe, J. M. 1991. Development of a single-stage submaximal treadmill walking test. *Medicine and science in sports and exercise*, 23 (8): 966-973.

Edoardo, A. and Alan, P. 2014. The Systematic Review. AJN. *American Journal of Nursing*, 114 (3): 47.

Efthimiou, P. and Kukar, M. 2010. Complementary and alternative medicine use in rheumatoid arthritis: proposed mechanism of action and efficacy of commonly used modalities. *Rheumatology international*, 30 (5): 571-586.

Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C. and Antes, G. 1997. Language bias in randomised controlled trials published in English and German. *The Lancet*, 350 (9074): 326-329.

Eisenberg, D. M., Davis, R. B., Ettner, S. L., Appel, S., Wilkey, S., Van Rompay, M. and Kessler, R. C. 1998. Trends in Alternative Medicine Use in the United States, 1990-1997Results of a Follow-up National Survey. *JAMA*, 280 (18): 1569-1575.

Emery, P. 2006. Treatment of rheumatoid arthritis. *Bmj*, 332 (7534): 152.

England, B. R., Tiong, B. K., Bergman, M. J., Curtis, J. R., Kazi, S., Mikuls, T. R., O'Dell, J. R., Ranganath, V. K., Limanni, A., Suter, L. G. and Michaud, K. 2019. 2019 Update of the American College of Rheumatology Recommended Rheumatoid Arthritis Disease Activity Measures. *Arthritis Care & Research*, 71 (12): 1540-1555.

Ernst, E. 2004. Musculoskeletal conditions and complementary/alternative medicine. *Best Practice & Research Clinical Rheumatology*, 18 (4): 539-556.

Ernst, E. and Posadzki, P. 2011. Complementary and alternative medicine for rheumatoid arthritis and osteoarthritis: an overview of systematic reviews. *Curr Pain Headache Rep*, 15 (6): 431-437.

Ethgen, O., Kahler, K. H., Kong, S. X., Reginster, J.-Y. and Wolfe, F. 2002. The effect of health related quality of life on reported use of health care resources in patients with osteoarthritis and rheumatoid arthritis: a longitudinal analysis. *J Rheumatol*, 29 (6): 1147-1155.

Eti Aslan, F. 2002. Pain assessment methods. *Journal of Cumhuriyet University School of Nursing*, 6 (1): 9-16.

Evans, D. 2003. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *J Clin Nurs*, 12 (1): 77-84.

Evans, S. R. 2010. Clinical trial structures. *Journal of experimental stroke & translational medicine*, 3 (1): 8.

Eversden, L., Maggs, F., Nightingale, P. and Jobanputra, P. 2007. A pragmatic randomised controlled trial of hydrotherapy and land exercises on overall well being and quality of life in rheumatoid arthritis. *BMC musculoskeletal disorders*, 8 (1): 1-7.

Fernandez-Llanio Comella, N., Fernandez Matilla, M. and Castellano Cuesta, J. A. 2016. Have complementary therapies demonstrated effectiveness in rheumatoid arthritis? *Reumatol Clin*, 12 (3): 151-157.

Ferraz, M. B., Oliveira, L. M., Araujo, P., Atra, E. and Walter, S. 1990. EPM-ROM Scale: an evaluative instrument to be used in rheumatoid arthritis trials. *Clinical and experimental rheumatology*, 8 (5): 491-494.

Fess, F. 1992. Grip strength. *Casanova JS clinical assessment recommendations*: 41-45.

Field, T., Diego, M., Cullen, C., Hartshorn, K., Gruskin, A., Hernandez-Reif, M. and Sunshine, W. 2004. Carpal tunnel syndrome symptoms are lessened following massage therapy. *Journal of Bodywork and Movement therapies*, 8 (1): 9-14.

Field, T., Diego, M., Delgado, J., Garcia, D. and Funk, C. 2011. Hand pain is reduced by massage therapy. *Complementary therapies in clinical practice*, 17 (4): 226-229.

Field, T., Diego, M., Delgado, J., Garcia, D. and Funk, C. 2013. Rheumatoid arthritis in upper limbs benefits from moderate pressure massage therapy. *Complementary therapies in clinical practice*, 19 (2): 101-103.

Field, T., Diego, M. and Hernandez-Reif, M. 2007. Massage therapy research. *Developmental Review*, 27 (1): 75-89.

Field, T., Diego, M., Hernandez-Reif, M. and Shea, J. 2007. Hand arthritis pain is reduced by massage therapy. *Journal of Bodywork and Movement therapies*, 11 (1): 21-24.

Field, T., Hernandez-Reif, M., Seligmen, S., Krasnegor, J., Sunshine, W., Rivas-Chacon, R., Schanberg, S. and Kuhn, C. 1997. Juvenile rheumatoid arthritis: benefits from massage theraphy. *Journal of pediatric Psychology*, 22 (5): 607-617.

Fischer, A. A. 1987. Pressure algometry over normal muscles. Standard values, validity and reproducibility of pressure threshold. *Pain*, 30 (1): 115-126.

Fitzcharles, M.-A. 2002. Is it time for rheumatologists to rethink the use of manual therapies? *J Rheumatol*, 29 (6): 1117-1120.

Flatt, A. 1974. Shoulder-hand syndrome. *The Lancet*, 303 (7866): 1107-1108.

Fogg, L. and Gross, D. 2000. Threats to validity in randomized clinical trials. *Research in nursing & health*, 23 (1): 79-87.

Foley, N. C., Teasell, R. W., Bhogal, S. K. and Speechley, M. R. 2003. Stroke rehabilitation evidence-based review: methodology. *Topics in stroke rehabilitation*, 10 (1): 1-7.

Forestier, R., Andre-Vert, J., Guillez, P., Coudeyre, E., Lefevre-Colau, M. M., Combe, B. and Mayoux-Benhamou, M. A. 2009. Non-drug treatment (excluding surgery) in rheumatoid arthritis: clinical practice guidelines. *Joint Bone Spine*, 76 (6): 691-698.

Francis, P. and Francis, L. 1988. If it hurts, don't do it. *Rocklin, CA: Prima*,

Franck, H. and Gottwalt, J. 2009. Peripheral bone density in patients with rheumatoid arthritis. *Clinical rheumatology*, 28 (10): 1141-1145.

Fransen, J., Creemers, M. and Van Riel, P. 2004. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology*, 43 (10): 1252-1255.

Fransen, J. and Van Riel, P. 2006. DAS remission cut points. *Clinical and experimental rheumatology*, 24 (6 Suppl 43): S-29.

Fraser, D. F. 2008. A prospective clinical trial to determine the relative effectiveness of cross friction massage versus Graston instrument assisted soft tissue mobilisation in treating patellar tendinopathy.

Fries, J. 1983. The assessment of disability: from first to future principles. *Rheumatology*, 22 (suppl_1): 48-58.

Fries, J. F., Spitz, P., Kraines, R. G. and Holman, H. R. 1980. Measurement of patient outcome in arthritis. *Arthritis & Rheumatism*, 23 (2): 137-145.

Fuchs, H. A. and Pincus, T. 1994. Reduced joint counts in controlled clinical trials in rheumatoid arthritis. *Arthritis & Rheumatism*, 37 (4): 470-475.

Fukuda, K., Straus, S. E., Hickie, I., Sharpe, M. C., Dobbins, J. G. and Komaroff, A. 1994. The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Annals of internal medicine*, 121 (12): 953-959.

Furlan, A. D., Yazdi, F., Tsertsvadze, A., Gross, A., Van Tulder, M., Santaguida, L., Gagnier, J., Ammendolia, C., Dryden, T. and Doucette, S. 2012. A systematic review and meta-analysis of efficacy, cost-effectiveness, and safety of selected complementary and alternative medicine for neck and low-back pain. *Evidence-Based Complementary and Alternative Medicine*, 2012

Gagnier, J. J., Kienle, G., Altman, D. G., Moher, D., Sox, H. and Riley, D. 2013. The CARE guidelines: consensus-based clinical case reporting guideline development. *BMJ Case Rep*, 2013

Gambrill, E. D. 2003. *From the Editor: Evidence-Based Practice: Sea Change or the Emperor's New Clothes?* : Taylor & Francis.

Garner, S., Fenton, T., Martin, L., Creaser, C., Johns, C. and Barnabe, C. 2018. Personalized diet and exercise recommendations in early rheumatoid arthritis: A feasibility trial. *Musculoskeletal care*, 16 (1): 167-172.

Gencay-Can, A. and Can, S. S. 2012. Validation of the Turkish version of the fatigue severity scale in patients with fibromyalgia. *Rheumatology international*, 32 (1): 27-31.

Gillman, T., Schmidtke, K. A., Manning, V. and Vlaev, I. 2018. General Practitioners' recommendations of self-directed-exercises for musculoskeletal problems and perceived barriers and facilitators to doing so: a mixed methods study. *BMC health services research*, 18 (1): 1-10.

Gillon, R. 1994. Medical ethics: four principles plus attention to scope. *Bmj*, 309 (6948): 184.

Gizińska, M., Rutkowski, R., Romanowski, W., Lewandowski, J. and Straburzyńska-Lupa, A. 2015. Effects of whole-body cryotherapy in comparison with other physical modalities used with kinesitherapy in rheumatoid arthritis. *BioMed research international*, 2015

Godin, G. and Shephard, R. 1985. A simple method to assess exercise behavior in the community. *Can J Appl Sport Sci*, 10 (3): 141-146.

Goeppinger, J., Arthur, M. W., Baglioni Jr, A. J., Brunk, S. E. and Brunner, C. M. 1989. A reexamination of the effectiveness of self-care education for persons with arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 32 (6): 706-716.

Graham-Engeland, J. E., Song, S., Mathur, A., Wagstaff, D. A., Klein, L. C., Whetzel, C. and Ayoub, W. T. 2019. Emotional state can affect inflammatory responses to pain among rheumatoid arthritis patients: preliminary findings. *Psychological reports*, 122 (6): 2026-2049.

Grand-View-Research;. 2018. *U.S. Chiropractic Market Size, Share & Trends Analysis Report By Entity Type (Solo DC, DC Group), By Location Spread (West, South, Midwest, East), Vendor Landscape, And Segment Forecasts, 2018 - 2025*. Available: https://www.grandviewresearch.com/industry-analysis/us-chiropractic-market (Accessed

Grant, M.-E., Steffen, K. and Palmer, D. 2020. The usage of multidisciplinary physical therapies at the Rio de Janeiro 2016 Olympic Summer Games: an observational study. *Brazilian Journal of Physical Therapy*,

Grindstaff, T. L., Hertel, J., Beazell, J. R., Magrum, E. M. and Ingersoll, C. D. 2009. Effects of lumbopelvic joint manipulation on quadriceps activation and strength in healthy individuals. *Manual therapy*, 14 (4): 415-420.

Grob, D., Frauenfelder, H. and Mannion, A. 2007. The association between cervical spine curvature and neck pain. *European Spine Journal*, 16 (5): 669-678.

Guillemin, F., Braincon, S. and Pourel, J. 1991. Measurement of the functional capacity in rheumatoid polyarthritis: a French adaptation of the Health Assessment Questionnaire (HAQ). *Revue du rhumatisme et des maladies osteo-articulaires*, 58 (6): 459-465.

Guillemin, F., Coste, J., Pouchot, J., Ghézail, M., Bregeon, C. and Sany, J. 1997. The AIMS2-SF. A short form of the arthritis impact measurement scales 2. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 40 (7): 1267-1274.

Gummesson, C., Atroshi, I. and Ekdahl, C. 2003. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC musculoskeletal disorders*, 4 (1): 1-6.

Gupta, S. K. 2011. Intention-to-treat concept: a review. *Perspectives in clinical research*, 2 (3): 109.

Häkkinen, A. 2004. Effectiveness and safety of strength training in rheumatoid arthritis. *Current opinion in rheumatology*, 16 (2): 132-137.

Häkkinen, A., Häkkinen, K. and Hannonen, P. 1994. Effects of strength training on neuromuscular function and disease activity in patients with recent-onset inflammatory arthritis. *Scandinavian journal of rheumatology*, 23 (5): 237-242.

Häkkinen, A., Hannonen, P. and Häkkinen, K. 1995. Muscle strength in healthy people and in patients suffering from recent-onset inflammatory arthritis. *Rheumatology*, 34 (4): 355-360.

Häkkinen, A., Hannonen, P., Nyman, K., Lyyski, T. and Häkkinen, K. 2003. Effects of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: comparison with healthy subjects. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 49 (6): 789-797.

Häkkinen, A., Mälkiä, E., Häkkinen, K., Jäppinen, I., Laitinen, L. and Hannonen, P. 1997. Effects of detraining subsequent to strength training on neuromuscular function in patients with inflammatory arthritis. *British journal of rheumatology*, 36 (10): 1075-1081.

Hakkinen, A., Pakarinen, A., Hannonen, P., Kautiainen, H., Nyman, K., Kraemer, W. and Hakkinen, K. 2005. Effects of prolonged combined strength and endurance training on physical fitness, body composition and serum hormones in women with rheumatoid arthritis and in healthy controls. *Clinical and experimental rheumatology*, 23 (4): 505.

Häkkinen, A., Sokka, T., Kautiainen, H., Kotaniemi, A. and Hannonen, P. 2004. Sustained maintenance of exercise induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5 year follow up. *Annals of the rheumatic diseases*, 63 (8): 910-916.

Häkkinen, A., Sokka, T., Kotaniemi, A. and Hannonen, P. 2001. A randomized two-year study of the effects of dynamic strength training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis. *Arthritis & Rheumatism*, 44 (3): 515-522.

Häkkinen, K., Alen, M., Kallinen, M., Izquierdo, M., Jokelainen, K., Lassila, H., Mälkiä, E., Kraemer, W. J. and Newton, R. U. 1998. Muscle CSA, force production, and activation of leg extensors during isometric and dynamic actions in middle-aged and elderly men and women. *Journal of aging and physical activity*, 6 (3): 232-247.

Häkkinen, K., Komi, P. V. and Kauhanen, H. 1987. Scientific Evaluation of Specific Loading of the Knee Extensors with Variable Resistance,'Isokinetic'and Barbell Exercises1. In: *Muscular function in exercise and training*. Karger Publishers, 224-237.

Haldeman, S. 1984. Pain physiology as a neurological model for manipulation. In: *Concepts and Mechanisms of Neuromuscular Functions*. Springer, 9-18.

Haldeman, S. 2005. *Principles and Practice of Chiropractic, Third Edition*. McGraw-Hill Education.

Hamer, M. and Stamatakis, E. 2008. The accumulative effects of modifiable risk factors on inflammation and haemostasis. *Brain, behavior, and immunity*, 22 (7): 1041-1043.

Hamilton, D., Bywaters, E. and Please, N. 1959. A controlled trial of various forms of physiotherapy in arthritis. *British medical journal*, 1 (5121): 542.

Han, M., Sung, Y. K., Cho, S. K., Kim, D., Won, S., Choi, C. B., Bang, S. Y., Cha, H. S., Choe, J. Y., Chung, W. T., Hong, S. J., Jun, J. B., Jung, Y. O., Kim, S. K., Kim, T. H., Koh, E., Lee, H. S., Lee, J., Lee, J. H., Lee, S. S., Nah, S. S., Shim, S. C., Yoo, D. H., Yoo, W. H., Yoon, B. Y., Jee, S. H. and Bae, S. C. 2015. Factors Associated with the Use of Complementary and Alternative Medicine for Korean Patients with Rheumatoid Arthritis. *J Rheumatol*, 42 (11): 2075-2081.

Han, S.-H., Nam, E.-S., Uhm, D.-C., Kim, K.-S., Paik, S.-I. and Park, S.-H. 2010. Effects of aromatherapy on pain and inflammatory responses in patients with rheumatoid arthritis. *Journal of muscle and joint health*, 17 (1): 25-34.

Hansen, T., Hansen, G., Langgaard, A. and Rasmussen, J. 1993. Longterm physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. *Scandinavian journal of rheumatology*, 22 (3): 107-112.

Harden, A., Thomas, J., Cargo, M., Harris, J., Pantoja, T., Flemming, K., Booth, A., Garside, R., Hannes, K. and Noyes, J. 2018. Cochrane Qualitative and Implementation Methods Group guidance series—paper 5: methods for integrating qualitative and implementation evidence within intervention effectiveness reviews. *Journal of clinical epidemiology*, 97: 70-78.

Harding, B., Black, T., Bruulsema, A., Maxwell, B. and Stratford, P. 1988. Reliability of a reciprocal test protocol performed on the kinetic communicator: an isokinetic test of knee extensor and flexor strength. *Journal of Orthopaedic & Sports Physical Therapy*, 10 (6): 218-223.

Harkcom, T. M., Lampman, R. M., Banwell, B. F. and Castor, C. W. 1985. Therapeutic value of graded aerobic exercise training in rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 28 (1): 32-39.

Harris, K. J. 2014. The state of current knowledge regarding evidence-based conservative management of iliotibial band syndrome: a systematic review.

Harris, M. and Richards, K. C. 2010. The physiological and psychological effects of slow-stroke back massage and hand massage on relaxation in older people. *J Clin Nurs*, 19 (7-8): 917-926.

Hartling, L., Milne, A., Hamm, M. P., Vandermeer, B., Ansari, M., Tsertsvadze, A. and Dryden, D. M. 2013. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *Journal of clinical epidemiology*, 66 (9): 982-993.

Hawkes, J., Care, G., Dixon, J., Bird, H. and Wright, V. 1986. A comparison of three different physiotherapy treatments for rheumatoid arthritis of the hands. *Physiotherapy Practice*, 2 (4): 155-160.

Hays, R. D., Sherbourne, C. D. and Mazel, R. M. 1993. The rand 36-item health survey 1.0. *Health economics*, 2 (3): 217-227.

Heidari, P., Cross, W., Weller, C., Team, V., Nazarinia, M. and Crawford, K. 2019. Rheumatologists' insight into medication adherence in patients with rheumatoid arthritis: A qualitative study. *International Journal of Rheumatic Diseases*, 22 (9): 1695-1705.

Heine, P., Williams, M., Williamson, E., Bridle, C., Adams, J., O'Brien, A., Evans, D. and Lamb, S. 2012. Development and delivery of an exercise intervention for rheumatoid arthritis: strengthening and stretching for rheumatoid arthritis of the hand (SARAH) trial. *Physiotherapy*, 98 (2): 121-130.

Hemingway, P. and Brereton, N. 2009. What is a systematic review. *What is*: 1-8.

Hengstman, G., De Bleecker, J., Feist, E., Vissing, J., Denton, C., Manoussakis, M., Jensen, H. S., Van Engelen, B. and Van Den Hoogen, F. 2008. Open-label trial of anti-TNF-α in dermato-and polymyositis treated concomitantly with methotrexate. *European neurology*, 59 (3-4): 159-163.

Herbison, P., Hay-Smith, J. and Gillespie, W. J. 2006. Adjustment of meta-analyses on the basis of quality scores should be abandoned. *Journal of clinical epidemiology*, 59 (12): 1249. e1241-1249. e1211.

Heritier, S. R., Gebski, V. J. and Keech, A. C. 2003. Inclusion of patients in clinical trial analysis: the intention-to-treat principle. *Medical Journal of Australia*, 179 (8): 438-440.

Hewlett, S., Hehir, M. and Kirwan, J. R. 2007. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. *Arthritis Care & Research*, 57 (3): 429-439.

Higgins, J. P., Green, S. and Collaboration, C. 2011. Section 13.5. 2.3. Tools for assessing methodological quality or risk of bias in non-randomized studies. *Cochrane Handbook for Systematic Reviews of Interventions, Version 5. 1*,

Higgins, J. P., Green, S. and Scholten, R. 2008. Maintaining reviews: updates, amendments and feedback. *Cochrane handbook for systematic reviews of interventions*: 31.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J. and Welch, V. A. 2019. *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hill, A. B. 1965. *The environment and disease: association or causation?* : Sage Publications.

Hill, J. J. 2018. Obesity: An Emerging Threat. *Journal of National Black Nurses' Association: JNBNA*, 29 (2): 36-39.

Hochberg, M. C., Chang, R. W., Dwosh, I., Lindsey, S., Pincus, T. and Wolfe, F. 1992. The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 35 (5): 498-502.

Hoenig, H., Groff, G., Pratt, K., Goldberg, E. and Franck, W. 1993. A randomized controlled trial of home exercise on the rheumatoid hand. *J Rheumatol*, 20 (5): 785-789.

Hogeweg, J., Langereis, M., Bernards, A., Faber, J. and Helders, P. 1992. Algometry. Measuring pain threshold, method and characteristics in healthy subjects. *Scandinavian Journal of Rehabilitation Medicine*, 24 (2): 99-103.

Hogeweg, J. A., Kuis, W., Oostendorp, R. A. and Helders, P. J. 1995. General and segmental reduced pain thresholds in juvenile chronic arthritis. *Pain*, 62 (1): 11-17.

Home, D. and Carr, M. 2009. Rheumatoid arthritis: the role of early intervention and self-management. *British journal of community nursing*, 14 (10): 432-436.

Hongratanaworakit, T. 2011. Aroma-therapeutic effects of massage blended essential oils on humans. *Natural Product Communications*, 6 (8): 1934578X1100600838.

Hoskin, T. 2012. Parametric and nonparametric: Demystifying the terms. In: Proceedings of *Mayo Clinic*. 1-5.

Hosseini, A. M., Majidi, J., Baradaran, B. and Yousefi, M. 2015. Toll-like receptors in the pathogenesis of autoimmune diseases. *Advanced pharmaceutical bulletin*, 5 (Suppl 1): 605.

Hughes, G. C. 2012. Progesterone and autoimmune disease. *Autoimmunity reviews*, 11 (6-7): A502-A514.

Hurkmans, E. J., Jones, A., Li, L. C. and Vliet Vlieland, T. P. 2011. Quality appraisal of clinical practice guidelines on the use of physiotherapy in rheumatoid arthritis: a systematic review. *Rheumatology*, 50 (10): 1879-1888.

Hurley, M. and Scott, D. 1998. Improvements in quadriceps sensorimotor function and disability of patients with knee osteoarthritis following a clinically practicable exercise regime. *British journal of rheumatology*, 37 (11): 1181-1187.

Huskisson, E. 1983. Visual analogue scales. *Pain measurement and assessment*,

Huskisson, E. C. 1974. Measurement of pain. *The Lancet*, 304 (7889): 1127-1131.

Huth, E. 1999. Writing and publishing in medicine 3rd. *Edition, Williams and Wilkins*,

Ibrahim, J. G., Chu, H. and Chen, M.-H. 2012. Missing data in clinical studies: issues and methods. *Journal of clinical oncology*, 30 (26): 3297.

Ilar, A., Alfredsson, L., Wiebert, P., Klareskog, L. and Bengtsson, C. 2018. Occupation and risk of developing rheumatoid arthritis: results from a population-based case–control study. *Arthritis Care & Research*, 70 (4): 499-509.

Intriago, M., Maldonado, G., Cardenas, J. and Rios, C. 2019. Quality of life in Ecuadorian patients with established rheumatoid arthritis. *Open access rheumatology: research and reviews*, 11: 199.

Izudi, J., Semakula, D., Sennono, R., Tamwesigire, I. K. and Bajunirwe, F. 2018. Protocol for systematic review and meta-analysis of treatment success rate among adult patients with tuberculosis in sub-Saharan Africa. *BMJ open*, 8 (12): e024559.

Jackson, J. L., Kuriyama, A., Anton, A., Choi, A., Fournier, J.-P., Geier, A.-K., Jacquerioz, F., Kogan, D., Scholcoff, C. and Sun, R. 2019. The accuracy of Google Translate for abstracting data from non–English-language trials for systematic reviews. *Annals of internal medicine*, 171 (9): 677-679.

Jadhav, M. P., Jadhav, P. M., Shelke, P., Sharma, Y. and Nadkar, M. 2011. Assessment of use of complementary alternative medicine and its impact on quality of life in the patients attending rheumatology clinic, in a tertiary care centre in India. *Indian J Med Sci*, 65 (2): 50-57.

Jager, J., Putnick, D. L. and Bornstein, M. H. 2017. II. More than just convenient: The scientific merits of homogeneous convenience samples. *Monographs of the Society for Research in Child Development*, 82 (2): 13-30.

James, M. J., Cleland, L. G., Gaffney, R. D., Proudman, S. and Chatterton, B. 1994. Effect of exercise on 99mTc-DTPA clearance from knees with effusions. *J Rheumatol*, 21 (3): 501-504.

Jastrząbek, R., Straburzyńska-Lupa, A., Rutkowski, R. and Romanowski, W. 2013. Effects of different local cryotherapies on systemic levels of TNF-α, IL-6, and clinical parameters in active rheumatoid arthritis. *Rheumatology international*, 33 (8): 2053-2060.

Jegalian, A. G., Facchetti, F. and Jaffe, E. S. 2009. Plasmacytoid dendritic cells: physiologic roles and pathologic states. *Advances in anatomic pathology*, 16 (6): 392.

Jette, A. M. 1980. Functional Status Index: reliability of a chronic disease evaluation instrument. *Archives of physical medicine and rehabilitation*, 61 (9): 395-401.

Joseph, S. 2004. Client-centred therapy, post-traumatic stress disorder and post-traumatic growth: Theoretical perspectives and practical implications. *Psychology and Psychotherapy: Theory, Research and Practice*, 77 (1): 101-119.

Kajiyama, H., Akama, H., Yamanaka, H., Shoji, A., Matsuda, Y., Tanaka, E., Nakajima, A., Terai, C., Hara, M. and Tomatsu, T. 2006. One third of Japanese patients with rheumatoid arthritis use complementary and alternative medicine. *Modern rheumatology*, 16 (6): 355.

Kaltenborn, F. M. 1993. The spine. Basic evaluation and mobilization techniques. *Olaf Norlis Bokhandel*,

Kaltenborn FM. 2011. *Manual Mobilization of the Joints: The Extremities.* . 7th ed. Oslo, Norway: Orthopedic Physical Therapy Products.

Kaneda, T., Greenbaum, C. and Kline, K. 2020. *2020 World population data sheet shows older populations growing, total fertility rates declining. Population Reference Bureau, July 10, 2020*.

Kapale, P., Vardharajulu, G. and Warude, T. 2017. Effect of Free Exercise and Rheumatoid Arthritis. *Website: www. ijpot. com*, 11 (3): 62.

Kaptchuk, T. J. 1998. Powerful placebo: the dark side of the randomised controlled trial. *The Lancet*, 351 (9117): 1722-1725.

Karatepe, A. G., Günaydin, R., Türkmen, G. and Kaya, T. 2011. Effects of home-based exercise program on the functional status and the quality of life in patients with rheumatoid arthritis: 1-year follow-up study. *Rheumatology international*, 31 (2): 171-176.

Kasapis, C. and Thompson, P. D. 2005. The effects of physical activity on serum C-reactive protein and inflammatory markers: a systematic review. *Journal of the American College of Cardiology*, 45 (10): 1563-1569.

Kato, M. 2020. New insights into IFN-γ in rheumatoid arthritis: role in the era of JAK inhibitors. *Immunological Medicine*: 1-7.

Katz, N. 2002. The impact of pain management on quality of life. *Journal of pain and symptom management*, 24 (1): S38-S47.

Kauppi, M. 1996. Conservative treatment for rheumatoid cervical spine. *The Lancet*, 347 (9016): 1695.

Kauppi, M. and Anttila, P. 1996. A stiff collar for the treatment of rheumatoid atlantoaxial subluxation. *Rheumatology*, 35 (8): 771-774.

Kauppi, M., Leppänen, L., Heikkilä, S., Lahtinen, T. and Kautiainen, H. 1998. Active conservative treatment of atlantoaxial subluxation in rheumatoid arthritis. *British journal of rheumatology*, 37 (4): 417-420.

Kaur, S., White, S. and Bartold, M. 2012. Periodontal Disease as a Risk Factor for Rheumatoid Arthritis: A Systematic Review. *JBI Evidence Synthesis*, 10 (42): 1-12.

Kelly, A. M. 1998. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? *Academic Emergency Medicine*, 5 (11): 1086-1090.

Kelly, A. M. 2009. Evidence-based practice: an introduction and overview. In: Proceedings of *Seminars in roentgenology*. 131-139.

Keltikangas-Järvinen, L. and Rimon, R. 1987. Rimon's Brief Depression Scale, a rapid method for screening depression. *Psychological reports*, 60 (1): 111-119.

Kendall, J. 2003. Designing a research project: randomised controlled trials and their principles. *Emergency medicine journal: EMJ*, 20 (2): 164.

Khan, S., Otter, S. and Springett, K. 2006. The effects of reflexology on foot pain and quality of life in a patient with rheumatoid arthritis: a case report. *The Foot*, 16 (2): 112-116.

Kiecolt-Glaser, J. K., Gouin, J.-P. and Hantsoo, L. 2010. Close relationships, inflammation, and health. *Neuroscience & biobehavioral reviews*, 35 (1): 33-38.

Kim, M. J., Nam, E. S. and Paik, S. I. 2005. The effects of aromatherapy on pain, depression, and life satisfaction of arthritis patients. *Journal of Korean Academy of nursing*, 35 (1): 186-194.

Kimberlin, C. L. and Winterstein, A. G. 2008. Validity and reliability of measurement instruments used in research. *American journal of health-system pharmacy*, 65 (23): 2276-2284.

Kirkham, J. J., Dwan, K. M., Altman, D. G., Gamble, C., Dodd, S., Smyth, R. and Williamson, P. R. 2010. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *Bmj*, 340

Klinkhoff, A., Bellamy, N., Bombardier, C., Carette, S., Chalmers, A., Esdaile, J., Goldsmith, C., Tugwell, P., Smythe, H. and Buchanan, W. 1988. An experiment in reducing interobserver variability of the examination for joint tenderness. *J Rheumatol*, 15 (3): 492-494.

Kłodziński, Ł. and Wisłowska, M. 2018. Comorbidities in rheumatic arthritis. *Reumatologia*, 56 (4): 228.

Knobloch, K., Yoon, U. and Vogt, P. M. 2011. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. *Journal of Cranio-Maxillofacial Surgery*, 39 (2): 91-92.

Knott, M. and Voss, D. E. 1956. *Proprioceptive Neuromuscular Facilitation: Patterns and Techniques*. Hoeber.

Konttinen, Y., Honkanen, V., Grönblad, M., Keinonen, M., Santavirta, N. and Santavirta, S. 1992. The relation of extraarticular tenderness to inflammatory joint disease and personality in patients with rheumatoid arthritis. *J Rheumatol*, 19 (6): 851.

Konttinen, Y. T., Kemppinen, P., Segerberg, M., Hukkanen, M., Rees, R., Santavirta, S., Sorsa, T., Pertovaara, A. and Polak, J. M. 1994. Peripheral and spinal neural mechanisms in arthritis, with particular reference to treatment of inflammation and pain. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 37 (7): 965-982.

Krippendorff, K. 2004. Content Analysis: An introduction to its methodology 2nd Ed.

Küçükdeveci, A. A., Sahin, H., Ataman, S., Griffiths, B. and Tennant, A. 2004. Issues in cross-cultural validity: Example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. *Arthritis Care & Research*, 51 (1): 14-19.

Kutlay, S., Küçükdeveci, A. A., Gönül, D. and Tennant, A. 2003. Adaptation and validation of the Turkish version of the Rheumatoid Arthritis Quality of Life Scale. *Rheumatology international*, 23 (1): 21-26.

Larsen, A. 1995. How to apply Larsen score in evaluating radiographs of rheumatoid arthritis in long-term studies. *J Rheumatol*, 22 (10): 1974-1975.

Larsen, A., Dale, K. and Eek, M. 1977. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. *Acta Radiologica. Diagnosis*, 18 (4): 481-491.

Lavie, C. J., Thomas, R. J., Squires, R. W., Allison, T. G. and Milani, R. V. 2009. Exercise training and cardiac rehabilitation in primary and secondary prevention of coronary heart disease. In: Proceedings of *Mayo Clinic Proceedings*. Elsevier, 373-383. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2665984/pdf/mayoclinproc_84_4_013.pdf (Accessed

Law, R. J., Saynor, Z. L., Gabbitas, J., Jones, J., Kraus, A., Breslin, A., Maddison, P. J. and Thom, J. M. 2015. The effects of aerobic and resistance exercise on markers of large joint health in stable rheumatoid arthritis patients: a pilot study. *Musculoskeletal care*, 13 (4): 222-235.

Leach, R. A. 2004. *The Chiropractic Theories: A Textbook of Scientific Research*. Lippincott Williams & Wilkins.

Lebowitz, M. D., Knudson, R. J., Robertson, G. and Burrows, B. 1982. Significance of intraindividual changes in maximum expiratory flow volume and peak expiratory flow measurements. *Chest*, 81 (5): 566-570.

Lee, M.-S., Lee, M. S., Yang, C.-Y., Lee, S.-I., Joo, M.-C., Shin, B.-C., Yoo, W.-H. and Shin, Y.-I. 2008. Use of complementary and alternative medicine by rheumatoid arthritis patients in Korea. *Clinical rheumatology*, 27 (1): 29-33.

Lee, M., Gandevia, S. C. and Carroll, T. J. 2009. Unilateral strength training increases voluntary activation of the opposite untrained limb. *Clinical Neurophysiology*, 120 (4): 802-808.

Lee, P., Baxter, A., Dick, W. C. and Webb, J. 1974. An assessment of grip strength measurement in rheumatoid arthritis. *Scandinavian journal of rheumatology*, 3 (1): 17-23.

Leelayuwat, N. 2017. Exercise therapy for physical therapist. In: *Clinical Physical Therapy*. IntechOpen.

Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M. I., Noel-Storr, A., Rader, T., Shokraneh, F. and Thomas, J. 2019. Searching for and selecting studies. *Cochrane handbook for systematic reviews of interventions*: 67-107.

Lehman, G. 2012. Kinesiological research: the use of surface electromyography for assessing the effects of spinal manipulation. *Journal of Electromyography and Kinesiology*, 22 (5): 692-696.

Lemmey, A. B., Marcora, S. M., Chester, K., Wilson, S., Casanova, F. and Maddison, P. J. 2009. Effects of high-intensity resistance training in patients with rheumatoid arthritis: A randomized controlled trial. *Arthritis Care & Research*, 61 (12): 1726-1734.

Levitsky, A., Kisten, Y., Lind, S., Nordström, P., Hultholm, H., Lyander, J., Hammelin, V., Gentline, C., Giannakou, I. and Faustini, F. 2019. Joint mobilization of the hands of patients with rheumatoid arthritis: results from an assessor-blinded, randomized crossover trial. *Journal of manipulative and physiological therapeutics*, 42 (1): 34-46.

Lewis, K. L. and Battaglia, P. J. 2019. Differentiating Bilateral Symptomatic Hand Osteoarthritis From Rheumatoid Arthritis Using Sonography When Clinical and Radiographic Features Are Nonspecific: A Case Report. *Journal of Chiropractic Medicine*, 18 (1): 56-60.

Liang, M. H. 2000. Longitudinal construct validity: establishment of clinical meaning in patient evaluative instruments. *Medical care*, 38 (9): II-84-II-90.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J. and Moher, D. 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of clinical epidemiology*, 62 (10): e1-e34.

Liddle, J., Williamson, M. and Irwig, I. 1996a. Method for evaluating research and

guideline evidence (MERGE).

Liddle, J., Williamson, M. and Irwig, L. 1996b. *Method for evaluating research & guideline evidence*. NSW Department of Health.

Lindgren, L., Westling, G., Brulin, C., Lehtipalo, S., Andersson, M. and Nyberg, L. 2012. Pleasant human touch is represented in pregenual anterior cingulate cortex. *Neuroimage*, 59 (4): 3427-3432.

Lindroth, Y., Bauman, A., Barnes, C., McCredie, M. and Brooks, P. 1989. A controlled evaluation of arthritis education. *Rheumatology*, 28 (1): 7-12.

Lineker, S., Badley, E., Hughes, E. and Bell, M. 1997. Development of an instrument to measure knowledge in individuals with rheumatoid arthritis: the ACREU rheumatoid arthritis knowledge questionnaire. *J Rheumatol*, 24 (4): 647-653.

Lineker, S. C., Bell, M. J., Wilkins, A. L. and Badley, E. M. 2001. Improvements following short term home based physical therapy are maintained at one year in people with moderate to severe rheumatoid arthritis. *J Rheumatol*, 28 (1): 165-168.

Liu, X., Tedeschi, S. K., Lu, B., Zaccardelli, A., Speyer, C. B., Costenbader, K. H., Karlson, E. W. and Sparks, J. A. 2019. Long-term physical activity and subsequent risk for rheumatoid arthritis among women: a prospective cohort study. *Arthritis & Rheumatology*, 71 (9): 1460-1471.

Lombard, M., Snyder-Duch, J. and Bracken, C. C. 2002. Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human communication research*, 28 (4): 587-604.

Longmore, M., Wilkinson, I., Baldwin, A. and Wallin, E. 2014. *Oxford Handbook of Clinical Medicine-Mini Edition*. OUP Oxford.

Lonta, M. K. 1960. Facilitation Technics in the Treatment of Early Rheumatoid Arthritis. *Physical therapy*, 40 (2): 119-120.

López-Castillo, C. A., Calderón-Rojas, R., Amaya-Amaya, J., DeSanVicente-Célis, Z., Mantilla, R. D. and Rojas-Villarraga, A. 2014. Impact of educational level on rheumatoid arthritis: a systematic review. *Revista Colombiana de Reumatología*, 21 (4): 201-212.

Lorenzetti, R., Janowska, I., Smulski, C. R., Frede, N., Henneberger, N., Walter, L., Schleyer, M.-T., Hüppe, J. M., Staniek, J., Salzer, U., Venhoff, A., Troilo, A., Voll, R. E., Venhoff, N., Thiel, J. and Rizzi, M. 2019. Abatacept modulates CD80 and CD86 expression and memory formation in human B-cells. *Journal of Autoimmunity*, 101: 145-152.

Lorig, K., Chastain, R. L., Ung, E., Shoor, S. and Holman, H. R. 1989. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 32 (1): 37-44.

Lorig K, F. J. 1988. *The arthritis helpbook. Reading*. Addison-Wesley.

Lorig, K. and Holman, H. 1993. Arthritis self-management studies: a twelve-year review. *Health education quarterly*, 20 (1): 17-28.

Lorig, K., Lubeck, D., Kraines, R. G., Seleznick, M. and Holman, H. R. 1985. Outcomes of self-help education for patients with arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 28 (6): 680-685.

Lorig, K., Stewart, A., Ritter, P., Gonzalez, V., Lynch, J. and Laurent, D. 1996. *Outcome measures for health education and other health care interventions*. Sage.

Lowe, W. 2006. *Orthopedic Assessment in Massage Therapy*. Daviau Scott.

Lowman, E. W. 1958. Rehabilitation of the rheumatoid cripple: a five year study. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 1 (1): 38-43.

Lugli, E. B., Correia, R. E. S. M., Fischer, R., Lundberg, K., Bracke, K. R., Montgomery, A. B., Kessler, B. M., Brusselle, G. G. and Venables, P. J. 2015. Expression of citrulline and homocitrulline residues in the lungs of non-smokers and smokers: implications for autoimmunity in rheumatoid arthritis. *Arthritis research & therapy*, 17 (1): 9.

Lyngberg, K., Danneskiold-Samsøe, B. and Halskov, O. 1988. The effect of physical training on patients with rheumatoid arthritis: changes in disease activity, muscle strength and aerobic capacity. A clinically controlled minimized cross-over study. *Clinical and experimental rheumatology*, 6 (3): 253-260.

Lyngberg, K. K., Harreby, M., Bentzen, H., Frost, B. and Danneskiold-Samsøe, B. 1994. Elderly rheumatoid arthritis patients on steroid treatment tolerate physical training without an increase in disease activity. *Archives of physical medicine and rehabilitation*, 75 (11): 1189-1195.

Macedo, L. G., Elkins, M. R., Maher, C. G., Moseley, A. M., Herbert, R. D. and Sherrington, C. 2010. There was evidence of convergent and construct validity of Physiotherapy Evidence Database quality scale for physiotherapy trials. *Journal of clinical epidemiology*, 63 (8): 920-925.

Macfarlane, G. J., Paudyal, P., Doherty, M., Ernst, E., Lewith, G., MacPherson, H., Sim, J. and Jones, G. T. 2012. A systematic review of evidence for the effectiveness of practitioner-based complementary and alternative therapies in the management of rheumatic diseases: rheumatoid arthritis. *Rheumatology (Oxford)*, 51 (9): 1707-1713.

MacGregor, A. J., Griffiths, G. O., Baker, J. and Spector, T. D. 1997. Determinants of pressure pain threshold in adult twins: evidence that shared environmental influences predominate. *Pain*, 73 (2): 253-257.

Maher, C. G., Sherrington, C., Herbert, R. D., Moseley, A. M. and Elkins, M. 2003. Reliability of the PEDro scale for rating quality of randomized controlled trials. *Physical therapy*, 83 (8): 713-721.

Mainland, D. 1965. A seven-day variability study of 499 patients with peripheral rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 8 (2): 302-334.

Makaram, S. 1995. Interprofessional cooperation. *Medical education*, 29: 65-69.

Mallett, R., Hagen-Zanker, J., Slater, R. and Duvendack, M. 2012. The benefits and challenges of using systematic reviews in international development research. *Journal of development effectiveness*, 4 (3): 445-455.

Malm, K., Bergman, S., Andersson, M. L., Bremander, A. and Larsson, I. 2017. Quality of life in patients with established rheumatoid arthritis: a phenomenographic study. *SAGE open medicine*, 5: 2050312117713647.

Manek, N. J., Crowson, C. S., Ottenberg, A. L., Curlin, F. A., Kaptchuk, T. J. and Tilburt, J. C. 2010. What rheumatologists in the United States think of complementary and alternative medicine: results of a national survey. *BMC Complement Altern Med*, 10: 5.

Mannerkorpi, K. and Bjelle, A. 1994. Evaluation of a home training programme to improve shoulder function in rheumatoid arthritis patients. *Physiotherapy Theory and Practice*, 10 (2): 69-76.

Manning, V. 2013. Exercise-based upper limb rehabilitation in rheumatoid arthritis. King's College London (University of London).

Mansournia, M. A., Higgins, J. P., Sterne, J. A. and Hernán, M. A. 2017. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology (Cambridge, Mass.)*, 28 (1): 54.

Maradit-Kremers, H., Crowson, C., Nicola, P., Ballman, K., Roger, V., Jacobsen, S. and Gabriel, S. 2004. Increased unrecognized coronary heart disease in rheumatoid arthritis: a population-based cohort study. In: Proceedings of *ARTHRITIS AND RHEUMATISM*. WILEY-LISS DIV JOHN WILEY & SONS INC, 111 RIVER ST, HOBOKEN, NJ 07030 USA, S688-S688.

Maravic, M., Bozonnat, M., Sevezan, A., Gasqueres, D., Pastor, J., Pere, M., Neil, V., Roch-Bras, F., Daures, J. and Sany, J. 2000. Preliminary evaluation of medical outcomes (including quality of life) and costs in incident RA cases receiving hospital-based multidisciplinary management. *Joint Bone Spine*, 67 (5): 425-433.

Marcora, S. M., Lemmey, A. B. and Maddison, P. J. 2005. Can progressive resistance training reverse cachexia in patients with rheumatoid arthritis? Results of a pilot study. *J Rheumatol*, 32 (6): 1031-1039.

Martins, P. and Fonseca, J. E. 2019. How to investigate: Pre-clinical rheumatoid arthritis. *Best Practice & Research Clinical Rheumatology*, 33 (4): 101438.

Maska, L., Anderson, J. and Michaud, K. 2011. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). *Arthritis Care & Research*, 63 (S11): S4-S13.

Mathiowetz, V. 1991. Reliability and validity of grip and pinch strength measurements. *Critical Reviews in Physical and Rehabilitation Medicine*, 2: 201-212.

Mathiowetz, V., Weber, K., Kashman, N. and Volland, G. 1985. Adult norms for the nine hole peg test of finger dexterity. *The Occupational Therapy Journal of Research*, 5 (1): 24-38.

Matschke, V., Murphy, P., Lemmey, A. B., Maddison, P. and Thom, J. M. 2010a. Skeletal muscle properties in rheumatoid arthritis patients. *Med Sci Sports Exerc*, 42 (12): 2149-2155.

Matschke, V., Murphy, P., Lemmey, A. B., Maddison, P. J. and Thom, J. M. 2010b. Muscle quality, architecture, and activation in cachectic patients with rheumatoid arthritis. *J Rheumatol*, 37 (2): 282-284.

Matsushita, M., Yamaji, K. E. N. and Tamura, N. 2020. Extra-articular Manifestations of Rheumatoid Arthritis. *Juntendo Medical Journal*, 66 (1): 21-26.

Mayo-Wilson, E., Li, T., Fusco, N., Bertizzolo, L., Canner, J. K., Cowley, T., Doshi, P., Ehmsen, J., Gresham, G. and Guo, N. 2017. Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. *Journal of clinical epidemiology*, 91: 95-110.

McDonald, R. and Strang, J. 2016. Are take-home naloxone programmes effective? Systematic review utilizing application of the Bradford Hill criteria. *Addiction*, 111 (7): 1177-1187.

McDowell, I. and Newell, C. 1996. A guide to rating scales and questionnaires. *A guide to rating scales and questionnaires*,

McGowan, P. 1997. Evaluation of the national implementation of the Arthritis Self Management Course. Preliminary analysis. *The Arthritis Society*,

McInnes, I. B. and Schett, G. 2011. The pathogenesis of rheumatoid arthritis. *New England Journal of Medicine*, 365 (23): 2205-2219.

McManus, K. M., Visker, J. D. and Cox, C. C. 2015. Effect of an arthritis foundation exercise program on sleep quality/sleep disturbance in seniors with rheumatoid arthritis: A pilot study. *Activities, Adaptation & Aging*, 39 (1): 56-63.

McMeeken, J., Stillman, B., Story, I., Kent, P. and Smith, J. 1999. The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis. *Physiotherapy research international*, 4 (1): 55-67.

McNair, D., Lorr, M. and Droppleman, L. 1971a. POMS profile—Profile of mood states. *San Diego, CA: Educational and Industrial Testing Services*,

McNair, D. M., Lorr, M. and Droppleman, L. F. 1971b. Manual profile of mood states.

McQuoid-Mason, D. J. and Dada, M. A. 2011. *A-Z of Medical Law*. Juta.

Meehan RF, G. P., Mason JH,. 1980. Arthritis impact measurement scales. *Arthritis Rheum* 23: 146-152.

Meenan, R. F., Mason, J. H., Anderson, J. J., Guccione, A. A. and Kazis, L. E. 1992. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 35 (1): 1-10.

Meeus, M., Hermans, L., Ickmans, K., Struyf, F., Van Cauwenbergh, D., Bronckaerts, L., De Clerck, L. S., Moorken, G., Hans, G. and Grosemans, S. 2015. Endogenous pain modulation in response to exercise in patients with rheumatoid arthritis, patients with chronic fatigue syndrome and comorbid fibromyalgia, and healthy controls: a double-blind randomized controlled trial. *Pain Practice*, 15 (2): 98-106.

Meeus, M., Ickmans, K., Struyf, F., Hermans, L., Van Noesel, K., Oderkerk, J., De Clerck, L., Moorkens, G., Hans, G. and Grosemans, S. 2013. Does acetaminophen activate endogenous pain inhibition in chronic fatigue syndrome/fibromyalgia and rheumatoid arthritis?: a double-blind randomized controlled cross-over trial. *Pain Physician*, 16 (2): E61-E70.

Meeus, M., Vervisch, S., De Clerck, L. S., Moorkens, G., Hans, G. and Nijs, J. 2012. Central sensitization in patients with rheumatoid arthritis: a systematic literature review. *Semin Arthritis Rheum*, 41 (4): 556-567.

Melia, M., Geissler, B., König, J., Ottersbach, H. J., Umbreit, M., Letzel, S. and Muttray, A. 2019. Pressure pain thresholds: subject factors and the meaning of peak pressures. *European Journal of Pain*, 23 (1): 167-182.

Melzack, R. 1983. *Pain measurement and assessment*. Raven Press.

Melzack, R. and Wall, P. D. 1965. Pain mechanisms: a new theory. *Science*, 150 (3699): 971-979.

Metin, Z. G. and Ozdemir, L. 2016. The effects of aromatherapy massage and reflexology on pain and fatigue in patients with rheumatoid arthritis: a randomized controlled trial. *Pain Management Nursing*, 17 (2): 140-149.

Metsios, G. S., Stavropoulos-Kalinoglou, A., Panoulas, V. F., Wilson, M., Nevill, A. M., Koutedakis, Y. and Kitas, G. D. 2009. Association of physical inactivity with increased cardiovascular risk in patients with rheumatoid arthritis. *European Journal of Cardiovascular Prevention & Rehabilitation*, 16 (2): 188-194.

Michalsen, A. 2013. The role of complementary and alternative medicine (CAM) in rheumatology--it's time for integrative medicine. *J Rheumatol*, 40 (5): 547-549.

Mills, C. 2012. M1 and M2 Macrophages: Oracles of Health and Disease. 32 (6): 463-488.

Minor, M. A. 1991. Physical activity and management of arthritis. *Annals of Behavioral Medicine*, 13 (3): 117-124.

Moayedi, M. and Davis, K. D. 2013. Theories of pain: from specificity to gate control. *Journal of neurophysiology*, 109 (1): 5-12.

Mochan, E. and Ebell, M. H. 2008. Predicting rheumatoid arthritis risk in adults with undifferentiated arthritis. *Am Fam Physician*, 77 (10): 1451-1453.

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P., Elbourne, D., Egger, M. and Altman, D. G. 2012. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *International journal of surgery*, 10 (1): 28-55.

Moher, D., Liberati, A., Tetzlaff, J. and Altman, D. 2009a. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*, 6 (7)

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. and Group, P. 2009b. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, 6 (7): e1000097.

Montoye, H. J., Washburn, R., Servais, S., Ertl, A., Webster, J. G. and Nagle, F. J. 1983. Estimation of energy expenditure by a portable accelerometer. *Medicine and science in sports and exercise*, 15 (5): 403-407.

Mooney, V. and Robertson, J. 1976. The facet syndrome. *Clinical Orthopaedics and Related Research (1976-2007)*, 115: 149-157.

Moquin, B., Blackman, M. R., Mitty, E. and Flores, S. 2009. Complementary and alternative medicine (CAM). *Geriatric Nursing*, 30 (3): 196-203.

Morris, V. H., Cruwys, S. C. and Kidd, B. L. 1997. Characterisation of capsaicin-induced mechanical hyperalgesia as a marker for altered nociceptive processing in patients with rheumatoid arthritis. *Pain*, 71 (2): 179-186.

Moustgaard, H., Bello, S., Miller, F. G. and Hróbjartsson, A. 2014. Subjective and objective outcomes in randomized clinical trials: definitions differed in methods publications and were often absent from trial reports. *Journal of clinical epidemiology*, 67 (12): 1327-1334.

Mouton, J. 2006. *How to succeed in your in your Master's and Doctoral Studies*: Van Schaik Publishers. Pretoria.

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A. and Aromataris, E. 2018a. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, 18 (1): 143.

Munn, Z., Stern, C., Aromataris, E., Lockwood, C. and Jordan, Z. 2018b. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Medical Research Methodology*, 18 (1): 5.

Murad, M. H., Sultan, S., Haffar, S. and Bazerbachi, F. 2018. Methodological quality and synthesis of case series and case reports. *BMJ evidence-based medicine*, 23 (2): 60-63.

Muskardin, T. L. W. and Niewold, T. B. 2018. Type I interferon in rheumatic diseases. *Nature Reviews Rheumatology*, 14 (4): 214-228.

Muthny, F. 1989. Freiburg questionnaire of coping with illness–manual. *Beltz Test, Weinheim, Germany*,

Muthny, F. 1996. References for evaluation scales in quality assurance in rehabilitation--6. Assessment of coping processes with the Freiburg Questionnaire of Illness Coping. *Die Rehabilitation*, 35 (2): IX-XVI.

Mutlu, E. K. and Ozdincler, A. R. 2015. Reliability and responsiveness of algometry for measuring pressure pain threshold in patients with knee osteoarthritis. *Journal of physical therapy science*, 27 (6): 1961-1965.

Myasoedova, E., Davis, J., Matteson, E. L. and Crowson, C. S. 2020. Is the epidemiology of rheumatoid arthritis changing? Results from a population-based incidence study, 1985–2014. *Annals of the rheumatic diseases*, 79 (4): 440-444.

Nair, B. 2019. Clinical trial designs. *Indian dermatology online journal*, 10 (2): 193.

Nakagawa, S., Nakaishi, M., Hashimoto, M., Ito, H., Yamamoto, W., Nakashima, R., Tanaka, M., Fujii, T., Omura, T., Imai, S., Nakagawa, T., Yonezawa, A., Imai, H., Mimori, T. and Matsubara, K. 2018. Effect of medication adherence on disease activity among Japanese patients with rheumatoid arthritis. *PLOS ONE*, 13 (11): e0206943.

Nandgude, T. D., Hasabe, P. S. and Kolsure, A. K. 2018. Clinical Features and Treatment of Rheumatoid Arthritis: A Review. *Research Journal of Pharmacy and Technology*, 11 (12): 5701-5706.

Naqvi, A. A., Hassali, M. A. and Aftab, M. T. 2019. Epidemiology of rheumatoid arthritis, clinical aspects and socio-economic determinants in Pakistani patients: A systematic review and meta-analysis. *JPMA. The Journal of the Pakistan Medical Association*, 69 (3): 389-398.

Neill, J., Belan, I. and Ried, K. 2006. Effectiveness of non-pharmacological interventions for fatigue in adults with multiple sclerosis, rheumatoid arthritis, or systemic lupus erythematosus: a systematic review. *Journal of advanced nursing*, 56 (6): 617-635.

Neuberger, G. B., Press, A. N., Lindsley, H. B., Hinton, R., Cagle, P. E., Carlson, K., Scott, S., Dahl, J. and Kramer, B. 1997. Effects of exercise on fatigue, aerobic fitness, and disease activity measures in persons with rheumatoid arthritis. *Research in nursing & health*, 20 (3): 195-204.

Newsome, G. 2002. Guidelines for the management of rheumatoid arthritis: 2002 update. *Journal of the American Academy of Nurse Practitioners*, 14 (10): 432-437.

Noack, M. and Miossec, P. 2014. Th17 and regulatory T cell balance in autoimmune and inflammatory diseases. *Autoimmunity reviews*, 13 (6): 668-677.

Noreau, L., Martineau, H., Roy, L. and Belzile, M. 1995. Effects of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. *American journal of physical medicine & rehabilitation*, 74 (1): 19-27.

O'Brien, A., Jones, P., Mullis, R., Mulherin, D. and Dziedzic, K. 2006. Conservative hand therapy treatments in rheumatoid arthritis—a randomized controlled trial. *Rheumatology*, 45 (5): 577-583.

O'Connor, S. R., Tully, M. A., Ryan, B., Bradley, J. M., Baxter, G. D. and McDonough, S. M. 2015. Failure of a numerical quality assessment scale to identify potential risk of bias in a systematic review: a comparison study. *BMC research notes*, 8 (1): 1-7.

O'Neil, L. J. and Kaplan, M. J. 2019. Neutrophils in rheumatoid arthritis: breaking immune tolerance and fueling disease. *Trends in molecular medicine*, 25 (3): 215-227.

O'Reilly, S. C., Muir, K. R. and Doherty, M. 1999. Effectiveness of home exercise on pain and disability from osteoarthritis of the knee: a randomised controlled trial. *Annals of the rheumatic diseases*, 58 (1): 15-19.

Oh, H. and Seo, W. 2003. Decreasing pain and depression in a health promotion program for people with rheumatoid arthritis. *Journal of Nursing Scholarship*, 35 (2): 127-132.

Ohrbach, R., Crow, H. and Kamer, A. 1998. Examiner expectancy effects in the measurement of pressure pain thresholds. *Pain*, 74 (2-3): 163-170.

Okada, Y., Eyre, S., Suzuki, A., Kochi, Y. and Yamamoto, K. 2019. Genetics of rheumatoid arthritis: 2018 status. *Annals of the rheumatic diseases*, 78 (4): 446-453.

Olivo, S. A., Macedo, L. G., Gadotti, I. C., Fuentes, J., Stanton, T. and Magee, D. J. 2008. Scales to assess the quality of randomized controlled trials: a systematic review. *Physical therapy*, 88 (2): 156-175.

Olsson, Å. R., Skogh, T. and Wingren, G. 2001. Comorbidity and lifestyle, reproductive factors, and environmental exposures associated with rheumatoid arthritis. *Annals of the rheumatic diseases*, 60 (10): 934-939.

Ooijendijk, W., Hildebrandt, V. and Stiggelbout, M. 2000. Physical activity in the Netherlands 2000: first results from the monitoring study Physical Activity and Health. *Trendrapport Bewegen en Gezondheid*, 2001: 7-23.

Orange, D. E., Blachere, N. E., DiCarlo, E. F., Mirza, S., Pannellini, T., Jiang, C. S., Frank, M. O., Parveen, S., Figgie, M. P., Gravallese, E. M., Bykerk, V. P., Orbai, A.-M., Mackie, S. L. and Goodman, S. M. 2020. Rheumatoid Arthritis Morning Stiffness Is Associated With Synovial Fibrin and Neutrophils. *Arthritis & Rheumatology*, 72 (4): 557-564.

Organization, W. H. 2019. WHO global report on trends in prevalence of tobacco use 2000-2025.

Osborn, K. 2005. Chasing the pain away.(electronic version). *Massage & Bodywork*, 20 (3): 138-143.

Otter, S., Church, A., Murray, A., Lucas, J., Creasey, N., Woodhouse, J., Grant, R. and Cooper, H. 2010. The effects of reflexology in reducing the symptoms of fatigue in people with rheumatoid arthritis: a preliminary study. *The journal of alternative and complementary medicine*, 16 (12): 1251-1252.

Ovayolu, Ö. and Ovayolu, N. 2013. Onkolojide semptom yönetiminde kullanilan kanit temelli tamamlayici yöntemler ve etkileri. *ERÜ Sağlık Bilimleri Fakültesi Dergisi*, 1 (1): 83-98.

Palmefors, H., DuttaRoy, S., Rundqvist, B. and Börjesson, M. 2014. The effect of physical activity or exercise on key biomarkers in atherosclerosis–a systematic review. *Atherosclerosis*, 235 (1): 150-161.

Park, G., Kim, C. W., Park, S. B., Kim, M. J. and Jang, S. H. 2011. Reliability and usefulness of the pressure pain threshold measurement in patients with myofascial pain. *Annals of rehabilitation medicine*, 35 (3): 412.

Pascual-Salcedo, D., Plasencia, C., Ramiro, S., Nuño, L., Bonilla, G., Nagore, D., Ruiz del Agua, A., Martínez, A., Aarden, L. and Martín-Mola, E. 2011. Influence of immunogenicity on the efficacy of long-term treatment with infliximab in rheumatoid arthritis. *Rheumatology*, 50 (8): 1445-1452.

Pater, J., Osoba, D., Zee, B., Lofters, W., Gore, M., Dempsey, E., Palmer, M. and Chin, C. 1998. Effects of altering the time of administration and the time frame of quality of life assessments in clinical trials: an example using the EORTC QLQ-C30 in a large anti-emetic trial. *Quality of life Research*, 7 (3): 273-278.

Pearson, A. 2004. Balancing the evidence: incorporating the synthesis of qualitative data into systematic reviews. *JBI reports*, 2 (2): 45-64.

Pearson, A., Jordan, Z. and Munn, Z. 2012. Translational science and evidence-based healthcare: a clarification and reconceptualization of how knowledge is generated and used in healthcare. *Nursing research and practice*, 2012

Pedersen, M., Jacobsen, S., Klarlund, M. and Frisch, M. 2006a. Socioeconomic status and risk of rheumatoid arthritis: a Danish case-control study. *J Rheumatol*, 33 (6): 1069-1074.

Pedersen, M., Jacobsen, S., Klarlund, M., Pedersen, B. V., Wiik, A., Wohlfahrt, J. and Frisch, M. 2006b. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis research & therapy*, 8 (4): R133.

PEDro. 1999. *PEDro Scale*. Available: https://pedro.org.au/english/resources/pedro-scale/ (Accessed

*PEDro Scale*. 1999. Available: https://www.pedro.org.au/english/downloads/pedro-scale/ (Accessed 20/07/2017).

Peled, M. and Fisher, E. A. 2014. Dynamic Aspects of Macrophage Polarization during Atherosclerosis Progression and Regression. *Frontiers in Immunology*, 5 (579)

Perera, S., Mody, S. H., Woodman, R. C. and Studenski, S. A. 2006. Meaningful change and responsiveness in common physical performance measures in older adults. *Journal of the American Geriatrics Society*, 54 (5): 743-749.

Perlman, S. G., Connell, K. J., Clark, A., Robinson, M. S., Conlon, P., Gecht, M., Caldron, P. and Sinacore, J. M. 1990. Dance-based aerobic exercise for rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 3 (1): 29-35.

Peyron, J. G. 1986. Review of the main epidemiologic-etiologic evidence that implies mechanical forces as factors in osteoarthritis. *Engineering in medicine*, 15 (2): 77-79.

Phang, J. K., Kwan, Y. H., Goh, H., Tan, V. I. C., Thumboo, J., Ostbye, T. and Fong, W. 2018. Complementary and alternative medicine for rheumatic diseases: A systematic review of randomized controlled trials. *Complement Ther Med*, 37: 143-157.

Pickar, J. G. 1999. An in vivo preparation for investigating neural responses to controlled loading of a lumbar vertebra in the anesthetized cat. *Journal of neuroscience methods*, 89 (2): 87-96.

Pickar, J. G. 2002. Neurophysiological effects of spinal manipulation. *The spine journal*, 2 (5): 357-371.

Pierson, D. J. 2009. How to read a case report (or teaching case of the month). *Respiratory Care*, 54 (10): 1372-1378.

Pincus, T., Callahan, L. and Vaughn, W. 1987. Questionnaire, walking time and button test measures of functional capacity as predictive markers for mortality in rheumatoid arthritis. *J Rheumatol*, 14 (2): 240-251.

Piva, S. R., Khoja, S. S., Toledo, F. G., Chester-Wasko, M., Fitzgerald, G. K., Goodpaster, B. H., Smith, C. N. and Delitto, A. 2019. Neuromuscular electrical stimulation compared to volitional exercise for improving muscle function in rheumatoid arthritis: A randomized pilot study. *Arthritis Care & Research*, 71 (3): 352-361.

Platt, R. W., Schisterman, E. F. and Cole, S. R. 2009. Time-modified confounding. *American journal of epidemiology*, 170 (6): 687-694.

Pocock, S. J., Assmann, S. E., Enos, L. E. and Kasten, L. E. 2002. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practiceand problems. *Statistics in medicine*, 21 (19): 2917-2930.

Pocock, S. J. and Simon, R. 1975. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*: 103-115.

Podsiadlo, D. and Richardson, S. 1991. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *Journal of the American geriatrics Society*, 39 (2): 142-148.

Pollock, M. L., Gaesser, G. A., Butcher, J., Després, J. P., Dishman, R. K., Franklin, B. A. and Garber, C. E. 1998. ACSM position stand: the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Journals AZ> Medicine & Science*, 30: 6.

Prevoo, M., Van'T Hof, M. A., Kuper, H., Van Leeuwen, M., Van De Putte, L. and Van Riel, P. 1995. Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 38 (1): 44-48.

Price, D. D., McGrath, P. A., Rafii, A. and Buckingham, B. 1983. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain*, 17 (1): 45-56.

Primdahl, J., Wagner, L. and Hørslev-Petersen, K. 2011. Self-efficacy as an outcome measure and its association with physical disease-related variables in persons with rheumatoid arthritis: A literature review. *Musculoskeletal care*, 9 (3): 125-140.

Pugatch, D., Haskell, D. and McNair, D. 1969. Predictors and patterns of change associated with the course of time-limited psychotherapy. *Archives of General Psychol*, 21: 546-552.

Quigley, J. M., Thompson, J. C., Halfpenny, N. J. and Scott, D. A. 2019. Critical appraisal of nonrandomized studies—a review of recommended and commonly used tools. *Journal of evaluation in clinical practice*, 25 (1): 44-52.

Ramos-Remus, C., Gutierrez-Ureña, S. and Davis, P. 1999. Epidemiology of complementary and alternative practices in rheumatology. *Rheumatic Disease Clinics of North America*, 25 (4): 789-804.

Rantanen, P., Airaksinen, O. and Penttinen, E. 1994. Paradoxical variation of strength determinants with different rotation axes in trunk flexion and extension strength tests. *European journal of applied physiology and occupational physiology*, 68 (4): 322-326.

Rao, J. K., Arick, R., Mihaliak, K. and Weinberger, M. 1998. Using focus groups to understand arthritis patients' perceptions about unconventional therapy. *Arthritis Care Res*, 11 (4): 253-260.

Raven, E. E., Haverkamp, D., Sierevelt, I. N., Van Montfoort, D. O., Pöll, R. G., Blankevoort, L. and Tak, P. P. 2008. Construct validity and reliability of the disability ofArm, shoulder and hand questionnaire for upper extremity complaints in rheumatoid arthritis. *J Rheumatol*, 35 (12): 2334-2338.

Raza, K., Holers, V. M. and Gerlag, D. 2019. Nomenclature for the Phases of the Development of Rheumatoid Arthritis. *Clinical Therapeutics*, 41 (7): 1279-1285.

Reeves, J. L., Jaeger, B. and Graff-Radford, S. B. 1986. Reliability of the pressure algometer as a measure of myofascial trigger point sensitivity. *Pain*, 24 (3): 313-321.

Reeves, N. D., Narici, M. V. and Maganaris, C. N. 2006. Myotendinous plasticity to ageing and resistance exercise in humans. *Experimental physiology*, 91 (3): 483-498.

Rezaei, H., Af Klint, E., Hammer, H. B., Terslev, L., D'Agostino, M. A., Kisten, Y. and Arnaud, L. 2017. Analysis of correlation and causes for discrepancy between quantitative and semi-quantitative Doppler scores in synovitis in rheumatoid arthritis. *Rheumatology*, 56 (2): 255-262.

Rezaei, H., Torp-Pedersen, S., af Klint, E., Backheden, M., Kisten, Y., Györi, N. and van Vollenhoven, R. F. 2014. Diagnostic utility of musculoskeletal ultrasound in patients with suspected arthritis–a probabilistic approach. *Arthritis research & therapy*, 16 (5): 1-8.

Richardson, G. W. 2007. The effect of differing clinical settings on chiropractic patients suffering from mechanical low back pain.

Ricker, T. J., Vergauwe, E. and Cowan, N. 2016. Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology*, 69 (10): 1969-1995.

Ridgley, L. A., Anderson, A. E. and Pratt, A. G. 2018. What are the dominant cytokines in early rheumatoid arthritis? *Current opinion in rheumatology*, 30 (2): 207.

Rindfleisch, A. J. and Muller, D. 2005. Diagnosis and management of rheumatoid arthritis. *Am Fam Physician*, 72 (6): 1037-1047.

Ritchie, D. M., Boyle, J. A., MCINNES, J. M., Jasani, M. K., Dalakos, T. G., Grieveson, P. and Buchanan, W. W. 1968. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *QJM: An International Journal of Medicine*, 37 (3): 393-406.

Robins, J. M., Hernán, M. A. and Siebert, U. 2004. Effects of multiple interventions. *Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors*, 1: 2191-2230.

Roelofs, M. F., Boelens, W. C., Joosten, L. A., Abdollahi-Roodsaz, S., Geurts, J., Wunderink, L. U., Schreurs, B. W., van den Berg, W. B. and Radstake, T. R. 2006. Identification of small heat shock protein B8 (HSP22) as a novel TLR4 ligand and potential involvement in the pathogenesis of rheumatoid arthritis. *The Journal of Immunology*, 176 (11): 7021-7027.

Romanowski, M., Romanowska, J. and Grzeskowiak, M. 2012. A comparison of the effects of deep tissue massage and therapeutic massage on chronic low back pain. *Stud Health Technol Inform*, 176: 411-414.

Romanowski, M., Straburzyńska-Lupa, A., Romanowska, A. and Lorenc, R. 2014. AB1125 A Comparison of the Effects of Kinesiotherapy and Post-Isometric Relaxation on Knee Pain in Patients with Rheumatoid Arthritis. *Annals of the rheumatic diseases*, 73 (Suppl 2): 1174-1174.

Romanowski, M. W., Majchrzak, M., Kostiukow, A., Malak, R. and Samborski, W. 2018. Effect of Manual Therapy on Pain Sensation and Hand Dexterity and Grip Strength in a Patient with RA. Didactic Case Study. *Ortopedia, traumatologia, rehabilitacja*, 20 (2): 157-162.

Romanowski, M. W., Špiritović, M., Romanowski, W. and Straburzyńska-Lupa, A. 2020. Manual Therapy (Postisometric Relaxation and Joint Mobilization) in Knee Pain and Function Experienced by Patients with Rheumatoid Arthritis: A Randomized Clinical Pilot Study. *Evidence-Based Complementary and Alternative Medicine*, 2020

Rønningen, A. and Kjeken, I. 2008. Effect of an intensive hand exercise programme in patients with rheumatoid arthritis. *Scandinavian journal of occupational therapy*, 15 (3): 173-183.

Roopchand, A. K. 2014. A Systematic review of the non-invasive therapeutice modalities in the treatment of myofascial pain and dysfunction. Masters of Technology: Chiropractic, Durban University of Technology.

Ropes, M. W., Bennett, G. A., Cobb, S., Jacox, R. and Jessar, R. A. 1959. 1958 revision of diagnostic criteria for rheumatoid arthritis. *JBJS*, 41 (4): 781-782.

Rothwell, P. M. 2006. Factors that can affect the external validity of randomised controlled trials. *PLoS clinical trials*, 1 (1): e9.

Roubenoff, R., Roubenoff, R. A., Cannon, J. G., Kehayias, J. J., Zhuang, H., Dawson-Hughes, B., Dinarello, C. A. and Rosenberg, I. H. 1994. Rheumatoid cachexia: cytokine-driven hypermetabolism

accompanying reduced body cell mass in chronic inflammation. *The Journal of clinical investigation*, 93 (6): 2379-2386.

Russell, M. A., Hill, K. D., Blackberry, I., Day, L. M. and Dharmage, S. C. 2008. The reliability and predictive accuracy of the falls risk for older people in the community assessment (FROP-Com) tool. *Age and ageing*, 37 (6): 634-639.

Ryan, R., Santesso, N., Lowe, D., Hill, S., Grimshaw, J., Prictor, M., Kaufman, C., Cowie, G. and Taylor, M. 2014. Interventions to improve safe and effective medicines use by consumers: an overview of systematic reviews. *The Cochrane database of systematic reviews*, 2014 (4): CD007768-CD007768.

Sackett, D. L., Rosenberg, W., Gray, J., Haynes, R. B. and Richardson, W. S. 2007. Evidence based medicine: what it is and what it isn't. 1996. *Clinical orthopaedics and related research*, 455: 3.

Safiri, S., Kolahi, A. A., Hoy, D., Smith, E., Bettampadi, D., Mansournia, M. A., Almasi-Hashiani, A., Ashrafi-Asgarabad, A., Moradi-Lakeh, M. and Qorbani, M. 2019. Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the Global Burden of Disease study 2017. *Annals of the rheumatic diseases*, 78 (11): 1463-1471.

Sakellariou, G., Scire, C. A., Zambon, A., Caporali, R. and Montecucco, C. 2013. Performance of the 2010 classification criteria for rheumatoid arthritis: a systematic literature review and a meta-analysis. *PLOS ONE*, 8 (2): e56528.

Sambajon, V. V., Cillo Jr, J. E., Gassner, R. J. and Buckley, M. J. 2003. The effects of mechanical strain on synovial fibroblasts. *Journal of oral and maxillofacial surgery*, 61 (6): 707-712.

Sandoz, R. 1976. Some physical mechanisms and effects of spinal adjustments. *Ann Swiss Chirop Assoc*, 6: 91-141.

Sayers, E., MacGregor, A. and Carding, S. R. 2018. Drug-microbiota interactions and treatment response: Relevance to rheumatoid arthritis. *AIMS microbiology*, 4 (4): 642-654.

Sbarra, D. A. 2009. Marriage protects men from clinically meaningful elevations in C-reactive protein: results from the National Social Life, Health, and Aging Project (NSHAP). *Psychosomatic medicine*, 71 (8): 828.

Schaible, H.-G. and Grubb, B. D. 1993. Afferent and spinal mechanisms of joint pain. *Pain*, 55 (1): 5-54.

Schlaeger, J. M., Roach, K., Golas, M., Takayama, M. and Wilkie, D. J. 2018. Treatment-seeking behaviors of persons with rheumatoid arthritis. *Journal of Holistic Nursing*, 36 (2): 179-191.

Scholten, C., Brodowicz, T., Graninger, W., Gardavsky, I., Pils, K., Pesau, B., Eggl-Tyl, E., Wanivenhaus, A. and Zielinski, C. C. 1999. Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. *Archives of physical medicine and rehabilitation*, 80 (10): 1282-1287.

Schreiber, E. H. and Schreiber, K. N. 1998. Use of hypnosis and Jacobson's relaxation techniques for improving academic achievement of college students. *Perceptual and motor skills*, 86 (1): 85-86.

Schulz, K. F., Altman, D. G. and Moher, D. 2010. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med*, 8: 18.

Schwartz, J. E., Jandorf, L. and Krupp, L. B. 1993. The measurement of fatigue: a new instrument. *Journal of psychosomatic research*, 37 (7): 753-762.

Scollon, R., Scollon, S. W. and Jones, R. H. 2011. *Intercultural Communication: A Discourse Approach*. Wiley.

Scott, D. and Houssien, D. 1996. Joint assessment in rheumatoid arthritis. *Rheumatology*, 35 (suppl_2): 14-18.

Scott, D. L., Panayi, G.S.,, Vanriel, P. L. C. M., Smolen, J., and Vandeputte, L. B. A. 1992. Assessing disease-activity in rheumatoid-arthritis. In: Proceedings of *Arthritis and rheumatism*. Lippincott-Raven Publ 227 east Washington SQ, Philadelphia, PA 19106, S202-S202.

Scott, J. and Huskisson, E. 1979. Vertical or horizontal visual analogue scales. *Annals of the rheumatic diseases*, 38 (6): 560.

Sedgwick, P. and Greenwood, N. 2015. Understanding the Hawthorne effect. *Bmj*, 351

Sellgren, C., Frisell, T., Lichtenstein, P., Landèn, M. and Askling, J. 2014. The Association Between Schizophrenia and Rheumatoid Arthritis: A Nationwide Population-Based Swedish Study on Intraindividual and Familial Risks. *Schizophrenia Bulletin*, 40 (6): 1552-1559.

Senderovich, H., Ip, M. L., Berall, A., Karuza, J., Gordon, M., Binns, M., Wignarajah, S., Grossman, D. and Dunal, L. 2016. Therapeutic Touch® in a geriatric Palliative Care Unit–A retrospective review. *Complementary therapies in clinical practice*, 24: 134-138.

Seo, J.-W., Yang, E.-J., Yoo, K.-H. and Choi, I.-H. 2015. Macrophage Differentiation from Monocytes Is Influenced by the Lipid Oxidation Degree of Low Density Lipoprotein. *Mediators of Inflammation*, 2015: 235797.

Sharma, A., Khan, R., Gupta, N., Sharma, A., Zaheer, M. S., Abbas, M. and Khan, S. A. 2018. Acute phase reactant, Pentraxin 3, as a novel marker for the diagnosis of rheumatoid arthritis. *Clin Chim Acta*, 480: 65-70.

Sharma, L. and Pai, Y.-C. 1997. Impaired proprioception and osteoarthritis. *Current opinion in rheumatology*, 9 (3): 253-258.

Shegarfi, H., Naddafi, F. and Mirshafiey, A. 2012. Natural Killer Cells and Their Role in Rheumatoid Arthritis: Friend or Foe? *The Scientific World Journal*, 2012: 491974.

Sherrington, C., Herbert, R., Maher, C. and Moseley, A. 2000. PEDro. A database of randomized trials and systematic reviews in physiotherapy. *Manual therapy*, 5 (4): 223-226.

Shinde, S. B. and Varadharajulu, G. 2017. Effect of therapeutic exercise programme in adults with early rheumatoid arthritis. *Indian J Physiother Occup Ther*, 11: 76-80.

Shmueli, A. and Shuval, J. 2004. Use of complementary and alternative medicine in Israel: 2000 vs. 1993. *IMAJ-RAMAT GAN-*, 6 (1): 3-8.

Shumway-Cook, A. and Horak, F. B. 1986. Assessing the influence of sensory interaction on balance: suggestion from the field. *Physical therapy*, 66 (10): 1548-1550.

Siegert, C., Vleming, L.-J., Van-Denbroucke, J. and Cats, A. 1984. Measurement of disability in Dutch rheumatoid arthritis patients. *Clinical rheumatology*, 3 (3): 305-309.

Siering, U., Eikermann, M., Hausner, E., Hoffmann-Eßer, W. and Neugebauer, E. A. 2013. Appraisal tools for clinical practice guidelines: a systematic review. *PLOS ONE*, 8 (12): e82915.

Sinigaglia, L., Varenna, M., Girasole, G. and Bianchi, G. 2006. Epidemiology of osteoporosis in rheumatic diseases. *Rheumatic Disease Clinics*, 32 (4): 631-658.

Slim, K., Nini, E., Forestier, D., Kwiatkowski, F., Panis, Y. and Chipponi, J. 2003. Methodological index for non-randomized studies (MINORS): development and validation of a new instrument. *ANZ journal of surgery*, 73 (9): 712-716.

Slosberg, M. 1990. Spinal learning: central modulation of pain processing and long-term alteration of interneuronal excitability as a result of nociceptive peripheral input. *Journal of manipulative and physiological therapeutics*, 13 (6): 326-336.

Smedstad, L., Moum, T., Guillemin, F., Kvien, T., Finch, M., Suurmeijer, T. and Van Den Heuvel, W. 1996. Correlates of functional disability in early rheumatoid arthritis: a cross-sectional study of 706 patients in four European countries. *Rheumatology*, 35 (8): 746-751.

Smith, R. C., Fortin, A. H., Dwamena, F. and Frankel, R. M. 2013. An evidence-based patient-centered method makes the biopsychosocial model scientific. *Patient education and counseling*, 91 (3): 265-270.

Smolen, J. S., Aletaha, D., Koeller, M., Weisman, M. H. and Emery, P. 2007. New therapies for treatment of rheumatoid arthritis. *The Lancet*, 370 (9602): 1861-1874.

Smolen, J. S., Aletaha, D. and McInnes, I. B. 2016. Rheumatoid arthritis. *The Lancet*, 388 (10055): 2023-2038.

Smolen, J. S., Eberl, G., Breedveld, F. C., Jones, I., Leeming, M., Wylie, G. L. and Kirkpatrick, J. 1995. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 38 (1): 38-43.

Snibsøer, A. K., Ciliska, D., Yost, J., Graverholt, B., Nortvedt, M. W., Riise, T. and Espehaug, B. 2018. Self-reported and objectively assessed knowledge of evidence-based practice terminology among healthcare students: A cross-sectional study. *PLOS ONE*, 13 (7): e0200313.

Snyder-Halpern, R. and Verran, J. A. 1987. Instrumentation to describe subjective sleep characteristics in healthy subjects. *Research in nursing & health*, 10 (3): 155-163.

So, A., Varisco, P. A., Kemkes-Matthes, B., Herkenne-Morard, C., Chobaz-Peclat, V., Gerster, J. C. and Busso, N. 2003. Arthritis is linked to local and systemic activation of coagulation and fibrinolysis pathways. *Journal of Thrombosis and Haemostasis*, 1 (12): 2510-2515.

Soubrier, M. and Dougados, M. 2005. Selecting criteria for monitoring patients with rheumatoid arthritis. *Joint Bone Spine*, 72 (2): 129-134.

Spiegel, J. S., Spiegel, T. M., Ward, N. B., Paulus, H. E., Leake, B. and Kane, R. L. 1986. Rehabilitation for rheumatoid arthritis patients. A controlled trial. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 29 (5): 628-637.

Spielberger, C. D. 1970. Manual for the State-Trait Anxietry, Inventory. *Consulting Psychologist*,

Spielberger, C. D. 1972. Current trends in theory and research. *Anxiety: current trends in theory and research*, 1

Spitz, P. and Fries, J. 1987. The present and future of comprehensive outcome measures for rheumatic diseases. *Clinical rheumatology*, 6 (2): 105-111.

Stang, A. 2010. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European journal of epidemiology*, 25 (9): 603-605.

Steflitsch, W. and Steflitsch, M. 2008. Clinical aromatherapy. *Journal of Men's Health*, 5 (1): 74-85.

Stein, M., Wallston, K. and Nicassio, P. 1988. Factor structure of the arthritis helplessness index. *J Rheumatol*, 15 (3): 427-432.

Steinberg, E., Greenfield, S., Wolman, D. M., Mancher, M. and Graham, R. 2011. *Clinical practice guidelines we can trust*. National Academies Press.

Steinbrocker, O., Traeger, C. H. and Batterman, R. C. 1949. Therapeutic criteria in rheumatoid arthritis. *Journal of the American Medical Association*, 140 (8): 659-662.

Stenström, C. H. and Minor, M. A. 2003. Evidence for the benefit of aerobic and strengthening exercise in rheumatoid arthritis. *Arthritis Care & Research*, 49 (3): 428-434.

Stern, C. and Kleijnen, J. 2020. *Language bias in systematic reviews: you only get out what you put in*: LWW.

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T. and Boutron, I. 2016. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Bmj*, 355

Stillman, B. and McMeeken, J. 1995. A video-based version of the pendulum test: technique and normal response. *Archives of physical medicine and rehabilitation*, 76 (2): 166-176.

Stillman, B. C., McMeeken, J. M. and Macdonell, R. A. 1998. Aftereffects of resisted muscle contractions on the accuracy of joint position sense in elite male athletes. *Archives of physical medicine and rehabilitation*, 79 (10): 1250-1254.

Strasser, B., Leeb, G., Strehblow, C., Schobersberger, W., Haber, P. and Cauza, E. 2011. The effects of strength and endurance training in patients with rheumatoid arthritis. *Clinical rheumatology*, 30 (5): 623-632.

Straub, R. H. 2007. The complex role of estrogens in inflammation. *Endocrine reviews*, 28 (5): 521-574.

Sutoo, D. e. and Akiyama, K. 2003. Regulation of brain function by exercise. *Neurobiology of disease*, 13 (1): 1-14.

Swanson, K. V., Deng, M. and Ting, J. P. Y. 2019. The NLRP3 inflammasome: molecular activation and regulation to therapeutics. *Nature Reviews Immunology*, 19 (8): 477-489.

Sweaney, J. 2005. *WHO giudelines on basic training and safety in chiropractic*: Geneva.

Symmons, D., Turner, G., Webb, R., Asten, P., Barrett, E., Lunt, M., Scott, D. and Silman, A. 2002. The prevalence of rheumatoid arthritis in the United Kingdom: new estimates for a new century. *Rheumatology*, 41 (7): 793-800.

Tan, I. J., Peeva, E. and Zandman-Goddard, G. 2015. Hormonal modulation of the immune system—a spotlight on the role of progestogens. *Autoimmunity reviews*, 14 (6): 536-542.

Taneja, V. and David, C. S. 2001. Lessons from animal models for human autoimmune diseases. *Nature Immunology*, 2 (9): 781-784.

Thompson, P. D., Buchner, D., Piña, I. L., Balady, G. J., Williams, M. A., Marcus, B. H., Berra, K., Blair, S. N., Costa, F. and Franklin, B. 2003. Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease: a statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Rehabilitation, and Prevention) and the Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity). *Circulation*, 107 (24): 3109-3116.

Tick, H., Nielsen, A., Pelletier, K. R., Bonakdar, R., Simmons, S., Glick, R., Ratner, E., Lemmon, R. L., Wayne, P. and Zador, V. 2018. Evidence-based nonpharmacologic strategies for comprehensive pain care: the consortium pain task force white paper. *Explore*, 14 (3): 177-211.

Tierney, W. M. 2001. Improving clinical decisions and outcomes with information: a review. *International Journal of Medical Informatics*, 62 (1): 1-9.

Tiran, D. and Chummun, H. 2005. The physiological basis of reflexology and its use as a potential diagnostic tool. *Complementary therapies in clinical practice*, 11 (1): 58-64.

Tokem, Y., Parlar Kilic, S., Ozer, S., Nakas, D. and Argon, G. 2014. A multicenter analysis of the use of complementary and alternative medicine in Turkish patients with rheumatoid arthritis: holistic nursing practice review copy. *Holist Nurs Pract*, 28 (2): 98-105.

Torgerson, D. J. and Sibbald, B. 1998. Understanding controlled trials. What is a patient preference trial? *BMJ: British Medical Journal*, 316 (7128): 360.

Tranfield, D., Denyer, D. and Smart, P. 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management*, 14 (3): 207-222.

Tucker, M. and Kirwan, J. R. 1991. Does patient education in rheumatoid arthritis have therapeutic potential? *Annals of the rheumatic diseases*, 50 (Suppl 3): 422.

Uhlig, T., Kvien, T. K. and Pincus, T. 2009. Test–retest reliability of disease activity core set measures and indices in rheumatoid arthritis. *Annals of the rheumatic diseases*, 68 (6): 972-975.

Unbiased Research. 2017. *Pre-hoc vs Post-hoc analysis: Whats the difference*. Available: http://unbiasedresearch.blogspot.com/2017/04/pre-hoc-vs-post-hoc-analysis-whats.html (Accessed

Valesini, G., Gerardi, M. C., Iannuccelli, C., Pacucci, V. A., Pendolino, M. and Shoenfeld, Y. 2015. Citrullination and autoimmunity. *Autoimmunity reviews*, 14 (6): 490-497.

Van den Berg, M., Ronday, H., Peeters, A., Le Cessie, S., Van Der Giesen, F., Breedveld, F. and Vliet Vlieland, T. 2006. Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 55 (6): 935-945.

Van den Ende, C., Breedveld, F., Le Cessie, S., Dijkmans, B., De Mug, A. and Hazes, J. 2000. Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial. *Annals of the rheumatic diseases*, 59 (8): 615-621.

Van den Ende, C., Hazes, J., le Cessie, S., Mulder, W. J., Belfor, D. G., Breedveld, F. C. and Dijkmans, B. 1996. Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomised clinical trial. *Annals of the rheumatic diseases*, 55 (11): 798-805.

Van der Heijde, D., Dankert, T., Nieman, F., Rau, R. and Boers, M. 1999. Reliability and sensitivity to change of a simplification of the Sharp/van der Heijde radiological assessment in rheumatoid arthritis. *Rheumatology*, 38 (10): 941-947.

van Lankveld, W., van't Pad Bosch, P., Bakker, J., Terwindt, S., Franssen, M. and van Kiel, P. 1996. Sequential occupational dexterity assessment (SODA): a new test to measure hand disability. *Journal of hand therapy*, 9 (1): 27-32.

van Riel, P. L. and Fransen, J. 2007. Established rheumatoid arthritis: clinical assessments. *Best Practice & Research Clinical Rheumatology*, 21 (5): 807-825.

Vernon, H., Dhami, M., Howley, T. and Annett, R. 1986. Spinal manipulation and beta-endorphin: a controlled study of the effect of a spinal manipulation on plasma beta-endorphin levels in normal males. *Journal of manipulative and physiological therapeutics*, 9 (2): 115-123.

Vernon, H. and Mrozek, J. 2005. A revised definition of manipulation. *Journal of Manipulative & Physiological Therapeutics*, 28 (1): 68-72.

Vigotsky, A. D. and Bruhns, R. P. 2015. The Role of Descending Modulation in Manual Therapy and Its Analgesic Implications: A Narrative Review. *Pain Res Treat*, 2015: 292805.

Villafañe, J. H., Cleland, J. A. and Fernandez-de-Las-Peñas, C. 2013. Bilateral sensory effects of unilateral passive accessory mobilization in patients with thumb carpometacarpal osteoarthritis. *Journal of manipulative and physiological therapeutics*, 36 (4): 232-237.

Vlayen, J., Aertgeerts, B., Hannes, K., Sermeus, W. and Ramaekers, D. 2005. A systematic review of appraisal tools for clinical practice guidelines: multiple similarities and one common deficit. *International Journal for Quality in Health Care*, 17 (3): 235-242.

Vliet Vlieland, T., van Den Ende, C., Breedveld, F. and Hazes, J. 1993. Evaluation of joint mobility in rheumatoid arthritis trials: the value of the EPM-range of motion scale. *J Rheumatol*, 20 (12): 2010-2014.

Vljeland, T. V., Zwinderman, A., Vandenbroucke, J., Breedveld, F. and Hazes, J. 1996. A randomized clinical trial of in-patient multidisciplinary treatment versus routine out-patient care in active rheumatoid arthritis. *Rheumatology*, 35 (5): 475-482.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C. and Vandenbroucke, J. P. 2007. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Bmj*, 335 (7624): 806-808.

Walker, B. R., Colledge, N. R., Ralston, S. H. and Penman, I. D. 2013. *Davidson's Principles and Practice of Medicine E-Book*. Elsevier Health Sciences.

Wall, P. D., McMahon, S. B. and Koltzenburg, M. 2006. *Wall and Melzack's textbook of pain*. Elsevier/Churchill Livingstone.

Wallman, K. E., Morton, A. R., Goodman, C. and Grove, R. 2004. Physiological responses during a submaximal cycle test in chronic fatigue syndrome. *Medicine and science in sports and exercise*, 36: 1682-1688.

Walpole, S. C. 2019. Including papers in languages other than English in systematic reviews: important, feasible, yet often omitted. *Journal of clinical epidemiology*, 111: 127-134.

Wang, Y., Chen, X., Gong, J. and Yan, Y. 2016. Relationships between stress, negative emotions, resilience, and smoking: Testing a moderated mediation model. *Substance use & misuse*, 51 (4): 427-438.

Wasserman, A. M. 2011. Diagnosis and management of rheumatoid arthritis. *Am Fam Physician*, 84 (11): 1245-1252.

Wei, S.-T., Sun, Y.-H., Zong, S.-H. and Xiang, Y.-B. 2015. Serum levels of IL-6 and TNF-α may correlate with activity and severity of rheumatoid arthritis. *Medical science monitor: international medical journal of experimental and clinical research*, 21: 4030.

Weinberg, J. B., Pippen, A. M. and Greenberg, C. S. 1991. Extravascular fibrin formation and dissolution in synovial tissue of patients with osteoarthritis and rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 34 (8): 996-1005.

Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M. and P Tugwell. 2003. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*.

Wells, G., Shea, B. and O'Connell, D. 2003. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Slide 1–39, Manual: 1–4, and Scale: 1–2*.

Wells, M., Williams, B., Treweek, S., Coyle, J. and Taylor, J. 2012. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials*, 13 (1): 1-17.

Westby, M. D., Wade, J. P., Rangno, K. K. and Berkowitz, J. 2000. A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. *J Rheumatol*, 27 (7): 1674-1680.

Whaley, M., Brubaker, P., Otto, R. and Armstrong, L. 2006. Health-related physical fitness testing and interpretation. *ACSM's Guidelines for Exercise Testing and Prescription. Baltimore: Lippincott Williams & Wilkins*: 55-92.

Whiting, P., Savović, J., Higgins, J. P., Caldwell, D. M., Reeves, B. C., Shea, B., Davies, P., Kleijnen, J. and Churchill, R. 2016. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *Journal of clinical epidemiology*, 69: 225-234.

Więckiewicz, W., Woźniak, K., Piątkowska, D., Szyszka-Sommerfeld, L. and Lipski, M. 2015. The diagnostic value of pressure algometry for temporomandibular disorders. *BioMed research international*, 2015

Wilczynska, M. M., Condliffe, A. M. and McKeon, D. J. 2013. Coexistence of Bronchiectasis and Rheumatoid Arthritis: Revisited. *Respiratory Care*, 58 (4): 694.

Williams, M. A., Williamson, E. M., Heine, P. J., Nichols, V., Glover, M. J., Dritsaki, M., Adams, J., Dosanjh, S., Underwood, M. and Rahman, A. 2015. Strengthening And stretching for Rheumatoid Arthritis of the Hand (SARAH). A randomised controlled trial and economic evaluation.

Williams, S. B., Brand, C. A., Hill, K. D., Hunt, S. B. and Moran, H. 2010. Feasibility and outcomes of a home-based exercise program on improving balance and gait stability in women with lower-limb osteoarthritis or rheumatoid arthritis: a pilot study. *Archives of physical medicine and rehabilitation*, 91 (1): 106-114.

Wirtz, P. H., Ehlert, U., Emini, L. and Suter, T. 2008. Higher body mass index (BMI) is associated with reduced glucocorticoid inhibition of inflammatory cytokine production following acute psychosocial stress in men. *Psychoneuroendocrinology*, 33 (8): 1102-1110.

Wolfe, F., Clauw, D. J., Fitzcharles, M. A., Goldenberg, D. L., Katz, R. S., Mease, P., Russell, A. S., Russell, I. J., Winfield, J. B. and Yunus, M. B. 2010. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care & Research*, 62 (5): 600-610.

Wright, H. L., Moots, R. J., Bucknall, R. C. and Edwards, S. W. 2010. Neutrophil function in inflammation and inflammatory diseases. *Rheumatology*, 49 (9): 1618-1631.

Yang, M., Feng, X., Ding, J., Chang, F. and Chen, X. 2017. Nanotherapeutics relieve rheumatoid arthritis. *Journal of Controlled Release*, 252: 108-124.

Yip, Y. B. and Tam, A. C. Y. 2008. An experimental study on the effectiveness of massage with aromatic ginger and orange essential oil for moderate-to-severe knee pain among the elderly in Hong Kong. *Complement Ther Med*, 16 (3): 131-138.

Zegarra-Parodi, R., Park, P. Y. S., Heath, D. M., Makin, I. R. S., Degenhardt, B. F. and Roustit, M. 2015. Assessment of skin blood flow following spinal manual therapy: A systematic review. *Manual therapy*, 20 (2): 228-249.

Zeidler, H. 2013. Systemic literature review of the performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: good news of debatable significance. *Annals of the rheumatic diseases*, 72 (8): e21-e21.

Zeng, X., Zhang, Y., Kwong, J. S., Zhang, C., Li, S., Sun, F., Niu, Y. and Du, L. 2015. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *Journal of evidence-based medicine*, 8 (1): 2-10.

Zernicke, J., Kedor, C., Müller, A., Burmester, G.-R., Reißhauer, A. and Feist, E. 2016. A prospective pilot study to evaluate an animated home-based physical exercise program as a treatment option for patients with rheumatoid arthritis. *BMC musculoskeletal disorders*, 17 (1): 1-12.

Zhang, X., Zhang, D., Jia, H., Feng, Q., Wang, D., Liang, D., Wu, X., Li, J., Tang, L., Li, Y., Lan, Z., Chen, B., Li, Y., Zhong, H., Xie, H., Jie, Z., Chen, W., Tang, S., Xu, X., Wang, X., Cai, X., Liu, S., Xia, Y., Li, J., Qiao, X., Al-Aama, J. Y., Chen, H., Wang, L., Wu, Q.-j., Zhang, F., Zheng, W., Li, Y., Zhang, M., Luo, G., Xue, W., Xiao, L., Li, J., Chen, W., Xu, X., Yin, Y., Yang, H., Wang, J., Kristiansen, K., Liu, L., Li, T., Huang, Q., Li, Y. and Wang, J. 2015. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nature Medicine*, 21 (8): 895-905.

Zhao, S., Otieno, F., Akpan, A. and Moots, R. J. 2017. Complementary and Alternative Medicine Use in Rheumatoid Arthritis: Considerations for the Pharmacological Management of Elderly Patients. *Drugs Aging*, 34 (4): 255-264.

# APPENDICES

Appendix A: Faculty of Research Committee (FRC) letter of approval

Appendix B: PROSPERO registration / approval

Appendix C: DALRO approval (Copyright permission)

Appendix D: PRISMA flow diagram / outline

Appendix E: Master list

Appendix F: Memorandum of Agreement (MoA)

Appendix G: PEDro scale template

Appendix H: NOS scale template

Appendix I: Liddle scale template

Appendix J: Permission letter from Mrs A Finlayson

Appendix K: PRISMA checklist

## Appendix A: Faculty of Research Committee (FRC) letter of approval


21 August, 2020

Mr K Sclanders
Student No: 21516881

Berwin Farm
Bergville
3500

Dear Mr Sclanders

**MASTER OF TECHNOLOGY: CHIROPRACTIC**

I am pleased to advise that:

1. The Faculty Research Committee approved the following:

    (i) Your research proposal and dissertation title, being:

**A systematic review of non-invasive manual therapy in the management of rheumatoid arthritis patients.**

**Please note: ANY PROPOSED CHANGES in the DISSERTATION TITLE require the approval of your supervisor and the Faculty Research Committee.**

    (ii) Supervisor – **Dr C Korporaal**

2. Your request for funding totalling **R 8 000.00** subject to any literature referred to in Section A of the PG 4a form being accessioned by this University, and any equipment purchased shall become the property of the department.

**NOTE: - This funding is not paid directly to you but is controlled by the Faculty Office. Any proposed changes to this funding allocation needs the approval of your supervisor, and Faculty Research Committee**

The University Research Committee has stipulated that:

(a) Ownership of any patent registered in respect of the results of your Master's studies is retained by you as the initiator of the project;

(b) Should you make any drift from the results of your Master's studies, you will be required to repay pro rata, the **R 5 000.00** investment which the University Research Committee has made in approving your request for funding;

(c) If the Durban University of Technology provided the equipment/materials for the creation of artefacts, this cost would be refunded to the University if such artefacts were sold and

(d) Durban University of Technology is given first refusal in respect of any possible future sale by you of any patent that may be registered in respect of your said project.

(e) All journal articles, referenced in your dissertation, are to accompany your ring-bound copies when submitting for examination purposes.

Should you experience any problems relating to your research studies, your supervisor must be informed as soon as possible. If the difficulty persists, you must then approach your Head of Department and thereafter the Executive Dean of the Faculty.

Yours sincerely

**Ms S Perumal**
**FACULTY RESEARCH OFFICER**

_____ _____

**Student's signature in acceptance of the conditions contained herein.** **Date:**

## Appendix B: PROSPERO registration / approval



You have 1 records

### My other records

*These are records that have either been published or rejected and are not currently being worked on.*

| ID | Title | Status | Last edited |
|---|---|---|---|
| CRD42020205554 | A systematic review of non-invasive manual therapy in the management of rheumatoid arthritis patients. To enable PROSPERO to focus on COVID-19 registrations during the 2020 pandemic, this registration record was automatically published exactly as submitted. The PROSPERO team has not checked eligibility. | Registered | 22/09/2020 |

## Appendix C: DALRO approval (Copyright permission)

DRAMATIC, ARTISTIC and LITERARY RIGHTS ORGANISATION (Pty) LIMITED

www.dalro.co.za
dalro@dalro.co.za

SAMRO Place
20 De Korte Street
Braamfontein
2001 Johannesburg

P O Box 31627
2017 Braamfontein
Telephone 086 12 DALRO
International +27 (0)11 712 8000
Fax +27 (0)86 648 6299

Affiliated to the International Confederation of Copyright Societies – CISAC
Member of the International Federation of Reproduction Rights Organisations – IFRRO
Reg. No. 1967/005018/07

**TO:** DURBAN UNIVERSITY OF TECHNOLOGY

Intellectual Property
Durban University of Technology
P O Box 1334
DURBAN

Dear Ramika,

**REPROGRAPHIC REPRODUCTION RIGHTS QUOTATION FOR (CHIROPRACTIC/RESEARCH DISSERTATION/KYLE DONALD SCLANDERS) (0000010362/01):**

DALRO hereby grants permission for the reprographic reproduction of extract(s) from the publication(s) as per the schedule following. The reproduction royalties are indicated next to the title(s).

**Please note that this permission is subject to the following conditions:**

1. The copied extract(s) may only be distributed by the licensing entity identified on the schedule, and only to persons enrolled for the specific course of study at the licensing entity.
2. Should the licensing entity decide to sell the copied extract(s) to the persons enrolled for the course of study, the selling price per unit shall not exceed the cost per unit incurred by the licensing entity.
3. The copied extract(s) may only be distributed from a point which is under the direct control of, and staffed by, the licensing entity, and no external entity or concern, whether commercial or non-commercial, may act as distributor.

Only upon acceptance, in writing, of this quotation will a Tax Invoice be issued. Please let us know if this reproduction will not be going ahead at all to enable us to cancel the quotation. Should the information on which the quotation is based, have changed, the new information need to be supplied to enable us to prepare a revised quotation.

Yours sincerely

**RIGHTS ADMINISTRATION**

**REPROGRAPHIC REPRODUCTION RIGHTS QUOTATION: 14/08/2020**

**QUOTATION NUMBER: 0000010362/01**

DALRO
DRAMATIC, ARTISTIC and LITERARY RIGHTS ORGANISATION (Pty) LIMITED

| | |
|---|---|
| LICENSING ENTITY: | DURBAN UNIVERSITY OF TECHNOLOGY |
| LECTURER: | KYLE DONALD SCLANDERS |
| DEPARTMENT: | CHIROPRACTIC |
| COURSE: | RESEARCH DISSERTATION |
| DISSEMINATION: | SINGLE ITEM HANDOUT |
| USAGE PERIOD: | 01/06/2020 - 31/12/2020 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 001/<br>001 | **JOURNAL OF RHEUMATOLOGY, THE**<br>RANDOMIZED CONTROLLED TRIAL TO EVALUATE THE EFFICACY OF COMMUNITY BASED PHYSICAL THERAPY IN THE TREATMENT OF PEOPLE WITH RHEUMATOID ARTHRITIS.<br>(BELL M.J.)<br>159329 | 2 | 7 | 14 | R1.07 | R42.63* | R6.39 | R49.02 |
| 002/<br>002 | **PHYSICAL THERAPY**<br>EFFECTS OF A GROUP-BASED EXERCISE AND EDUCATIONAL PROGRAM ON PHYSICAL PERFORMANCE AND DISEASE SELF-MANAGEMENT IN RHEUMATOID ARTHRITIS: A RANDOMIZED CONTROLLED STUDY.<br>(BREEDLAND I. (et al))<br>159330 | 2 | 15 | 30 | R1.07 | R42.63* | R6.39 | R49.02 |
| 003/<br>003 | **ARTHRITIS AND RHEMATISM**<br>COACHING PATIENTS WITH EARLY RHEUMATOID ARTHRITIS TO HEALTHY PHYSICAL ACTIVITY: A MULTICENTER, RANDOMIZED, CONTROLLED STUDY<br>(BRODIN Nina (et al))<br>159331 | 2 | 7 | 14 | R1.07 | R42.63* | R6.39 | R49.02 |
| 004/<br>004 | **ARTHRITIS AND RHEMATISM**<br>PHYSICAL AND EXERCISE THERAPY FOR TREATMENT OF THE RHEUMATOID HAND.<br>(BULJINA A.I. (et al))<br>159332 | 2 | 6 | 12 | R1.07 | R42.63* | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 005/<br>005 | **PSYCHOLOGY, HEALTH AND MEDICINE**<br>OUTCOME EVALUATION STUDY OF THE RHEUMATOID ARTHRITIS SELF-MANAGEMENT PROGRAMME IN HONG KONG<br>(CHUI D.Y.Y., LAU J.S.K., YAU I.T.Y.)<br>159333 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 006/<br>006 | **JOURNAL OF THE CANADIAN CHIROPRACTIC ASSOCIATION**<br>USE OF SPINAL MANIPULATION IN A RHEUMATOID PATIENT PRESENTING WITH ACUTE THORACIC PAIN: A CASE REPORT.<br>(CHUNG C.L., MIOR S.A.)<br>159334 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 007/<br>007 | **INTERNATIONAL JOURNAL OF THERAPEUTIC MASSAGE AND BODYWORK: RESEARCH, EDUCATION, AND PRACTICE**<br>SUSTAINED RELEASE MYOFASCIAL RELEASE AS TREATMENT FOR A PATIENT WITH COMPLICATIONS OF RHEUMATOID ARTHRITIS AND COLLAGENOUS COLITIS: A CASE REPORT<br>(CUBICKI E.E. (et al))<br>159335 | 2 | 9 | 18 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 008/<br>008 | **ARTHRITIS AND RHEMATISM**<br>IS A LONG-TERM HIGH-INTENSITY EXERCISE PROGRAM EFFECTIVE AND SAFE IN PATIENTS WITH RHEUMATOID ARTHRITIS? RESULTS OF A RANDOMIZED CONTROLLED TRIAL<br>(DE JONG Zuzana (et al))<br>159337 | 2 | 10 | 20 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 009/<br>009 | **ARTHRITIS AND RHEMATISM**<br>SLOWING OF BONE LOSS IN PATIENTS WITH RHEUMATOID ARTHRITIS BY LONG-TERM HIGH-INTENSITY EXERCISE: RESULTS OF A RANDOMIZED, CONTROLLED TRIAL.<br>(DE JONG Z. (et al))<br>159338 | 2 | 11 | 22 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 010/<br>010 | **ARTHRITIS CARE AND RESEARCH**<br>EFFECT OF ACTIVE HAND EXERCISE AND WAX BATH TREATMENT IN RHEUMATOID ARTHRITIS PATIENTS.<br>(DELLHAG B., WOLLERSJO I., BJELLE A.)<br>159339 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 011/<br>011 | **SCANDINAVIAN JOURNAL OF RHEUMATOLOGY**<br>PAIN THRESHOLD IN PATIENTS WITH RHEUMATOID ARTHRITIS AND EFFECT OF MANUAL OSCILLATIONS<br>(DHONDT W. (et al))<br>159340 | 2 | 6 | 12 | R1.07 | R42.63* | R6.39 | R49.02 |
| 012/<br>012 | **JOURNAL OF RHEUMATOLOGY, THE**<br>EFFECT OF EXERCISE ON SLEEP AND FATIGUE IN RHEUMATOID ARTHRITIS: A RANDOMIZED CONTROLLED STUDY<br>(DURCAN L. (et al))<br>159342 | 2 | 8 | 16 | R1.07 | R42.63* | R6.39 | R49.02 |
| 013/<br>013 | **BMC MUSCULOSKELETAL DISORDERS**<br>PRAGMATIC RANDOMISED CONTROLLED TRIAL OF HYDROTHERAPY AND LAND EXERCISES ON OVERALL WELL BEING AND QUALITY OF LIFE IN RHEUMATOID ARTHRITIS<br>(EVERSDEN Lis (et al))<br>159343 | 2 | 7 | 14 | R1.07 | R42.63* | R6.39 | R49.02 |
| 014/<br>014 | **COMPLEMENTARY THERAPIES IN CLINICAL PRACTICE: AN INTERNATIONAL JOURNAL**<br>RHEUMATOID ARTHRITIS IN UPPER LIMBS BENEFITS FROM MODERATE PRESSURE MASSAGE THERAPY<br>(FIELD Tiffany (et al))<br>159344 | 2 | 3 | 6 | R1.07 | R42.63* | R6.39 | R49.02 |
| 015/<br>015 | **MUSCULOSKELETAL CARE**<br>PERSONALIZED DIET AND EXERCISE RECOMMENDATIONS IN EARLY RHEUMATOID ARTHRITIS: A FEASIBILITY TRIAL.<br>(GARNER S. (et al))<br>159346 | 2 | 6 | 12 | R1.07 | R42.63* | R6.39 | R49.02 |
| 016/<br>016 | **BIOMED RESEARCH INTERNATIONAL**<br>EFFECTS OF WHOLE-BODY CRYOTHERAPY IN COMPARISON WITH OTHER PHYSICAL MODALITIES USED WITH KINESITHERAPY IN RHEUMATOID ARTHRITIS<br>(GIZINSKA M. (et al))<br>159348 | 2 | 7 | 14 | R1.07 | R42.63* | R6.39 | R49.02 |
| 017/<br>017 | **PAIN MANAGEMENT NURSING**<br>EFFECTS OF AROMATHERAPY MASSAGE AND REFLEXOLOGY ON PAIN AND FATIGUE IN PATIENTS WITH RHEUMATOID ARTHRITIS: A RANDOMIZED CONTROLLED TRIAL<br>(METIN Zehra Gok, OZDEMIR Leyla)<br>159350 | 2 | 10 | 20 | R1.07 | R42.63* | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 018/<br>018 | **RHEUMATOLOGY INTERNATIONAL: CLINICAL AND EXPERIMENTAL INVEST**<br>EFFECTS OF HOME-BASED EXERCISE PROGRAM ON THE FUNCTIONAL STATUS AND THE QUALITY OF LIFE IN PATIENTS WITH RHEUMATOID ARTHRITIS: 1-YEAR FOLLOW-UP STUDY<br>(KARATEPE Altnay Goksel (et al))<br>159352 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 019/<br>019 | **ARTHRITIS AND RHEMATISM**<br>RANDOMIZED TWOYEAR STUDY OF THE EFFECTS OF DYNAMIC STRENGTH TRAINING ON MUSCLE STRENGTH, DISEASE ACTIVITY, FUNCTIONAL CAPACITY, AND BONE MINERAL DENSITY IN EARLY RHEUMATOID ARTHRITIS<br>(HAKKINEN Arja (et al))<br>159353 | 2 | 8 | 16 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 020/<br>020 | **ANNALS OF THE RHEUMATIC DISEASES**<br>SUSTAINED MAINTENANCE OF EXERCISE INDUCED MUSCLE STRENGTH GAINS AND NORMAL BONE MINERAL DENSITY IN PATIENTS WITH EARLY RHEUMATOID ARTHRITIS: A 5 YEAR FOLLOW UP<br>(HAKKINEN A. (et al))<br>159354 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 021/<br>021 | **BRITISH MEDICAL JOURNAL**<br>CONTROLLED TRIAL OF PHYSIOTHERAPY IN ARTHRITIS<br>(HAMILTON D.E., BYWATERS E.G.L., PLEASE N.W.)<br>159355 | 2 | 3 | 6 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 022/<br>022 | **SCANDINAVIAN JOURNAL OF RHEUMATOLOGY**<br>LONGTERM PHYSICAL TRAINING IN RHEUMATOID ARTHRITIS. A RANDOMIZED TRIAL WITH DIFFERENT TRAINING PROGRAMS AND BLINDED OBSERVERS.<br>(HANSEN T.M. (et al))<br>159356 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 023/<br>023 | **PHYSIOTHERAPY THEORY AND PRACTICE**<br>COMPARISON OF THREE DIFFERENT PHYSIOTHERAPY TREATMENTS FOR RHEUMATOID ARTHRITIS OF THE HANDS<br>(HAWKES J. (et al))<br>159357 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 024/<br>024 | **PHYSIOTHERAPY**<br>DEVELOPMENT AND DELIVERY OF AN EXERCISE INTERVENTION FOR RHEUMATOID ARTHRITIS: STRENGTHENING AND STRETCHING FOR RHEUMATOID ARTHRITIS OF THE HAND (SARAH) TRIAL.<br>(HEINE P.J. (et al))<br>159358 | 2 | 9 | 18 | R1.07 | R42.63* | R6.39 | R49.02 |
| 025/<br>025 | **RHEUMATOLOGY**<br>MAINTENANCE OF PHYSICAL ACTIVITY AFTER INTERNET-BASED PHYSICAL ACTIVITY INTERVENTIONS IN PATIENTS WITH RHEUMATOID ARTHRITIS<br>(HURKMANS E.J. (et al))<br>159359 | 2 | 6 | 12 | R1.07 | R42.63* | R6.39 | R49.02 |
| 026/<br>026 | **PHYSICAL THERAPY**<br>FACILITATION TECHNICS IN THE TREATMENT OF EARLY RHEUMATOID ARTHRITIS*<br>(LONTA Marjorie K.)<br>159360 | 2 | 2 | 4 | R1.07 | R42.63* | R6.39 | R49.02 |
| 027/<br>027 | **INDIAN JOURNAL OF PHYSIOTHERAPY AND OCCUPATIONAL THERAPY**<br>EFFECT OF FREE EXERCISE AND RHEUMATOID ARTHRITIS<br>(KAPALE P.V., VARDHARAJULU G., WARUDE T.)<br>159361 | 2 | 4 | 8 | R1.07 | R42.63* | R6.39 | R49.02 |
| 028/<br>028 | **RHEUMATOLOGY**<br>ACTIVE CONSERVATIVE TREATMENT OF ATLANTOAXIAL SUBLUXATION IN RHEUMATOID ARTHRITIS<br>(KAUPPI M. (et al))<br>159362 | 2 | 4 | 8 | R1.07 | R42.63* | R6.39 | R49.02 |
| 029/<br>029 | **MUSCULOSKELETAL CARE**<br>EFFECTS OF AEROBIC AND RESISTANCE EXERCISE ON MARKERS OF LARGE JOINT HEALTH IN STABLE RHEUMATOID ARTHRITIS PATIENTS: A PILOT STUDY<br>(LAW R.J. (et al))<br>159363 | 2 | 14 | 28 | R1.07 | R42.63* | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 030/<br>030 | **JOURNAL OF MANIPULATIVE AND PHYSIOLOGICAL THERAPEUTICS**<br>JOINT MOBILIZATION OF THE HANDS OF PATIENTS WITH RHEUMATOID ARTHRITIS: RESULTS FROM AN ASSESSOR-BLINDED, RANDOMIZED CROSSOVER TRIAL<br>(LEVITSKY Adrian (et al))<br>159364 | 2 | 13 | 26 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 031/<br>031 | **JOURNAL OF RHEUMATOLOGY, THE**<br>IMPROVEMENTS FOLLOWING SHORT TERM HOME BASED PHYSICAL THERAPY ARE MAINTAINED AT ONE YEAR IN PEOPLE WITH MODERATE TO SEVERE RHEUMATOID ARTHRITIS.<br>(LINEKER S.C. (et al))<br>159365 | 2 | 4 | 8 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 032/<br>032 | **PHYSIOTHERAPY CANADA**<br>INVESTIGATING CARE PROVIDED BY PHYSICAL THERAPISTS TREATING PEOPLE WITH RHEUMATOID ARTHRITIS: PILOT STUDY<br>(LINEKER Sydney C., HURLEY Laurie, WILKINS Annette)<br>159366 | 2 | 8 | 16 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 033/<br>033 | **ARTHRITIS AND RHEMATISM**<br>REHABILITATION OF THE RHEUMATOID CRIPPLE; A FIVE YEAR STUDY.<br>(LOWMAN E.W.)<br>159367 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 034/<br>034 | **ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION**<br>ELDERLY RHEUMATOID ARTHRITIS PATIENTS ON STEROID TREATMENT TOLERATE PHYSICAL TRAINING WITHOUT AN INCREASE IN DISEASE ACTIVITY<br>(LYNGBERG K. (et al))<br>159368 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 035/<br>035 | **RHEUMATOLOGY INTERNATIONAL: CLINICAL AND EXPERIMENTAL INVEST**<br>IMPACT OF EXERCISE ON SLEEP (TIME, QUALITY, AND DISTURBANCE) IN PATIENTS WITH RHEUMATOID ARTHRITIS: A STUDY PROTOCOL FOR A PILOT RANDOMISED CONTROLLED TRIAL<br>(MCKENNA Sean G. (et al))<br>159369 | 2 | 8 | 16 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 036/<br>036 | **ACTIVITIES, ADAPTATION, AND AGING**<br>EFFECT OF AN ARTHRITIS FOUNDATION EXERCISE PROGRAM ON SLEEP QUALITY/SLEEP DISTURBANCE IN SENIORS WITH RHEUMATOID ARTHRITIS: A PILOT STUDY<br>(MCMANUS Kaitlyn M., VISKER Joseph D., COX Carolyn C.)<br>159370 | 2 | 8 | 16 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 037/<br>037 | **PAIN PRACTICE**<br>ENDOGENOUS PAIN MODULATION IN RESPONSE TO EXERCISE IN PATIENTS WITH RHEUMATOID ARTHRITIS, PATIENTS WITH CHRONIC FATIGUE SYNDROME AND COMORBID FIBROMYALGIA, AND HEALTHY CONTROLS: A DOUBLE-BLIND RANDOMIZED CONTROLLED TRIAL<br>(MEEUS Mira (et al))<br>159371 | 2 | 9 | 18 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 038/<br>038 | **RESEARCH IN NURSING AND HEALTH**<br>EFFECTS OF EXERCISE ON FATIGUE, AEROBIC FITNESS, AND DISEASE ACTIVITY MEASURES IN PERSONS WITH RHEUMATOID ARTHRITIS<br>(NEUBERGER Geri B. (et al))<br>159373 | 2 | 10 | 20 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 039/<br>039 | **AMERICAN JOURNAL OF PHYSICAL MEDICINE AND REHABILITATION**<br>EFFECTS OF A MODIFIED DANCE-BASED EXERCISE ON CARDIORESPIRATORY FITNESS, PSYCHOLOGICAL STATE AND HEALTH STATUS OF PERSONS WITH RHEUMATOID ARTHRITIS<br>(NOREAU L. (et al))<br>159374 | 2 | 9 | 18 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 040/<br>040 | **RHEUMATOLOGY**<br>CONSERVATIVE HAND THERAPY TREATMENTS IN RHEUMATOID ARTHRITIS--A RANDOMIZED CONTROLLED TRIAL.<br>(O'BRIEN A.V. (et al))<br>159375 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 041/<br>041 | **JOURNAL OF NURSING SCHOLARSHIP**<br>DECREASING PAIN AND DEPRESSION IN A HEALTH PROMOTION PROGRAM FOR PEOPLE WITH RHEUMATOID ARTHRITIS<br>(HYUNSOO Oh, WHASOOK Seo)<br>159376 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 042/<br>042 | **ARTHRITIS AND RHEMATISM**<br>DANCE-BASED AEROBIC EXERCISE FOR RHEUMATOID ARTHRITIS<br>(PERLMAN Susan G. (et al))<br>159377 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 043/<br>043 | **ARTHRITIS CARE & RESEARCH**<br>NEUROMUSCULAR ELECTRICAL STIMULATION COMPARED TO VOLITIONAL EXERCISE FOR IMPROVING MUSCLE FUNCTION IN RHEUMATOID ARTHRITIS: A RANDOMIZED PILOT STUDY<br>(PIVA Sara R. (et al))<br>159378 | 2 | 10 | 20 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 044/<br>044 | **ANNALS OF THE RHEUMATIC DISEASES**<br>AB1125 A COMPARISON OF THE EFFECTS OF KINESIOTHERAPY AND POST-ISOMETRIC RELAXATION ON KNEE PAIN IN PATIENTS WITH RHEUMATOID ARTHRITIS: TABLE 1.<br>(ROMANOWSKI M. (et al))<br>159380 | 2 | 1 | 2 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 045/<br>045 | **SCANDINAVIAN JOURNAL OF OCCUPATIONAL THERAPY**<br>EFFECT OF AN INTENSIVE HAND EXERCISE PROGRAMME IN PATIENTS WITH RHEUMATOID ARTHRITIS<br>(RØNNINGEN Aud, KJEKEN Ingvild)<br>159381 | 2 | 11 | 22 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 046/<br>046 | **ACR OPEN RHEUMATOLOGY**<br>IMPACT OF EXERCISE, LIFESTYLE, AND CLINICAL FACTORS ON PERCEIVED COGNITIVE FUNCTION IN PATIENTS WITH RHEUMATOID ARTHRITIS: RESULTS FROM A PROSPECTIVE COHORT STUDY<br>(SHADICK Nancy A. (et al))<br>159383 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 047/<br>047 | **INDIAN JOURNAL OF PHYSIOTHERAPY AND OCCUPATIONAL THERAPY**<br>EFFECT OF THERAPEUTIC EXERCISE PROGRAMME IN ADULTS WITH EARLY RHEUMATOID ARTHRITIS<br>(SHINDE Sandeep Babasaheb, VARADHARAJULU G.)<br>159385 | 2 | 5 | 10 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 048/<br>048 | **ANNALS OF THE RHEUMATIC DISEASES**<br>EFFECT OF INTENSIVE EXERCISE ON PATIENTS WITH ACTIVE RHEUMATOID ARTHRITIS: A RANDOMISED CLINICAL TRIAL<br>(VAN DEN ENDE C.H.M. (et al))<br>159386 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 049/<br>049 | **JOURNAL OF RHEUMATOLOGY, THE**<br>RANDOMIZED CONTROLLED TRIAL TO EVALUATE THE EFFECTIVENESS OF AN EXERCISE PROGRAM IN WOMEN WITH RHEUMATOID ARTHRITIS TAKING LOW DOSE PREDNISONE.<br>(WESTBY M.D. (et al))<br>159387 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 050/<br>050 | **ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION**<br>FEASIBILITY AND OUTCOMES OF A HOME-BASED EXERCISE PROGRAM ON IMPROVING BALANCE AND GAIT STABILITY IN WOMEN WITH LOWER-LIMB OSTEOARTHRITIS OR RHEUMATOID ARTHRITIS: A PILOT STUDY.<br>(WILLIAMS S.B. (et al))<br>159388 | 2 | 9 | 18 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 051/<br>051 | **HEALTH TECHNOLOGY ASSESSMENT**<br>STRENGTHENING AND STRETCHING FOR RHEUMATOID ARTHRITIS OF THE HAND (SARAH). A RANDOMISED CONTROLLED TRIAL AND ECONOMIC EVALUATION<br>(WILLIAMS Mark A. (et al))<br>159389 | 2 | 222 | 444 | R1.07 | R475.08 | R71.26 | R546.34 |
| 052/<br>052 | **BMC MUSCULOSKELETAL DISORDERS**<br>PROSPECTIVE PILOT STUDY TO EVALUATE AN ANIMATED HOME-BASED PHYSICAL EXERCISE PROGRAM AS A TREATMENT OPTION FOR PATIENTS WITH RHEUMATOID ARTHRITIS<br>(ZERNICKE Jan (et al))<br>159390 | 2 | 12 | 24 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 053/<br>053 | **BRITISH JOURNAL OF PHYSICAL MEDICINE: INCLUDING ITS APPLICATION TO INDUSTRY**<br>RHEUMATOID ARTHRITIS; PHYSIOTHERAPY AND PSYCHOTHERAPY IN SPA TREATMENT.<br>(ARMOUR G.)<br>159391 | 2 | 3 | 6 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 054/<br>054 | **FLORIDA MEDICAL ASSOCIATION. JOURNAL**<br>MEDICAL TREATMENT OF RHEUMATOID ARTHRITIS<br>(CALDWELL J.R.)<br>159393 | 2 | 5 | 10 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 055/<br>055 | **PHYSIOTHERAPY**<br>PHYSIOTHERAPY IN THE TREATMENT OF RHEUMATOID ARTHRITIS.<br>(CAMERON S.H.)<br>159395 | 2 | 5 | 10 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 056/<br>056 | **LAHEY CLINIC BULLETIN**<br>CONSERVATIVE MANAGEMENT OF RHEUMATOID ARTHRITIS.<br>(FERNANDEZ-HERLIHY L.)<br>159396 | 2 | 6 | 12 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 057/<br>057 | **CLINICAL PEDIATRICS**<br>SYSTEMATIC PHYSICAL MANAGEMENT OF THE PATIENT WITH RHEUMATOID ARTHRITIS.<br>(GRANGER C.V. JR)<br>159398 | 2 | 5 | 10 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 058/<br>058 | **PHYSIOTHERAPY CANADA**<br>EXERCISE TRAINING AND DEXAMETHASONE IONTOPHORESSIS IN RHEUMATOID ARTHRITIS: A CASE STUDY<br>(HASSON Scott (et al))<br>159399 | 2 | 4 | 8 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 059/<br>059 | **JOURNAL OF RHEUMATOLOGY, THE**<br>RANDOMIZED CONTROLLED TRIAL OF HOME EXERCISE ON THE RHEUMATOID HAND.<br>(HOENIG H. (et al))<br>159400 | 2 | 5 | 10 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 060/<br>060 | **AMERICAN JOURNAL OF OCCUPATIONAL THERAPY**<br>EFFECTIVENESS OF HAND EXERCISE AND THERMAL MODALITIES AGENTS IN MANAGING OSTEOARTHRITIS AND RHEUMATOID ARTHRITIS IN HANDS: RANDOMIZED CONTROLLED STUDY<br>(NAFAI Said (et al))<br>159401 | 2 | 1 | 2 | R1.07 | R42.63 * | R6.39 | R49.02 |
| 061/<br>061 | **ARCHIVES OF PHYSICAL MEDICINE AND REHABILITATION**<br>OPTIMUM REST-EXERCISE BALANCE IN THE TREATMENT OF RHEUMATOID ARTHRITIS.<br>(HOLLANDER J.L., PIERSOL G.M.)<br>159402 | 2 | 7 | 14 | R1.07 | R42.63 * | R6.39 | R49.02 |

| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 062/<br>062 | **PHYSIOTHERAPY CANADA**<br>COMPLIANCE OF PATIENTS WITH AN EXERCISE PROGRAM FOR RHEUMATOID ARTHRITIS.<br>(TERPSTRA S.J., DE WITTE L.P., DIEDERIKS J.P.M)<br>159403 | 2 | 5 | 10 | R1.07 | R42.63 * | R6.39 | R49.02 |
| | **TOTALS:** | **124** | **645** | **1,290** | | **R3,075.51** | **R461.05** | **R3,536.56** |

*** MINIMUM FEE PER TRANSACTIONAL CLEARANCE (R42.63 + VAT)**

**PLEASE NOTE: DALRO reserves the right to adjust the number of pages copied from a publication in all quotations if the calculation as per the application is found to be incorrect.**

***APPROVED***
***By Laura O'Connor at 8:06 am, Aug 26, 2020***

## ITEM(S) FOR WHICH REPROGRAPHIC REPRODUCTION RIGHTS CANNOT BE CLEARED THROUGH DALRO

| Item | Title | Reason | Comments |
|---|---|---|---|
| 001/<br>001 | Case Study of Treatment of a Patient after Rheumatoid Arthritis | NO COPYRIGHT VESTS IN WORK | NON-QULIFYING WORK |
| 002/<br>002 | King's College London<br>Exercise-based upper limb rehabilitation in rheumatoid arthritis | NO COPYRIGHT VESTS IN WORK | NON-QULIFYING WORK |
| 003/<br>003 | Revista Terapia Manual<br>Physical therapy intervention in rheumatoid arthritis: case report | DALRO DOES NOT CONTROL COPYRIGHT IN WORK | NO AGREEMENT WITH TERRITORY |

**This Quotation is valid for 90 days.**

DRAMATIC, ARTISTIC and LITERARY
RIGHTS ORGANISATION (Pty) LIMITED

www.dalro.co.za
dalro@dalro.co.za

SAMRO Place
20 De Korte Street
Braamfontein
2001 Johannesburg

P O Box 31627
2017 Braamfontein
Telephone 086 12 DALRO
International +27 (0)11 712 8000
Fax +27 (0)86 648 6299

Affiliated to the International Confederation of Copyright Societies – CISAC
Member of the International Federation of Reproduction Rights Organisations – IFRRO
Reg. No. 1967/005018/07

**TO:** DURBAN UNIVERSITY OF TECHNOLOGY

Intellectual Property
Durban University of Technology
P O Box 1334
DURBAN

Dear Ramika,

**REPROGRAPHIC REPRODUCTION RIGHTS QUOTATION FOR (CHIROPRACTIC/RESEARCH DISSERTATION/KYLE DONALD SCLANDERS/COST CENTRE: G942 ACCOUNT CODE: 201235) (0000010587/01):**

DALRO hereby grants permission for the reprographic reproduction of extract(s) from the publication(s) as per the schedule following. The reproduction royalties are indicated next to the title(s).

**Please note that this permission is subject to the following conditions:**

1. The copied extract(s) may only be distributed by the licensing entity identified on the schedule, and only to persons enrolled for the specific course of study at the licensing entity.
2. Should the licensing entity decide to sell the copied extract(s) to the persons enrolled for the course of study, the selling price per unit shall not exceed the cost per unit incurred by the licensing entity.
3. The copied extract(s) may only be distributed from a point which is under the direct control of, and staffed by, the licensing entity, and no external entity or concern, whether commercial or non-commercial, may act as distributor.

Only upon acceptance, in writing, of this quotation will a Tax Invoice be issued. Please let us know if this reproduction will not be going ahead at all to enable us to cancel the quotation. Should the information on which the quotation is based, have changed, the new information need to be supplied to enable us to prepare a revised quotation.

Yours sincerely

**RIGHTS ADMINISTRATION**

**REPROGRAPHIC REPRODUCTION RIGHTS QUOTATION: 30/10/2020**

**QUOTATION NUMBER: 0000010587/01**

DALRO
DRAMATIC, ARTISTIC and LITERARY RIGHTS ORGANISATION (Pty) LIMITED

| | |
|---|---|
| LICENSING ENTITY: | DURBAN UNIVERSITY OF TECHNOLOGY |
| REFERENCE: | COST CENTRE: G942 ACCOUNT CODE: 201235 |
| LECTURER: | KYLE DONALD SCLANDERS |
| DEPARTMENT: | CHIROPRACTIC |
| COURSE: | RESEARCH DISSERTATION |
| DISSEMINATION: | SINGLE ITEM HANDOUT |
| USAGE PERIOD: | 01/06/2020 - 31/12/2020 |

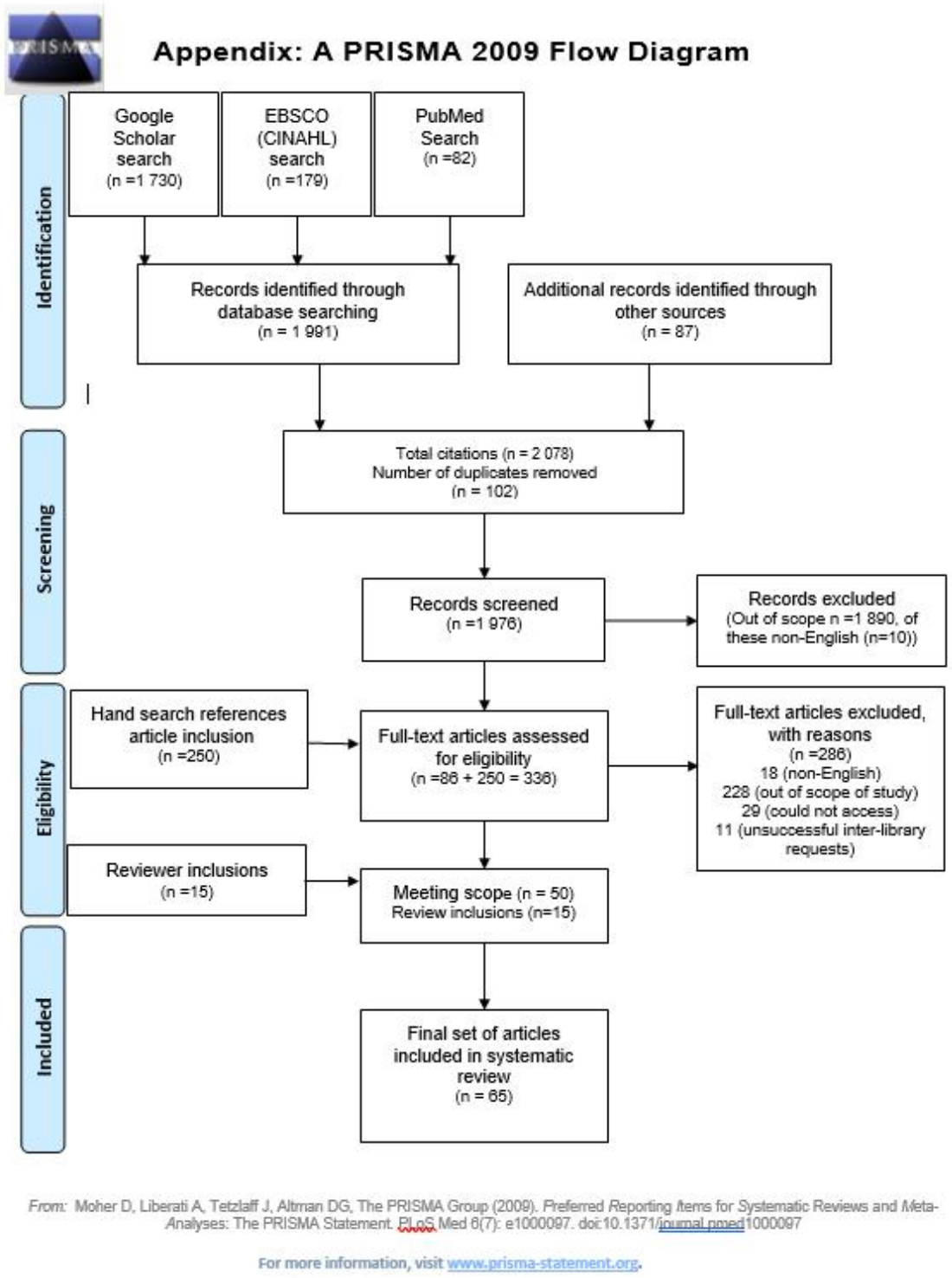
| Item | Title | Units | Pages | Copies | Rate | VAT Excl. | VAT | VAT Incl. |
|---|---|---|---|---|---|---|---|---|
| 001/ 001 | **EVIDENCE-BASED COMPLEMENTARY AND ALTERNATIVE MEDICINE**<br>MANUAL THERAPY (POSTISOMETRIC RELAXATION AND JOINT MOBILIZATION) IN KNEE PAIN AND FUNCTION EXPERIENCED BY PATIENTS WITH RHEUMATOID ARTHRITIS: A RANDOMIZED CLINICAL PILOT STUDY<br>(ROMANOWSKI, M.W (et al.))<br>193225 | 2 | 8 | 16 | R1.07 | R42.63 * | R6.39 | R49.02 |
| | TOTALS: | 2 | 8 | 16 | | R42.63 | R6.39 | R49.02 |

*** MINIMUM FEE PER TRANSACTIONAL CLEARANCE (R42.63 + VAT)**

**PLEASE NOTE: DALRO reserves the right to adjust the number of pages copied from a publication in all quotations if the calculation as per the application is found to be incorrect.**

**This Quotation is valid for 90 days.**

## Appendix D: PRISMA flow diagram / outline




*From:* Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.* PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

## Appendix E: Master list

**Master List:**

| | |
|---|---|
| RCT | 1. Baillet, A., Payraud, E., Niderprim, V.-A., Nissen, M. J., Allenet, B., Francois, P., Grange, L., Casez, P., Juvin, R. and Gaudin, P. 2009. A dynamic exercise programme to improve patients' disability in rheumatoid arthritis: a prospective randomized controlled trial. *Rheumatology*, 48 (4): 410-415.<br>2. Bell, M. J., Lineker, S.C., Wilkins, A.L., Goldsmith, C.H. and Badley E.M,. 1998. A randomized controlled trial to evaluate the efficacy of community based physical therapy in the treatment of people with rheumatoid arthritis.<br>3. Breedland, I., van Scheppingen, C., Leijsma, M., Verheij-Jansen, N. P. and van Weert, E. 2011. Effects of a group-based exercise and educational program on physical performance and disease self-management in rheumatoid arthritis: a randomized controlled study. *Physical therapy*, 91 (6): 879-893.<br>4. Brodin, N., Eurenius, E., Jensen, I., Nisell, R., Opava, C. H. and Group, P. S. 2008. Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. *Arthritis Care & Research*, 59 (3): 325-331.<br>5. Buljina, A. I., Taljanovic, M. S., Avdic, D. M. and Hunter, T. B. 2001. Physical and exercise therapy for treatment of the rheumatoid hand. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 45 (4): 392-397.<br>6. Cima, S. R., Barone, A., Porto, J. M. and de Abreu, D. C. C. 2013. Strengthening exercises to improve hand strength and functionality in rheumatoid arthritis with hand deformities: a randomized, controlled trial. *Rheumatology international*, 33 (3): 725-732.<br>7. de Jong, Z., Munneke, M., Lems, W. F., Zwinderman, A. H., Kroon, H. M., Pauwels, E. K., Jansen, A., Ronday, K. H., Dijkmans, B. A. and Breedveld, F. C. 2004. Slowing of bone loss in patients with rheumatoid arthritis by long-term high-intensity exercise: results of a randomized, controlled trial. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 50 (4): 1066-1076.<br>8. de Jong, Z., Munneke, M., Zwinderman, A. H., Kroon, H. M., Jansen, A., Ronday, K. H., van Schaardenburg, D., Dijkmans, B. A., Van den Ende, C. H. and Breedveld, F. C. 2003. Is a long-term high-intensity exercise program effective and safe in |

patients with rheumatoid arthritis?: Results of a randomized controlled trial. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 48 (9): 2415-2424.

9. Dellhag, B., Wollersjö, I. and Bjelle, A. 1992. Effect of active hand exercise and wax bath treatment in rheumatoid arthritis patients. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 5 (2): 87-92.
10. Dhondt, T. W., L. A Verbruggen, RA B Oostendorp, W Duquet, W. 1999. Pain threshold in patients with rheumatoid arthritis and effect of manual oscillations. *Scandinavian journal of rheumatology*, 28 (2): 88-93.
11. Durcan, L., Wilson, F. and Cunnane, G. 2014. The effect of exercise on sleep and fatigue in rheumatoid arthritis: a randomized controlled study. *J Rheumatol*, 41 (10): 1966-1973.
12. Eversden, L., Maggs, F., Nightingale, P. and Jobanputra, P. 2007. A pragmatic randomised controlled trial of hydrotherapy and land exercises on overall well being and quality of life in rheumatoid arthritis. *BMC musculoskeletal disorders*, 8 (1): 1-7.
13. Garner, S., Fenton, T., Martin, L., Creaser, C., Johns, C. and Barnabe, C. 2018. Personalized diet and exercise recommendations in early rheumatoid arthritis: A feasibility trial. *Musculoskeletal care*, 16 (1): 167-172.
14. Häkkinen, A., Mälkiä, E., Häkkinen, K., Jäppinen, I., Laitinen, L. and Hannonen, P. 1997. Effects of detraining subsequent to strength training on neuromuscular function in patients with inflammatory arthritis. *British journal of rheumatology*, 36 (10): 1075-1081.
15. Häkkinen, A., Sokka, T., Kotaniemi, A. and Hannonen, P. 2001. A randomized two-year study of the effects of dynamic strength training on muscle strength, disease activity, functional capacity, and bone mineral density in early rheumatoid arthritis. *Arthritis & Rheumatism*, 44 (3): 515-522.
16. Häkkinen, A., Sokka, T., Kautiainen, H., Kotaniemi, A. and Hannonen, P. 2004. Sustained maintenance of exercise induced muscle strength gains and normal bone mineral density in patients with early rheumatoid arthritis: a 5 year follow up. *Annals of the rheumatic diseases*, 63 (8): 910-916.
17. Hamilton, D., Bywaters, E. and Please, N. 1959. A controlled trial of various forms of physiotherapy in arthritis. *British medical journal*, 1 (5121): 542.

18. Hansen, T., Hansen, G., Langgaard, A. and Rasmussen, J. 1993. Longterm physical training in rheumatoid arthritis. A randomized trial with different training programs and blinded observers. *Scandinavian journal of rheumatology*, 22 (3): 107-112.
19. Harkcom, T. M., Lampman, R. M., Banwell, B. F. and Castor, C. W. 1985. Therapeutic value of graded aerobic exercise training in rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 28 (1): 32-39.
20. Hoenig, H., Groff, G., Pratt, K., Goldberg, E. and Franck, W. 1993. A randomized controlled trial of home exercise on the rheumatoid hand. *J Rheumatol*, 20 (5): 785-789.
21. Levitsky, A., Kisten, Y., Lind, S., Nordström, P., Hultholm, H., Lyander, J., Hammelin, V., Gentline, C., Giannakou, I. and Faustini, F. 2019. Joint mobilization of the hands of patients with rheumatoid arthritis: results from an assessor-blinded, randomized crossover trial. *Journal of manipulative and physiological therapeutics*, 42 (1): 34-46.
22. Lineker, S. C., Bell, M. J., Wilkins, A. L. and Badley, E. M. 2001. Improvements following short term home based physical therapy are maintained at one year in people with moderate to severe rheumatoid arthritis. *J Rheumatol*, 28 (1): 165-168.
23. Lyngberg, K. K., Harreby, M., Bentzen, H., Frost, B. and Danneskiold-Samsøe, B. 1994. Elderly rheumatoid arthritis patients on steroid treatment tolerate physical training without an increase in disease activity. *Archives of physical medicine and rehabilitation*, 75 (11): 1189-1195.
24. Mannerkorpi, K. and Bjelle, A. 1994. Evaluation of a home training programme to improve shoulder function in rheumatoid arthritis patients. *Physiotherapy Theory and Practice*, 10 (2): 69-76.
25. McManus, K. M., Visker, J. D. and Cox, C. C. 2015. Effect of an arthritis foundation exercise program on sleep quality/sleep disturbance in seniors with rheumatoid arthritis: A pilot study. *Activities, Adaptation & Aging*, 39 (1): 56-63.
26. McMeeken, J., Stillman, B., Story, I., Kent, P. and Smith, J. 1999. The effects of knee extensor and flexor muscle training on the timed-up-and-go test in individuals with rheumatoid arthritis. *Physiotherapy research international*, 4 (1): 55-67.
27. Meeus, M., Hermans, L., Ickmans, K., Struyf, F., Van Cauwenbergh, D., Bronckaerts, L., De Clerck, L. S., Moorken, G., Hans, G. and Grosemans, S. 2015. Endogenous pain modulation in response to exercise in patients with rheumatoid arthritis, patients with chronic fatigue syndrome and comorbid fibromyalgia, and healthy controls: a double-blind randomized controlled trial. *Pain Practice*, 15 (2): 98-106.

28. Metin, Z. G. and Ozdemir, L. 2016. The effects of aromatherapy massage and reflexology on pain and fatigue in patients with rheumatoid arthritis: a randomized controlled trial. *Pain Management Nursing*, 17 (2): 140-149.
29. O'Brien, A., Jones, P., Mullis, R., Mulherin, D. and Dziedzic, K. 2006. Conservative hand therapy treatments in rheumatoid arthritis—a randomized controlled trial. *Rheumatology*, 45 (5): 577-583.
30. Piva, S. R., Khoja, S. S., Toledo, F. G., Chester-Wasko, M., Fitzgerald, G. K., Goodpaster, B. H., Smith, C. N. and Delitto, A. 2019. Neuromuscular electrical stimulation compared to volitional exercise for improving muscle function in rheumatoid arthritis: A randomized pilot study. *Arthritis Care & Research*, 71 (3): 352-361.
31. Romanowski, M. W., Špiritović, M., Romanowski, W. and Straburzyńska-Lupa, A. 2020. Manual Therapy (Postisometric Relaxation and Joint Mobilization) in Knee Pain and Function Experienced by Patients with Rheumatoid Arthritis: A Randomized Clinical Pilot Study. *Evidence-Based Complementary and Alternative Medicine*, 2020.
32. Scholten, C., Brodowicz, T., Graninger, W., Gardavsky, I., Pils, K., Pesau, B., Eggl-Tyl, E., Wanivenhaus, A. and Zielinski, C. C. 1999. Persistent functional and social benefit 5 years after a multidisciplinary arthritis training program. *Archives of physical medicine and rehabilitation*, 80 (10): 1282-1287.
33. Shinde, S. B. and Varadharajulu, G. 2017. Effect of therapeutic exercise programme in adults with early rheumatoid arthritis. *Indian J Physiother Occup Ther*, 11: 76-80.
34. Strasser, B., Leeb, G., Strehblow, C., Schobersberger, W., Haber, P. and Cauza, E. 2011. The effects of strength and endurance training in patients with rheumatoid arthritis. *Clinical rheumatology*, 30 (5): 623-632.
35. Van den Ende, C., Breedveld, F., Le Cessie, S., Dijkmans, B., De Mug, A. and Hazes, J. 2000. Effect of intensive exercise on patients with active rheumatoid arthritis: a randomised clinical trial. *Annals of the rheumatic diseases*, 59 (8): 615-621.
36. Van den Ende, C., Hazes, J., le Cessie, S., Mulder, W. J., Belfor, D. G., Breedveld, F. C. and Dijkmans, B. 1996. Comparison of high and low intensity training in well controlled rheumatoid arthritis. Results of a randomised clinical trial. *Annals of the rheumatic diseases*, 55 (11): 798-805.
37. Vljeland, T. V., Zwinderman, A., Vandenbroucke, J., Breedveld, F. and Hazes, J. 1996. A randomized clinical trial of in-patient multidisciplinary treatment versus routine out-patient care in active rheumatoid arthritis. *Rheumatology*, 35 (5): 475-482.

| | |
|---|---|
| | 38. Westby, M. D., Wade, J. P., Rangno, K. K. and Berkowitz, J. 2000. A randomized controlled trial to evaluate the effectiveness of an exercise program in women with rheumatoid arthritis taking low dose prednisone. *J Rheumatol*, 27 (7): 1674-1680.<br>39. Williams, M. A., Williamson, E. M., Heine, P. J., Nichols, V., Glover, M. J., Dritsaki, M., Adams, J., Dosanjh, S., Underwood, M. and Rahman, A. 2015. Strengthening And stretching for Rheumatoid Arthritis of the Hand (SARAH). A randomised controlled trial and economic evaluation. |
| nRCT | 1. Brorsson, S., Hilliges, M., Sollerman, C. and Nilsdotter, A. 2009. A six-week hand exercise programme improves strength and hand function in patients with rheumatoid arthritis. *Journal of rehabilitation medicine*, 41 (5): 338-342.<br>2. Field, T., Diego, M., Delgado, J., Garcia, D. and Funk, C. 2013. Rheumatoid arthritis in upper limbs benefits from moderate pressure massage therapy. *Complementary therapies in clinical practice*, 19 (2): 101-103.<br>3. Gizińska, M., Rutkowski, R., Romanowski, W., Lewandowski, J. and Straburzyńska-Lupa, A. 2015. Effects of whole-body cryotherapy in comparison with other physical modalities used with kinesitherapy in rheumatoid arthritis. *BioMed research international*, 2015.<br>4. Häkkinen, A., Hannonen, P., Nyman, K., Lyyski, T. and Häkkinen, K. 2003. Effects of concurrent strength and endurance training in women with early or longstanding rheumatoid arthritis: comparison with healthy subjects. *Arthritis Care & Research: Official Journal of the American College of Rheumatology*, 49 (6): 789-797.<br>5. Hawkes, J., Care, G., Dixon, J., Bird, H. and Wright, V. 1986. A comparison of three different physiotherapy treatments for rheumatoid arthritis of the hands. *Physiotherapy Practice*, 2 (4): 155-160.<br>6. Hurkmans, E. J., Van den Berg, M. H., Ronday, K. H., Peeters, A. J., Le Cessie, S. and Vliet Vlieland, T. P. 2010. Maintenance of physical activity after Internet-based physical activity interventions in patients with rheumatoid arthritis. *Rheumatology*, 49 (1): 167-172.<br>7. Karatepe, A. G., Günaydin, R., Türkmen, G. and Kaya, T. 2011. Effects of home-based exercise program on the functional status and the quality of life in patients with rheumatoid arthritis: 1-year follow-up study. *Rheumatology international*, 31 (2): 171-176.<br>8. Law, R. J., Saynor, Z. L., Gabbitas, J., Jones, J., Kraus, A., Breslin, A., Maddison, P. J. and Thom, J. M. 2015. The effects of aerobic and resistance exercise on markers of large joint health in stable rheumatoid arthritis patients: a pilot study. *Musculoskeletal care*, 13 (4): 222-235. |

| | |
|---|---|
| | 9. Lowman, E. W. 1958. Rehabilitation of the rheumatoid cripple: a five year study. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 1 (1): 38-43.<br>10. Noreau, L., Martineau, H., Roy, L. and Belzile, M. 1995. Effects of a modified dance-based exercise on cardiorespiratory fitness, psychological state and health status of persons with rheumatoid arthritis. *American journal of physical medicine & rehabilitation*, 74 (1): 19-27.<br>11. Oh, H. and Seo, W. 2003. Decreasing pain and depression in a health promotion program for people with rheumatoid arthritis. *Journal of Nursing Scholarship*, 35 (2): 127-132.<br>12. Romanowski, M., Straburzyńska-Lupa, A., Romanowska, A. and Lorenc, R. 2014. AB1125 A Comparison of the Effects of Kinesiotherapy and Post-Isometric Relaxation on Knee Pain in Patients with Rheumatoid Arthritis. *Annals of the rheumatic diseases*, 73 (Suppl 2): 1174-1174.<br>13. Rønningen, A. and Kjeken, I. 2008. Effect of an intensive hand exercise programme in patients with rheumatoid arthritis. *Scandinavian journal of occupational therapy*, 15 (3): 173-183.<br>14. Spiegel, J. S., Spiegel, T. M., Ward, N. B., Paulus, H. E., Leake, B. and Kane, R. L. 1986. Rehabilitation for rheumatoid arthritis patients. A controlled trial. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 29 (5): 628-637.<br>15. Zernicke, J., Kedor, C., Müller, A., Burmester, G.-R., Reißhauer, A. and Feist, E. 2016. A prospective pilot study to evaluate an animated home-based physical exercise program as a treatment option for patients with rheumatoid arthritis. *BMC musculoskeletal disorders*, 17 (1): 1-12. |
| Case study/review | 1. Chui, D., Lau, J. and Yau, I. 2004. An outcome evaluation study of the rheumatoid arthritis self-management programme in Hong Kong. *Psychology, health & medicine*, 9 (3): 286-292.<br>2. Chung, C. L. and Mior, S. A. 2015. Use of spinal manipulation in a rheumatoid patient presenting with acute thoracic pain: a case report. *The Journal of the Canadian Chiropractic Association*, 59 (2): 143.<br>3. Cubick, E. E., Quezada, V. Y., Schumer, A. D. and Davis, C. M. 2011. Sustained release myofascial release as treatment for a patient with complications of rheumatoid arthritis and collagenous colitis: a case report. *International journal of therapeutic massage & bodywork*, 4 (3): 1. |

| | |
|---|---|
| | 4. Kapale, P., Vardharajulu, G. and Warude, T. 2017. Effect of Free Exercise and Rheumatoid Arthritis. *Website: www. ijpot. com*, 11 (3): 62.<br>5. Kauppi, M., Leppänen, L., Heikkilä, S., Lahtinen, T. and Kautiainen, H. 1998. Active conservative treatment of atlantoaxial subluxation in rheumatoid arthritis. *British journal of rheumatology*, 37 (4): 417-420.<br>6. Lonta, M. K. 1960. Facilitation Technics in the Treatment of Early Rheumatoid Arthritis. *Physical therapy*, 40 (2): 119-120.<br>7. Neuberger, G. B., Press, A. N., Lindsley, H. B., Hinton, R., Cagle, P. E., Carlson, K., Scott, S., Dahl, J. and Kramer, B. 1997. Effects of exercise on fatigue, aerobic fitness, and disease activity measures in persons with rheumatoid arthritis. *Research in nursing & health*, 20 (3): 195-204.<br>8. Perlman, S. G., Connell, K. J., Clark, A., Robinson, M. S., Conlon, P., Gecht, M., Caldron, P. and Sinacore, J. M. 1990. Dance-based aerobic exercise for rheumatoid arthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*, 3 (1): 29-35.<br>9. Williams, S. B., Brand, C. A., Hill, K. D., Hunt, S. B. and Moran, H. 2010. Feasibility and outcomes of a home-based exercise program on improving balance and gait stability in women with lower-limb osteoarthritis or rheumatoid arthritis: a pilot study. *Archives of physical medicine and rehabilitation*, 91 (1): 106-114. |
| Grey literature | 1. Almuhaysin, A. A. 2019. Case Study of Treatment of a Patient after Rheumatoid Arthritis.<br>2. Manning, V. 2013. Exercise-based upper limb rehabilitation in rheumatoid arthritis. King's College London (University of London). |

## Appendix F: Memorandum of Agreement (MoA)

**Appendix D: Memorandum of Agreement**

Dear Prospective Reviewer,

Title of the study: A systematic review of non-invasive manual therapy in the management of rheumatoid arthritis patients.

| | | | |
|---|---|---|---|
| Principle Investigator: | Mr K D Sclanders | BSc (Stell), MTech (student intern – DUT) | (Researcher) |
| Co-investigator: | Dr. C. Korporaal | M.Tech:Chiropractic | (Supervisor) |

Brief introduction and description of the study:

This study is a systematic review of literature pertaining to the use of Manual therapy (including manipulation, mobilization, massage and exercise) in the management of rheumatoid arthritis patients. Articles were collected in electronic form, from online databases by the researcher. The articles included in this study are divided into categories based on study design (randomised controlled studies (RCT), non-randomised controlled clinical trials (nRCT); case series/reports (CS); observational studies (OS)). Each category of articles are to be reviewed by a panel of external reviewers using rating scales (specific to the study types) and feedback from reviewers will be collated and presented in the study's Chapter 4 (data collection and analysis), from which the Chapter 5 (discussion) will be derived.

Outline of procedures:

Should you agree to participate, you will receive a set of (7 to 10) articles according to their allocated group, the articles as well as the corresponding scale rating sheet (RCT's using the PEDro Scale, nRCT's using the Liddle scale, CS and OS's using the Newcastle-Ottawa Scale), an explanation sheet explaining the use of each scale, and an explanation of what is expected of each reviewer (including timelines and submission of review protocol). You will then be asked to rate the articles according to its corresponding scale, and submit their findings (in the given period) to the co-investigator / supervisor (Dr C Korporaal). Once all the rating sheets from all the reviewers and principle investigator are collected, collation of the data will follow.

Reviewer Requirement:

You are herewith asked to submit a bio-sketch and CV outlining your qualifications and experience (clinical and research based), as an indication that you agree to participate in this study. The bio-sketch and CV are required for record keeping in this study to stand as evidence supporting that the proposal guidelines have been followed. You are also required to have read and signed this Memorandum.

Benefits and Remuneration:

The intention of publication of this study, will offer your participation as a co-author.

Should this study be published, you will be given the option (in writing) to co-author the article. Should you choose not to be included in the publication, you will need to stipulate in writing that you wish to be exempt from this, with

date and signature. An honorarium of one thousand rand (R 1000.00) will be awarded to each reviewer on completion of the data collection as a gesture of appreciation for your time and dedication to this project.

Contact persons:

Please do not hesitate to contact the supervisor/Co-investigator (Dr C Korporaal) and/or researcher (Kyle D Sclanders) regarding any questions or queries via the following methods:


| Dr. Charmaine Korporaal: | | Mr Kyle D Sclanders: | |
|---|---|---|---|
| Telephone: | (+27)31 373 2611 | Telephone: | (+27) 31 373 2205 |
| Cell no: | (+27)83 246 3562 | Cell no: | (+27) 72 593 9322 |
| Email: | charmak@dut.ac.za | Email: | 21516881@dut4life.ac.za / kylesclanders@yahoo.com |


Please note:

Prior to the data collection phase and on conclusion of the data collection phase there is no preference in line of communication. However, to limit bias during the data collection phase communication would be preferred via the co-investigator/supervisor (Dr C Korporaal).

Statement of Agreement to Participate in the Research Study:

I …… (Reviewer's full name),

…… (Identity number / passport number – please note that this information is only for purposes of auditing should be research be subject to a random audit),

have read this document in its entirety and understand its contents. Where I have had any questions or queries, these have been explained to me by .....to my satisfaction. Furthermore, I voluntarily agree to participate in this study as a reviewer.

Reviewer's name.....

Reviewer's Signature: …… Date: ……

Supervisor name.....

Supervisor Signature: …… Date: ……

Researcher name.....

Researcher Signature: …… Date: ……

*(Copies of the signed documents will also be sent to the reviewers for reference).*

## Appendix G: PEDro scale template

### PEDro scale

| | |
|---|---|
| 1. eligibility criteria were specified | no ❑ yes ❑ where: |
| 2. subjects were randomly allocated to groups (in a crossover study, subjects were randomly allocated an order in which treatments were received) | no ❑ yes ❑ where: |
| 3. allocation was concealed | no ❑ yes ❑ where: |
| 4. the groups were similar at baseline regarding the most important prognostic indicators | no ❑ yes ❑ where: |
| 5. there was blinding of all subjects | no ❑ yes ❑ where: |
| 6. there was blinding of all therapists who administered the therapy | no ❑ yes ❑ where: |
| 7. there was blinding of all assessors who measured at least one key outcome | no ❑ yes ❑ where: |
| 8. measures of at least one key outcome were obtained from more than 85% of the subjects initially allocated to groups | no ❑ yes ❑ where: |
| 9. all subjects for whom outcome measures were available received the treatment or control condition as allocated or, where this was not the case, data for at least one key outcome was analysed by "intention to treat" | no ❑ yes ❑ where: |
| 10. the results of between-group statistical comparisons are reported for at least one key outcome | no ❑ yes ❑ where: |
| 11. the study provides both point measures and measures of variability for at least one key outcome | no ❑ yes ❑ where: |

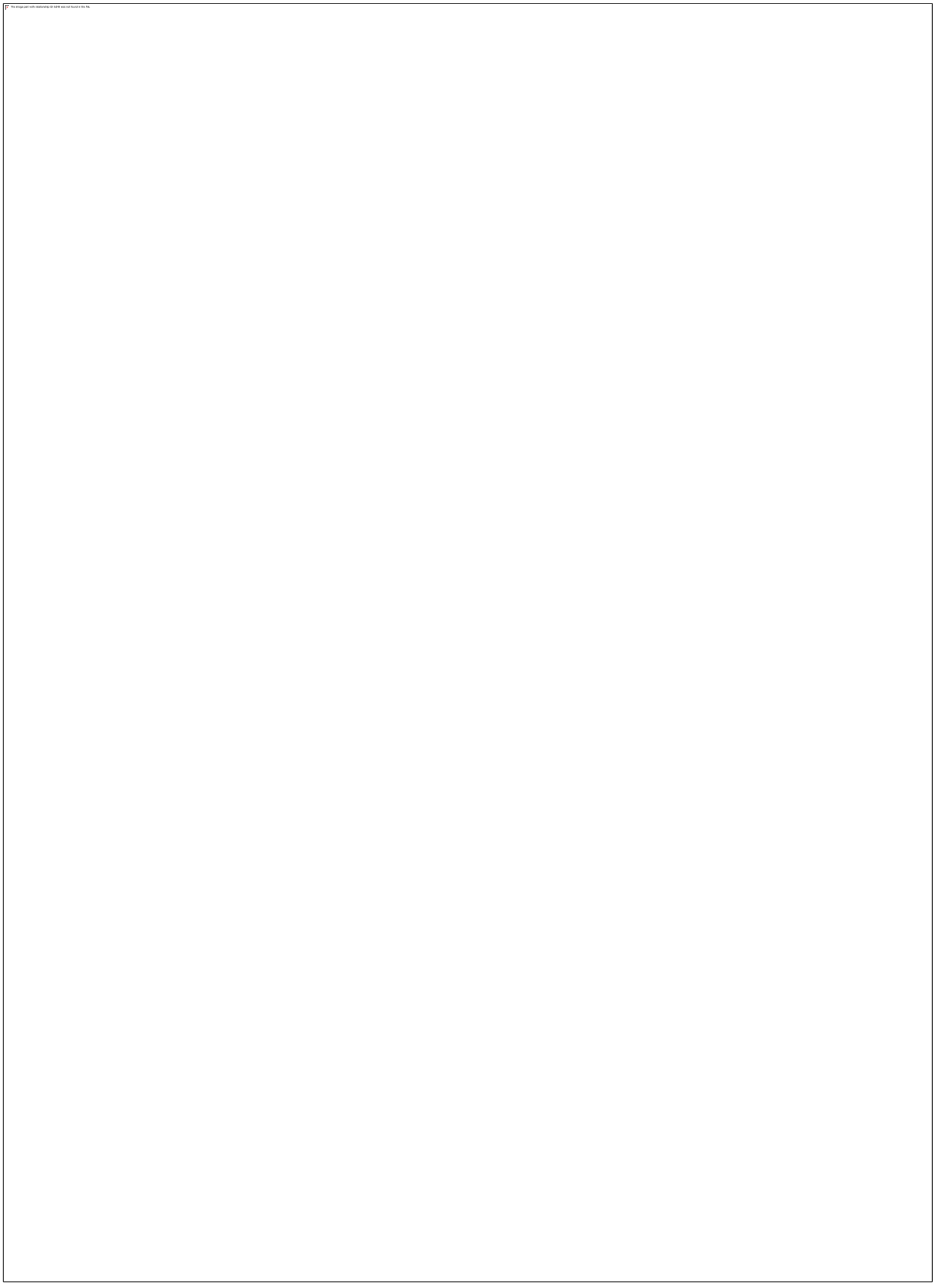
The PEDro scale is based on the Delphi list developed by Verhagen and colleagues at the Department of Epidemiology, University of Maastricht (*Verhagen AP et al (1998). The Delphi list: a criteria list for quality assessment of randomised clinical trials for conducting systematic reviews developed by Delphi consensus. Journal of Clinical Epidemiology, 51(12):1235-41*). The list is based on "expert consensus" not, for the most part, on empirical data. Two additional items not on the Delphi list (PEDro scale items 8 and 10) have been included in the PEDro scale. As more empirical data comes to hand it may become possible to "weight" scale items so that the PEDro score reflects the importance of individual scale items.

The purpose of the PEDro scale is to help the users of the PEDro database rapidly identify which of the known or suspected randomised clinical trials (ie RCTs or CCTs) archived on the PEDro database are likely to be internally valid (criteria 2-9), and could have sufficient statistical information to make their results interpretable (criteria 10-11). An additional criterion (criterion 1) that relates to the external validity (or "generalisability" or "applicability" of the trial) has been retained so that the Delphi list is complete, but this criterion will not be used to calculate the PEDro score reported on the PEDro web site.

The PEDro scale should not be used as a measure of the "validity" of a study's conclusions. In particular, we caution users of the PEDro scale that studies which show significant treatment effects and which score highly on the PEDro scale do not necessarily provide evidence that the treatment is clinically useful. Additional considerations include whether the treatment effect was big enough to be clinically worthwhile, whether the positive effects of the treatment outweigh its negative effects, and the cost-effectiveness of the treatment. The scale should not be used to compare the "quality" of trials performed in different areas of therapy, primarily because it is not possible to satisfy all scale items in some areas of physiotherapy practice.

Last amended June 21st, 1999

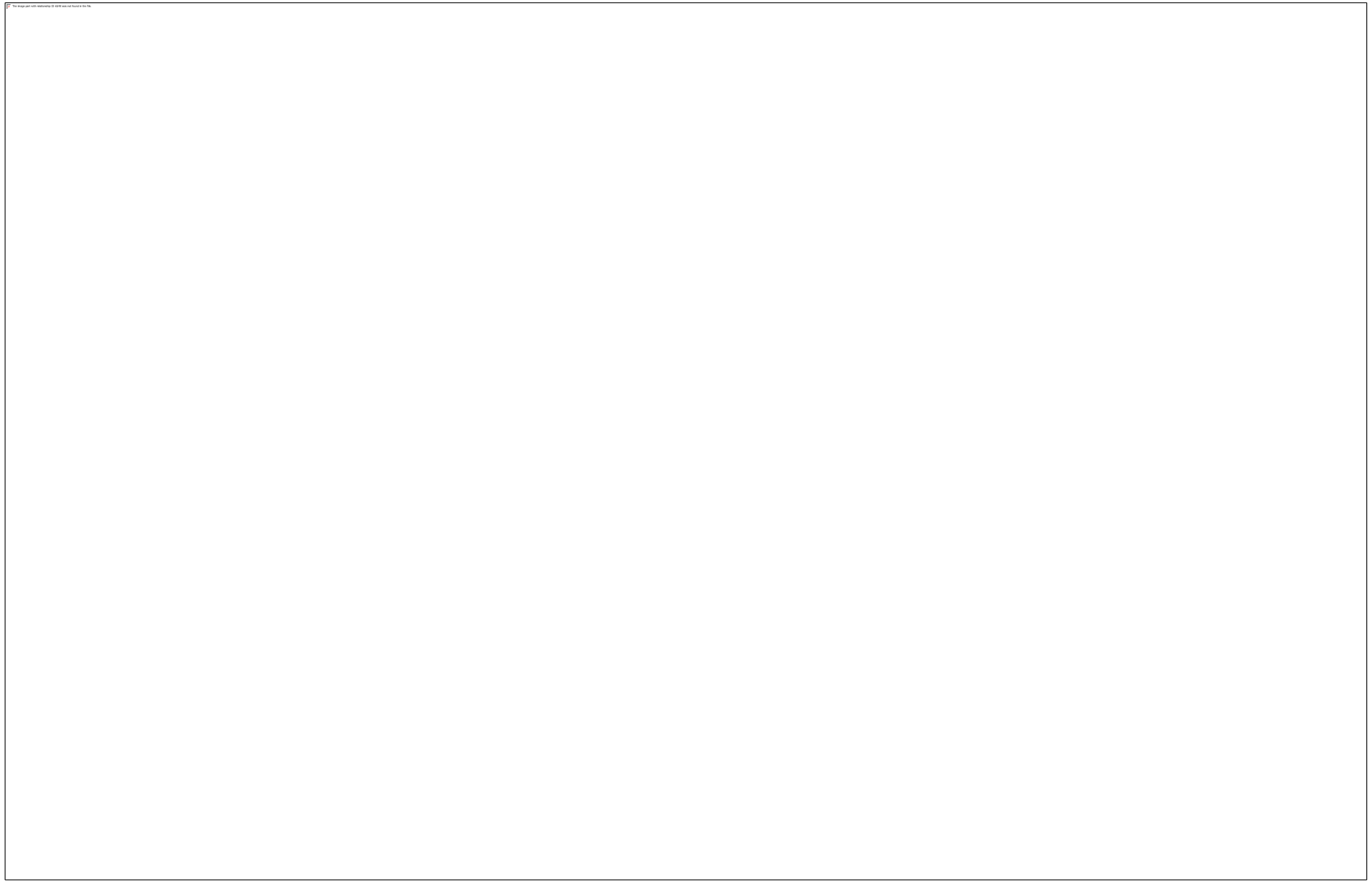

The image part with relationship ID rId4 was not found in the file.


The image part with relationship ID rId4 was not found in the file.

## Appendix H: NOS scale template


The image part with relationship ID rId4 was not found in the file.

## Appendix I: Liddle scale template




The image part with relationship ID rId4 was not found in the file.


The image part with relationship ID rId4 was not found in the file.

## Appendix J: Permission letter from Mrs A Finlayson

## Appendix K: PRISMA checklist

| Section and Topic | Item # | Checklist item | Location where item is reported |
|---|---|---|---|
| **TITLE** | | | |
| Title | 1 | Identify the report as a systematic review. | Front page. Title |
| **ABSTRACT** | | | |
| Abstract | 2 | See the PRISMA 2020 for Abstracts checklist. | Page iv |
| **INTRODUCTION** | | | |
| Rationale | 3 | Describe the rationale for the review in the context of existing knowledge. | Page 5 |
| Objectives | 4 | Provide an explicit statement of the objective(s) or question(s) the review addresses. | Page 5 |
| **METHODS** | | | |
| Eligibility criteria | 5 | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses. | Section 3.6 page 65 - 68 |
| Information sources | 6 | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | Section 3.6 page 65 – 68 |
| Search strategy | 7 | Present the full search strategies for all databases, registers and websites, including any filters and limits used. | Section 3.6 page 65 – 68 |
| Selection process | 8 | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | Section 3.6 page 65 – 68 |
| Data collection process | 9 | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | Section 3.9 and 3.10 page 71 - 77 |
| Data items | 10a | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | N/A |
| | 10b | List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | N/A |
| Study risk of bias assessment | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | Sections 3.9 |
| Effect measures | 12 | Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results. | Sections 1.2.5, 2.9, and 5.3 |
| Synthesis methods | 13a | Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | Section 3.6 Inclusion criteria |
| | 13b | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions. | Section 3.8 |
| | 13c | Describe any methods used to tabulate or visually display results of individual studies and syntheses. | Sections 3.9, 3.10, 4.2, and 5.2 |
| | 13d | Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used. | Section 3.10.4 and Chapter 5 |

| Section and Topic | Item # | Checklist item | Location where item is reported |
|---|---|---|---|
| | 13e | Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression). | Chapter 5 |
| | 13f | Describe any sensitivity analyses conducted to assess robustness of the synthesized results. | N/A |
| Reporting bias assessment | 14 | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases). | Section 2.9 and Chapter 5. |
| Certainty assessment | 15 | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome. | Section 3.6 page 65 - 68 |
| **RESULTS** | | | |
| Study selection | 16a | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | PRISMA |
| | 16b | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. | PRISMA |
| Study characteristics | 17 | Cite each included study and present its characteristics. | Master list, Chapter 4, and 5 |
| Risk of bias in studies | 18 | Present assessments of risk of bias for each included study. | Section 4.4.1 |
| Results of individual studies | 19 | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots. | See Tables 4._b for each article analysis in Chapter 4. |
| Results of syntheses | 20a | For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies. | See Tables 4._a, and 4._b of Chapter 4, and the analysis in Chapter 5. |
| | 20b | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | N/A |
| | 20c | Present results of all investigations of possible causes of heterogeneity among study results. | Section 4.4.1 and Chapter 5 |
| | 20d | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | N/A |
| Reporting biases | 21 | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | N/A |
| Certainty of evidence | 22 | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed. | N/A |
| **DISCUSSION** | | | |
| Discussion | 23a | Provide a general interpretation of the results in the context of other evidence. | Chapter 4 and 5 |
| | 23b | Discuss any limitations of the evidence included in the review. | Chapter 4 and 5 |
| | 23c | Discuss any limitations of the review processes used. | Section 6.1.3 |

| Section and Topic | Item # | Checklist item | Location where item is reported |
|---|---|---|---|
| | 23d | Discuss implications of the results for practice, policy, and future research. | Chapter 5 and 6 |
| **OTHER INFORMATION** | | | |
| Registration and protocol | 24a | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | Page 64 |
| | 24b | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | DUT repository |
| | 24c | Describe and explain any amendments to information provided at registration or in the protocol. | N/A |
| Support | 25 | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review. | Appendix A |
| Competing interests | 26 | Declare any competing interests of review authors. | N/A |
| Availability of data, code and other materials | 27 | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. | Chapter 3 |

*From:* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

For more information, visit: http://www.prisma-statement.org/