PHYSICAL ACTIVITY, SEDENTARY BEHAVIOUR, AND HEALTH OUTCOMES IN RHEUMATOID ARTHRITIS

By

SOPHIA BRADY

A thesis submitted to

The University of Birmingham

For the degree of

DOCTOR OF PHILOSOPHY

School of Sport, Exercise and Rehabilitation Sciences

College of Life and Environmental Sciences

University of Birmingham

May 2022

ABSTRACT

Rheumatoid Arthritis (RA) is a chronic autoimmune disease characterised by high-grade local and systemic inflammation. People with RA experience a multitude of symptoms, such as chronic pain, fatigue, poor mental health and psychological wellbeing and disability, which impact their overall quality of life. People with RA typically engage in low levels of physical activity (PA) and spend long periods of time in sedentary behaviours (SB). Research suggests that increasing PA and reducing SB may improve outcomes in RA. However, this research is limited by the use of non-validated or reliable measurement methods of PA, SB, and health outcomes. In addition, studies rarely assess the different dimensions and elements of PA and SB, and their relative and independent relationships with health in people with RA.

The overarching aim of this research was therefore to develop the understanding of the role of PA and SB for health in RA, through building on existing research in this domain. Specifically, the aim of this thesis was to contribute novel data examining the links between different dimensions of PA and SB with RA outcomes considered to be important by both patients and health professionals.

Initiatives such as Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT), have created core outcome sets of patient- and clinician-important symptoms experienced by people with RA. This thesis focuses on the links between PA and SB with the following OMERACT outcomes: pain, disease activity, functional ability, fatigue, depression, anxiety, subjective vitality, and quality of life.

First, the quality of the current evidence regarding lifestyle PA and SB interventions in people with RA was explored in a systematic review and meta-analysis (***Chapter 2***). In subsequent methodological chapters (***Chapters 3*** and ***4***), the reliability and validity of quantitative sensory testing (QST) modalities and reliability of different ActiGraph accelerometer model and placement site specific cut-points were investigated. The results of these methodological chapters were to inform the design

and methods of a longitudinal study, to be conducted as ***Chapter 5*** of this thesis. The aim of this study was to explore the relationships between ActiGraph-measured PA and activPAL™-measured SB with OMERACT health outcomes in people with RA.

However, due to the COVID-19 pandemic, the study to be conducted as part of ***Chapter 5*** was unable to proceed. As a consequence, ***Chapter 6*** comprised an online survey investigating the cross-sectional associations between different dimensions of self-reported of PA and SB with OMERACT health outcomes in people with RA during the COVID-19 pandemic.

Overall, thesis findings demonstrated individual links exist between lifestyle PA, non-exercise light intensity PA (LPA), walking, exercise, and sedentary time with core OMERACT patient- and clinician-important outcomes. More specifically, existing lifestyle PA and SB interventions are effective at increasing PA, reducing SB, and improving OMERACT outcomes in people with RA. Furthermore, methodological chapters suggested that QST, the ActiGraph GT9X and activPAL™ are reliable and valid assessments of pain, free-living PA, and SB, respectively. In addition, thesis findings also reported that non-exercise LPA, and walking in particular, demonstrated significant positive associations with OMERACT indicators of mental health and psychological wellbeing during the COVID-19 pandemic. These dimensions of PA should therefore be recommended to people with RA to improve mental health and psychological wellbeing, particularly during future pandemics.

To conclude, this thesis provides novel evidence regarding the complex and distinct relationships between different dimensions and elements of PA and SB with core OMERACT health outcomes in people with RA, particularly during the unique worldwide event of the COVID-19 pandemic.

I would like to dedicate this thesis to Peaches and to my Grandad, the one and only Mick Brady

Acknowledgements

Firstly, I am incredibly grateful to my primary supervisors Dr Sally Fenton and Dr Jet Veldhuijzen van Zanten for their support, insight and guidance throughout the past 3+ years. Your passion for research has been inspiring and contagious. Thank you for always answering my slightly stressed emails, and for your constant support and encouragement. You have both always steered me in the right direction, and I will always be in awe in how you both manage to do it all. Thank you both for giving me this opportunity and for believing in me along the way.

I would also like to thank my secondary supervisors and extended members of the research team, Professor Joan Duda, Professor George Kitas and Professor George Metsios, for their valued advice and input into the design, development and publication of this thesis. And thank you to CMAR for the studentship that allowed me to conduct this thesis.

In addition, thank you to my collaborators and supervisors at the University of Nottingham, Professor David Walsh, Dr Daniel McWilliams and Dr Vasileios Georgopoulos. I greatly appreciate your invaluable guidance and expertise. I must also thank my other collaborators, without whom, the systematic review would not have been possible - Thomas Nightingale, Petros Dinas and Saleh Elmsmari. To every one of my participants, thank you.

My friends have been a constant support throughout this PhD, and I am extremely lucky to be surrounded by such caring people. A big thank you goes to Grace for always being at the end of the phone, no matter what, and no questions asked.

I would like to thank every member of my family. For giving me every opportunity, and encouraging me to push myself beyond what I thought I could achieve. Thank you to my Mum for absolutely everything, and a shout out goes to Ailisha (and Jack) for providing me with endless entertainment. And, of course, thank you to Grandad and Peaches for always being with me. And finally, Luke - you've been there for me every step of the way and I couldn't have done this without you.

List of Papers

Sophia Brady conducted study design, recruitment, data collection, data and statistical analyses and writing comprising the studies as part of this thesis. Dr Sally Fenton, Dr Jet Veldhuijzen van Zanten, Professor Joan Duda, Professor George Kitas and Professor David Walsh advised on study design and statistical analysis, and provided feedback on written work and thesis chapters. Please see declaration sections at the end of each chapter for specific breakdown of author contributorship.

Chapter 6 of this thesis comprises the following paper:

- **Brady, S. M.**, Fenton, S. A. M., Metsios, G. S., Bosworth, A., Duda, J. L., Kitas, G. D., & Veldhuijzen van Zanten, J. J. C. S. (2021). Different types of physical activity are positively associated with indicators of mental health and psychological wellbeing in rheumatoid arthritis during COVID-19. *Rheumatology International*, *41*(2), 335-344. doi:10.1007/s00296-020-04751-w

During the period of PhD study within the School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham, the following paper was also accepted for publication:

- **Brady, S. M.**, Veldhuijzen van Zanten, J. J. C. S., Duda, J. L., Kitas, G. D., & Fenton, S. A. M. (2019). Physical Activity/Inactivity: Objective Measurement of. In M. Gellman (Ed.), *Encyclopedia of Behavioral Medicine* (pp. 1-4). New York, NY: Springer New York

During the period of study at the University of Birmingham the following paper was also published:

- Veldhuijzen van Zanten, J. J. C. S., Fenton, S. A. M., **Brady, S. M**, Metsios, G. S., Duda, J. L., & Kitas, G. D. (2020). Mental health and psychological wellbeing in rheumatoid arthritis during COVID-19 – can physical activity help? *Mediterranean Journal of Rheumatology*, *31*, 284-287. doi:10.31138/mjr.31.3.284

Contents

Tables and Figures

List of Abbreviations

ACPA	Anti-citrullinated protein antibody
ACR	American College of Rheumatology
ANOVA	Analysis of variance
AS	Ankylosing Spondylitis
B	Unstandardised beta coefficient
BGD	Between group difference
CDAI	Clinical disease activity index
CG	Control group
CI	Confidence interval
CINAHL	Cumulative Index to Nursing & Allied Health Literature
CMAR	Centre for Musculoskeletal Ageing Research
COMET	Core Outcome Measures in Effectiveness Trials
COVID-19	SARS-CoV-2, Coronavirus disease
CPM	Conditioned Pain Modulation
cpm	Counts per minute
$CPM^{PPT\text{-}mean}$	Single conditioned PPT measurement (PPT^{Con}) minus the arithmetic mean of all the replicated unconditioned PPT measurements from the study visit (PPT^{mean}) (=PPT^{Con} - PPT^{mean})
CPM^{Unc}	Single conditioned PPT measurement (PPT^{Con}) minus the interim unconditioned PPT measurement (PPT^{Unc}) (=PPT^{Con} – PPT^{Unc})
CRP	C-reactive protein
CVD	Cardiovascular Disease
DAS28	Disease activity score 28
DB	Device-based
DI	During intervention
DLW	Doubly-labelled water
DMARD	Disease-modifying antirheumatic drug
EMBASE	Excerpta Medica database
EQ-5D	EuroQoL 5-Dimensional Descriptive System
ESAI	Exercise Stage Assessment Instrument
ESR	Erythrocyte sedimentation rate
EULAR	European Alliance of Associations for Rheumatology
FA	Functional ability
FU	Follow-up
GRADE	Grading of Recommendations Assessment Development and Evaluation
HADS	Hospital anxiety and depression scale
HAQ	Health Assessment Questionnaire
HAQ-DI	Health Assessment Questionnaire disability index
$Healthy^{forearm}$	Healthy participants with quantitative sensory testing conducted on the forearm
$Healthy^{leg}$	Healthy participants with quantitative sensory testing conducted on the lower leg
HEPA	Health enhancing physical activity
HipGT3X	Hip-worn ActiGraph GT3X
HipGT9X	Hip-worn ActiGraph GT9X
HipGT9X-Sasaki	Hip-worn ActiGraph GT9X with triaxial cut-points applied
HipGT9X-Troiano	Hip-worn ActiGraph GT9X with uniaxial cut-points applied

ICC	Intraclass Correlation Coefficient
IG	Intervention group
IPAQ	International Physical Activity Questionnaire
IQR	Interquartile range
kPa	Kilopascals
LBP	Lower back pain
$LBP^{forearm}$	Lower back pain participants with Quantitative sensory testing conducted on the forearm
LoA	Limits of agreement
LPA	Light intensity physical activity
LTPAI	Leisure time physical activity index
M	Mean
MAF	Multidimensional assessment of fatigue
MD	Mean difference
MET	Metabolic equivalent of task
MFI	Multidimensional fatigue inventory
Min/day	Minutes per day
MPA	Moderate intensity physical activity
MPQ	McGill pain questionnaire
MRC	Medical Research Council
MVPA	Moderate to vigorous physical activity
NHANES	National Health and Nutrition Examination Survey
NIH	National Institute of health
NIH-AARP	National Institutes of Health-American Association of Retired Persons
NR	Not reported
NRAS	National Rheumatoid Arthritis Society
NRS	Numerical Rating Scale
NSAID	Non-steroidal anti-inflammatory drug
OA	Osteoarthritis
OMERACT	Outcome Measures in Rheumatoid Arthritis Clinical Trials
PA	Physical activity
PAS 2.1	Physical activity scale 2.1
PEDro	Physiotherapy Evidence Database
PI	Post-intervention
PPT	Pressure-Pain Threshold
PPT^{Con}	Single conditioned PPT measurement
PPT^{mean}	Arithmetic mean of 3 replicate pressure pain threshold measurements
PPT^{Unc}	Single PPT measure taken immediately before the conditioning stimulus
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	Prospective Register of Systematic Review database
QoL	Quality of life
QST	Quantitative Sensory Testing
RA	Rheumatoid Arthritis
RADAI	Rheumatoid Arthritis Disease Activity Index
RA^{leg}	Rheumatoid arthritis participants with quantitative sensory testing conducted on the lower leg
RCT	Randomised controlled trial
RF	Rheumatoid factor

RHH	Russells Hall Hospital
RoB2	Risk of Bias 2
SB	Sedentary behaviour
SBQ	Sedentary Behaviour Questionnaire
SD	Standard deviation
SF	Short form
SF-36	Short form-36
SMD	Standardised mean difference
SPSS	Statistical Package for the Social Sciences
SQUASH	Short Questionnaire to Assess Health-Enhancing Physical Activity
SR	Self-reported
ST	Sedentary Time
SVS	Subjective vitality scale
TS	Temporal Summation
TS^{WUD}	Temporal summation calculated as a wind-up difference
TS^{WUR}	Temporal summation calculated as a wind-up ratio
UK	United Kingdom
USA	United States of America
VAS	Visual analogue scale
VM	Vector magnitude
VPA	Vigorous intensity physical activity
VVPA	Very vigorous physical activity
WGD	Within group difference
WHO	World Health Organisation
WrGT9X	Wrist-worn ActiGraph GT9X
WrGT9X-Montoye	Wrist-worn ActiGraph GT9X with triaxial cut-points applied
WUD	Wind-up difference
WUR	Wind-up ratio
YPAS	Yale Physical Activity Survey
β	Standardised beta coefficient

CHAPTER 1: GENERAL INTRODUCTION

Rheumatoid Arthritis

Epidemiology

Rheumatoid Arthritis (RA) is an autoimmune disease, characterised by chronic systemic inflammation. This high-grade inflammation can lead to articular manifestations such as painful, stiff, tender and swollen joints and poor functional ability (Smolen et al., 2016; Uhlig et al., 2014). Extra-articular manifestations are highly prevalent in people with RA, including fatigue, poor mental health and psychological wellbeing and cardiovascular disease (CVD) (Katz, 2017b; Matcham et al., 2013; Metsios et al., 2015; Smolen et al., 2016). Disease activity is known to fluctuate in people with RA, with periods of acute disease activity, known as flares, interspersed with low disease activity periods, known as remission. If disease activity is poorly controlled in the long term, this can lead to irreversible structural joint damage and functional disability (Lee & Weinblatt, 2001; Smolen et al., 2016). The societal burden of people with RA is substantial, with healthcare and economic costs coming from reduced work capacity and informal care requirements (Smolen et al., 2016). The United Kingdom (UK) has an ageing population and RA is more common in older adults, so this societal burden of RA may continue to rise. A greater wealth of research is needed to investigate novel approaches to manage and treat this disease (Hsieh et al., 2020).

RA is the most common form of inflammatory arthritis, affecting approximately 0.5−1% of adults worldwide (Almutairi et al., 2021; Uhlig & Kvien, 2005; Uhlig et al., 2014), and 0.84% of UK adults aged >16 years (Arthritis Research UK, 2018; Gulati et al., 2018). RA is usually diagnosed in people aged 30−50 years (Smolen et al., 2016; Uhlig et al., 2014). Females are twice as likely to have RA, and there are sex differences in the reported symptoms. For example, hand and foot joints more frequently affected in women, whilst men experience greater large joint involvement (St. Clair et al., 2004). The exact pathophysiology of RA is unclear, but it is thought that disease development is precluded by a combination of genetic, lifestyle, and environmental factors. For example, a family

history of RA increases the risk of developing the disease by 3−5% (Smolen et al., 2016). In addition, there are environmental factors predisposing people to developing RA, which include infectious agents, and lifestyle risk factors such as smoking (St. Clair et al., 2004).

Classification and Diagnosis

RA can be classified according to the American College of Rheumatology−European Alliance of Associations for Rheumatology (ACR−EULAR) criteria (Aletaha et al., 2010). This criteria requires the presence of synovitis in at least 1 joint, and a score ≥6 (on a scale of 1-10) from 4 domains: 1) number and site of involved joints (score: 0= 1 large joint; 1= 2−10 large joints; 2= 1−3 small joints (with/without large joint involvement); 3= 4−10 small joints (with/without large joint involvement); 5= >10 joints (with at least 1 small joint)), 2) serological abnormality (score: 0= negative rheumatoid factor (RF) and negative anti-citrullinated protein antibody (ACPA); 2= low-positive RF or low-positive ACPA; 3= high-positive RF or high-positive ACPA), 3) elevated acute-phase response (score: 0= normal serum C-reactive protein (CRP) and normal erythrocyte sedimentation rate (ESR); 1= abnormal CRP or abnormal ESR) and 4) duration of symptoms (score: 0= <6 weeks; 1= ≥6 weeks) (Aletaha et al., 2010). The ACR criteria gives 91-94% sensitivity and 89% specificity at classifying RA (Arnett et al., 1988).

In early RA where little or no erosive joint damage is seen, diagnosis involves taking a medical history, physical examination, laboratory investigation and imaging techniques (St. Clair et al., 2004). Physical examination can include checking for joint swelling, tenderness and range of motion. This examination forms part of a clinically validated measure of RA disease activity, the Disease Activity Score 28 (DAS28) (Prevoo et al., 1995). The DAS28 includes physical examination of 28 joints (5 metacarpophalangeal finger joints, the interphalangeal joint of the thumb, and the second through fifth proximal interphalangeal joints, wrist, elbow, shoulder and knee joints) creating a count of the number of swollen and tender joints. In addition, patients rate their global health on a scale of 0 to

100 (whereby 0 = best health imaginable and 100 = worst health imaginable) on a visual analogue scale (VAS) to give an indicator of overall health. Finally, serological measures of inflammation, CRP or ESR levels from blood tests, are also used in the calculation of the DAS28. A DAS28 score of ≤3.2 is classified as low disease activity, 3.2–5.1 is moderate disease activity, and >5.1 classified as high disease activity (Gossec, 2018). The DAS28 is used in routine rheumatology appointments, and compared against ACR-EULAR criteria, to assess therapeutic efficacy and identify if treatment adaptations are needed (Prevoo et al., 1995; Smolen et al., 2016). In addition, imaging methods are increasingly used to identify and diagnose early RA, through ultrasonography, computerised tomography and magnetic resonance imaging (Østergaard et al., 2005).

Clinical Features

The most common clinical features of RA are joint pain, stiffness and swelling, with symptoms generally worse in the mornings. These can impair functional ability and ability to do activities of daily life (St. Clair et al., 2004). Disease progression can be highly variable, with some patients having mild inflammation and slow disease progression, and other patients having rapid disease onset over the course of days or weeks (Lee & Weinblatt, 2001). However, initial onset of RA is normally polyarticular and slow, taking place over weeks or months (Gulati et al., 2018). In early disease, the most frequently affected joints are the metacarpophalangeal, proximal interphalangeal and metatarsophalangeal joints of the fingers and toes, as well as the ankle and wrist joints (St. Clair et al., 2004). As the disease progresses, irreparable joint damage can lead to deformities, disability, impaired functional ability and mobility. Some studies have reported that only 17% of patients are free from disability after 10 years of having RA (Sherrer et al., 1986). Research has shown that disability can develop relatively soon after diagnosis and that initial aggressive treatment is required to combat this (Sherrer et al., 1986).

Comorbidities, such as CVD, osteoporosis, hypertension and pulmonary diseases, are common in people with RA (Gabriel et al., 1999; St. Clair et al., 2004). Many of these comorbidities develop due to extra-articular manifestations (e.g., rheumatoid nodules, myopathy, weight loss and rheumatoid cachexia) which are present in approximately 40% of individuals with RA (Cimmino et al., 2000; Cojocaru et al., 2010). Patients with extra-articular involvement often have more active and severe disease, and an increased risk of mortality (Cojocaru et al., 2010; St. Clair et al., 2004).

Pharmacological Treatment and Management

Pharmacological therapy is the first-line of treatment for the majority of people with RA, normally involving a combination of multiple drugs, to reduce the joint inflammation and slow or halt joint destruction (Smolen et al., 2017). Disease-modifying antirheumatic drugs (DMARDs) aim to suppress inflammation, and reduce disease progression and structural damage (Lee & Kavanaugh, 2003). Glucocorticoids provide control over inflammatory symptoms, and may be able to slow or halt radiographic changes associated with RA progression (Kirwan, 1995; Wassenberg et al., 2005). A combination therapy consisting of DMARDs and glucocorticoids are now commonly used in initial treatment of RA (Landewé et al., 2002), as joint damage can take place in the early stages of the disease, perhaps even before diagnosis (Smolen et al., 2017). Therefore, aggressive initial treatment to control disease activity is pivotal to prevent long term structural damage and maintain low disease activity, or even lead to remission in the long term (Boers et al., 1997; Smolen et al., 2016). Non-steroidal anti-inflammatory drugs (NSAIDs), are also prescribed to provide symptomatic relief through reducing inflammation, pain, and stiffness, and consequently improve function (Guidelines, 2002). However, they cannot slow joint damage or disease progression (Guidelines, 2002; Lee & Kavanaugh, 2003). In advanced stages of RA, biological DMARDs (e.g., infliximab and etanercept-examples of tumour necrosis factor inhibitors) can be prescribed, which target and inhibit specific immune pathways involved in RA (Lee & Kavanaugh, 2003; Smolen et al., 2016; Smolen et al., 2017).

Non-pharmacological Treatment

As RA is a multidimensional disease, affecting multiple aspects of both physical and mental health, non-pharmacological treatments are also recommended for both symptomatic relief and to slow disease progression. These can include self-management, education, and physical activity (PA) interventions. A multitude of research has already been done into these non-pharmacological approaches, with promising results (Christie et al., 2007). This thesis will examine non-pharmacological treatments for RA, with a particular focus on PA and sedentary behaviour (SB) interventions, which will be discussed in more detail later in this chapter.

Patient- and Clinician-important Health Outcomes in Rheumatoid Arthritis

As a consequence of the high disease activity, functional disability and pain associated with RA, many people with RA also experience poor mental health and psychological wellbeing, and high levels of fatigue. These can, in turn, further exacerbate RA symptoms. Together, these outcomes can significantly impact quality of life among people living with RA (Treharne et al., 2007). However, owing to the complex and multifaceted nature of RA, there has been a lack of consensus regarding which outcomes are the most important to assess in research seeking to improve management and treatment of RA, and how to measure such outcomes.

The ACR and EULAR have previously recommended that initiatives be created to develop outcome measures for rheumatic diseases. Since this, core sets of patient- and clinician-important outcomes have been determined by groups of experts, through the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) and Core Outcome Measures in Effectiveness Trials (COMET) (Bartlett et al., 2012; Boers et al., 1994; Van Tuyl & Boers, 2015; Williamson et al., 2017). OMERACT was developed in order to provide a consensus between researchers on optimal outcomes and measurements used in RA clinical trials. This was due to previous studies using highly varied

outcomes and endpoints, which were difficult to pool in meta-analyses (Boers et al., 2014). Individuals involved in the OMERACT initiative have consulted previous research, literature and expert opinions to derive core outcome sets specifically for people with rheumatic diseases since 1992 (Boers et al., 2014; Van Tuyl & Boers, 2015). OMERACT have consequently developed an index of core outcomes and linked tools to assess these outcomes, in different conditions and types of trials.

Due to being classed as a core outcome by OMERACT, and the most commonly reported symptom in individuals with RA (Heiberg & Kvien, 2002; Pollard et al., 2006), the experience of pain and its management in RA, was identified as a key focus of this thesis. In addition, further highly reported and relevant patient- and clinician-important indicators of disease activity, functional ability, fatigue, mental health and psychological wellbeing (i.e., depression, anxiety, and subjective vitality) and quality of life are investigated within the studies comprising this thesis. These outcomes are described in detail below.

Pain

Pain is associated with and causes the development of many of the other OMERACT patient- and clinician-important outcomes, such as poor psychological wellbeing, fatigue, impaired functional ability, as well as increased healthcare use (Walsh & McWilliams, 2014). Pain is defined as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" (Merskey & Bogduk, 1994). According to one study, arthritis pain is most commonly described as "aching", "throbbing", "tender" and "shooting" (Burckhardt, 1984). Investigation into the complex mechanisms behind the chronic pain prevalent in patients with RA has been carried out to gain understanding of how to target this debilitating symptom. Further understanding of pain mechanisms at play in people with RA, may then assist in the development of future interventions (Rolke et al., 2006a). It is thought that multiple pain processing mechanisms are

responsible for RA-related pain, and it may result from an interplay between central and peripheral sensory pathways and joint pathology (Walsh & McWilliams, 2014). Research has shown that augmented pain processing from both the central nervous system and peripheral nerves is present in patients with RA, which initiate and facilitate RA inflammation (Walsh & McWilliams, 2014).

Moreover, people with RA report experiencing more pain after periods of poor mental health and psychological wellbeing, with a bidirectional relationship observed between pain with depression and anxiety (Walsh & McWilliams, 2014). Large-scale observational studies have demonstrated that pain is associated with perceived health and activity limitation in people with RA (Demmelmaier et al., 2017). Interestingly, many of these associations were independent of disease activity, and in patients with controlled disease activity and inflammation, pain can still remain present (Lee et al., 2011; Pollard et al., 2006).

Measurement

Pain can be quantified in different ways, and the measurement tools used can reflect the presence or absence of different pain processing pathways. However, in order to better understand the relative contribution of the different mechanisms of pain in people with RA, we need to look critically at the different measurement methods available.

In RA clinical research, pain has been most frequently quantified using self-report methods, such as the McGill pain questionnaire (MPQ) (Melzack, 1987) or using a VAS (Scott & Huskisson, 1976), which have shown good reliability and accuracy at assessing pain (Van Lankveld et al., 1992). A VAS simply measures the overall intensity of pain, and is single-item in nature. Whereas, the MPQ, which was designed specifically for adults with chronic pain, including pain due to rheumatic conditions, assesses both the intensity and quality of subjective pain. Therefore, it may give a greater insight into the sensory aspects of pain and its potential effects on quality of life (Van Lankveld et al., 1992). Both

the full 78 item MPQ and short form version are widely used in research, for their ability to quantify affective and sensory aspects of pain (Hawker et al., 2011).

Quantitative Sensory Testing (QST), has been promoted by many researchers as the optimal method to explore the multidimensional aspects of pain in RA. QST can involve a number of different psychophysical tissue-stimulation laboratory-based tests which can quantify the central and peripheral mechanisms and processing pathways involved in pain (Arendt-Nielsen et al., 2015; Pavlakovic & Petzke, 2010). QST modalities can include tests which measure a pain threshold in response to a mechanical or thermal stimulus (known as Pressure-Pain Threshold [PPT] test), measuring a temporal response to repeated stimulation (known as Temporal Summation [TS] test), or measuring pain thresholds with the addition of a pain-inducing conditioning stimulus (known as Conditioned Pain Modulation [CPM]).

Measuring PPT involves increasing pressure applied to a pre-specified body site using an algometer. Once the sensation of pressure becomes painful, participants press a button and the pressure value is recorded (Joharatnam et al., 2015; Walton et al., 2011). TS is assessed by repeated application of a retractable blunt needle to the skin. A single stimulus is applied, followed by 10 repetitive stimuli, and participants rate the pain intensity on a 0 to 10 scale after the single and 10 stimuli (where 0 signifies no pain and 10 signifies the worst pain imaginable). The TS score can be calculated as a difference (between the score of the single stimulus and the average pain experienced during the 10 subsequent stimuli (TS^{WUD})) or as a ratio (as the average pain during the 10 stimuli divided by pain rating of single stimulus (TS^{WUR})) (Nie et al., 2005). Finally, CPM measurement involves PPT assessment whilst concurrent pain is induced by a conditioning stimulus. The conditioning stimulus can be a heat, cold or ischaemic stimulus, to induce pain inhibition (Kennedy et al., 2016). CPM is calculated by subtracting an unconditioned PPT measurement from the conditioned PPT measurement.

QST has been increasingly used for quantifying the multidimensional aspects of pain in recent clinical research among patients suffering from musculoskeletal and neuropathic pain (Moloney et al., 2012; Rolke et al., 2006a), individuals with osteoarthritis (Suokas et al., 2012) and chronic lower back pain (LBP) (Geletka et al., 2012). **Table 1.1** describes the design, methods, findings, and limitations of QST studies conducted in people with RA. In brief, studies have reported lower pain thresholds and increased pain sensitivity (as indicated by lower PPT and TS, respectively) in individuals with RA compared to healthy controls (Ayhan et al., 2014; Gerecz-Simon et al., 1989; Lofgren et al., 2018b; Vladimirova et al., 2015). Joharatnam et al. (2015) reported a more sensitive PPT was significantly associated with higher levels of inflammation, sleep disturbances, higher DAS28 and poorer mental health (Lee et al., 2009). In addition, PPT has been linked to other core OMERACT outcomes (i.e., functional ability, depression, anxiety and fatigue) in people with RA (Pollard et al., 2012), emphasising the link between multiple parameters of pain with other aspects of quality of life.

As well as quantifying pain, the use of different and multiple QST modalities has the ability to measure and reflect the relative contribution of both central and peripheral pain mechanisms involved in pain modulation (Arendt-Nielsen & Yarnitsky, 2009; Courtney et al., 2010; McWilliams & Walsh, 2017). This is a key advantage of QST over many single-item self-report methods. For example, when PPT is conducted on an anatomical site proximal to an inflamed joint, it has the ability to evaluate peripheral pain or sensitisation, and when conducted at a distant site, it assesses central sensitisation (Middlebrook et al., 2020). **Table 1.1** reports the clinical studies examining pain mechanisms in people with RA. In brief, results show that people with RA experience lower pain thresholds and increased sensitivity at anatomical sites distal to inflamed joints (Edwards et al., 2009; Joharatnam et al., 2015). This is evidence of abnormal and augmented central pain processing present in patients with RA (Walsh & McWilliams, 2014). Other QST studies have reported associations between pain threshold and augmented joint tenderness near affected joints, suggesting peripheral sensitisation is also involved in RA pain (**Table 1.1**) (Konttinen et al., 1992).

Although these studies demonstrate evidence of peripheral and central components in RA-related pain, very few of these studies established reliability of QST measures before use in their trials. Other limitations of these studies include small sample sizes of RA participants, and the majority only conducted PPT, with no other modalities to assess other aspects of pain processing (**Table 1.1**).

Before QST modalities can be used in clinical research to investigate RA-related pain, population and assessor-specific reliability of assessments must be confirmed. In order to ensure reliability when performing QST in research and clinical settings, environmental (e.g., room temperature) and methodological (e.g., test protocol and application) factors must be controlled and standardised (Middlebrook et al., 2020). To my knowledge, only 1 study has examined reliability of QST measures in RA participants (**Table 1.1**). Lee et al. (2018) assessed only inter-rater reliability (i.e., reliability of one rater compared to a 'reliable' second rater) of PPT, TS and CPM modalities before conducting these QST assessments in an RA cohort. However, authors did not provide sufficient detail of methods, assessors and participants involved in this reliability sub-study. Future studies assessing both inter-rater and test-retest (i.e., the reliability of one rater performing QST assessments at 2 different time points) reliability of a standardised QST protocol in patients with RA are required. Therefore, ***Chapter 3*** of this thesis will investigate inter-rater and test-retest reliability of PPT, TS and CPM QST modalities in healthy participants and individuals with RA and chronic LBP.

Table 1.1: Studies conducting pressure-pain threshold, temporal summation and/or conditioned pain modulation in people with Rheumatoid Arthritis

Study	*Title*	*Population*	*QST modality and location*	*Results/Findings*	*Limitations*
Lee Y. C., Bingham C.O., Edwards R.R., et al. (2018)	Pain Sensitization is Associated with Disease Activity in Rheumatoid Arthritis Patients: A Cross-Sectional Study (Follow up: Association of Pain Centralization and Patient-Reported Pain in Active Rheumatoid Arthritis)	139 RA	PPT – trapezius muscle TS- dorsal forearm CPM- cold water bath, left trapezius muscle. Sites were distal to articular sites, avoids peripheral sensitisation	• Initial inter-rater reliability- subgroup of 4 assessors • Associations with disease activity, tender joints, global assessment • ICC to assess reproducibility of QST between assessors • ICC ranged from .71 to .9 for PPTs and TS, whereas the ICC for CPM was .45	• Does not specify if participants for reliability testing are healthy or have RA • Single score ICCs used- not average measures so results cannot be extrapolated to other assessors.
Joharatnam N., McWilliams D.F., Wilson D. et al. (2015)	A cross-sectional study of pain sensitivity, disease-activity assessment, mental health, and fibromyalgia status in rheumatoid arthritis	50 stable long standing RA	PPT- knee, tibia and sternum	• More-sensitive (lower) PPTs at sites over or distant from joints were each associated with greater reported pain, higher patient-reported DAS28 components, and poorer mental health • Indicates central sensitisation contributes to pain in RA.	• Participants had high levels of disease activity (DAS28>3.1)- results are not relevant to those with lower disease activity. • No mention of assessor reliability testing done beforehand. • QST only conducted by 1 assessor • Only conducted PPT- this assesses only 1 aspect of pain processing
Konttinen Y.T., Honkanen V.E., Grönblad M., et al. (1992)	The relation of extraarticular tenderness to inflammatory joint disease and personality in patients with rheumatoid arthritis	44 RA	Pain tenderness threshold in 16 fibrositic tender points	• Associations between pain threshold and joint score index-tenderness is augmented near active joints.	• Small sample size • No mention of assessor reliability testing done beforehand. • Only conducted PPT- this assesses only 1 aspect of pain processing
Lee Y.C., Chibnik L.B., Lu B., et al. (2009)	The relationship between disease activity, sleep, psychiatric distress and pain sensitivity in rheumatoid arthritis: a cross-sectional study	59 RA	PPT at joint (wrist), close to joints (thumbnail) and distal (trapezium muscles) sites	• Significant correlations between PPT with tender joints, sleep problems and psychiatric distress • Multivariable models assessed - CRP inversely associated with wrist pain threshold (p= .003)	• Small sample size • No mention of assessor reliability testing done beforehand. • Female participants only • 1 rheumatologist performed all assessments • Only conducted PPT- this assesses only 1 aspect of pain processing

Vladimirova N., Jespersen A., Bartels E.M., et al. (2015)	Pain Sensitisation in Women with Active Rheumatoid Arthritis: A Comparative Cross-Sectional Study	38 active RA 38 healthy controls	Cuff pressure algometry (tourniquet and compressor) on dominant lower leg (pain detection and tolerance, TS)	• RA participants had lower pain detection and tolerance threshold and TS compared to healthy controls. • Indicates the presence of central pain processing in participants with RA.	• Female participants with active RA (DAS28>2.6) only • No mention of assessor reliability testing done beforehand • Only differences (t-tests) assessed • Small RA sample size
Ayhan, F., Gül, S., Uyar, S., et al. (2014)	The Decreased Sensory Thresholds in Rheumatoid Hand: Comparisons with Osteoarthritic and Normal Hands.	72 RA 43 OA 39 controls	Touch pressure thresholds- hand Pinch strength- median nerve in hand	• Participants with RA had greater sensitivity to light touch pressure threshold than OA and controls • Disease duration, DAS28, HAQ-DI deformities and inflammatory biomarkers were predictors of touch pressure threshold in RA	• Female participants only • Only assessed QST on hand anatomical sites • No mention of reliability testing done beforehand
Gerecz-Simon E.M., Tunks E.R., Heale J.A., et al. (1989)	Measurement of pain threshold in patients with rheumatoid arthritis, osteoarthritis, ankylosing spondylitis, and healthy controls	36 RA 36 OA 18 AS 36 Healthy controls	PPT- forehead, lateral aspect of the arm, midpoint of ulna, palm, midpoint of quadricep, midpoint of antero-medial aspect of tibia	• PPT in non-affected body site in RA participants were significantly lower than healthy controls • Females had lower PPT	• No mention of reliability testing done beforehand • Small RA sample size • Only conducted PPT- this assesses only 1 aspect of pain processing
Pollard L.C., Ibrahim F., Choy E.H., et al. (2012)	Pain thresholds in rheumatoid arthritis: the effect of tender point counts and disease duration	105 RA	PPT	• PPT significantly correlated with tender joint count, HAQ-DI, fatigue, depression and anxiety. • Regression analysis showed tender point count and disease duration were predictors of PPT	• Only conducted PPT- this assesses only 1 aspect of pain processing • No mention of reliability testing done beforehand
Löfgren M., Opava C.H., Demmelmaier I., et al. (2018)	Pain sensitivity at rest and during muscle contraction in persons with rheumatoid arthritis: a substudy within the Physical Activity in Rheumatoid Arthritis 2010 study	46 RA 20 Healthy controls	PPT, suprathreshold pressure pain at rest at 6 sites.	• RA participants had lower PPT, greater suprathreshold pressure pain than controls • Increased pain sensitivity in RA participants	• Participants were those participating in an exercise intervention- not generalisable to full RA population • No mention of reliability testing done beforehand

Note: RA= Rheumatoid Arthritis, OA= Osteoarthritis, AS= Ankylosing Spondylitis, PPT= Pressure-pain threshold, TS= Temporal Summation, VAS= Visual Analogue Scale, QST= Quantitative Sensory Testing, CPM= Conditioned Pain Modulation, CRP= C-reactive Protein, ICC= Intraclass Correlation Coefficient, DAS28= Disease Activity Score 28, HAQ-DI= Health Assessment Questionnaire Disability Index (measure of functional ability).

Functional Ability

Poor functional ability is a common symptom in people with RA, and can be caused by joint swelling and pain (reversible features of RA), as well as joint destruction and deformities (irreversible features of RA) (Aletaha et al., 2006). With the prevalence of joint damage and destruction increasing over time, as does the likelihood of functional impairment; whereby people with long-term RA are more likely to have irreversible disabilities (Aletaha et al., 2006). Longitudinal studies have shown that RA disease activity is also a determinant of functional ability (Drossaers-Bakker et al., 1999; Plant et al., 2005; Welsing et al., 2001), and poor function has been found to be an early feature of RA (Plant et al., 2005; Wolfe et al., 1991).

Impaired functional ability is a major factor contributing towards worsened health outcomes (e.g., anxiety, depression, independence and quality of life), leads to the development of comorbidities, and can exacerbate the already substantial societal burden of living with RA (Englbrecht et al., 2013; Smolen et al., 2016). Due to the impact of functional ability on patient-reported disease activity and wellbeing (Englbrecht et al., 2013), it is not surprising that it is deemed a core patient- and clinician-important outcome by OMERACT (Bartlett et al., 2015; Boers et al., 1994; Van Tuyl & Boers, 2015).

Measurement

Functional ability is measured routinely in clinical assessment and is a common outcome measure in observational studies and assessing the effectiveness of RA behavioural interventions. The Stanford Health Assessment Questionnaire disability index (HAQ-DI) is used in clinical practice and research to quantify functional ability (Fries et al., 1980). The HAQ-DI is made up of 8 subscales that reflect a different activity of daily living (i.e., dressing, rising, eating, walking, hygiene, grip, reach and activities). Patients rate their ability to complete a specific activity within these subscales (e.g., “open car doors”) on a scale from 0 − 3 (with 0 = without any difficulty, 1 = with some difficulty, 2 = with much difficulty, 3 = unable to do), and whether they require any devices/tools (e.g., walking stick,

raised toilet seat, stair lift) or help from another person to assist them with completion of these activities. From this, an overall disability index can be calculated as an average of the score from each of the subscales, with a higher score representing higher functional disability (Kirwan & Reeback, 1986). The HAQ-DI is a standardised, reliable and highly validated measure in RA (Fries et al., 1980).

Fatigue

Fatigue is one of the most frequently reported and debilitating symptoms reported by people with RA, with 40-70% of individuals reporting severe fatigue (Hewlett et al., 2010; Wolfe et al., 1996). It is the second most important domain that reflects remission, and significantly impacts wellbeing and quality of life for people with RA (Katz, 2017a). Hence, fatigue is a core patient-important outcome defined by OMERACT (Kirwan et al., 2007). People with RA report their fatigue as overwhelming, debilitating and uncontrollable, highlighting it's severity (Katz, 2017a).

Fatigue occurs in people with RA potentially as a result of the inflammation involved in the disease progress, and studies have demonstrated that fatigue is linked to multiple inflammatory biomarkers (e.g., ESR) and DAS28 (Madsen et al., 2016). A greater level of fatigue is associated with functional disability and pain (Nikolaus et al., 2013). In detail, pain thresholds have been found to predict levels of fatigue in people with RA, independently of disease activity, indicating disease activity may be a mediator in the relationship between pain and fatigue (Madsen et al., 2016; Walsh & McWilliams, 2014). Therefore, targeting pain clinically may have beneficial effects on fatigue (Katz, 2017b). Fatigue also related to psychological and behavioural outcomes in RA, with high levels of fatigue reported to affect mental health, quality of life, ability to do daily tasks, and increase healthcare utilisation (Katz, 2017a; Nikolaus et al., 2013; Repping-Wuts et al., 2008). However, despite the strong links between fatigue and other core OMERACT outcomes in RA, qualitative studies have demonstrated that RA patients rarely report being asked about fatigue by physicians (Ward et al., 2017).

Measurement

In order to develop more effective fatigue management in people with RA, it should be measured as part of clinical assessment. Consequently, in 2006, OMERACT recommended that fatigue should be a core outcome measure in clinical trials of RA treatments (Kirwan et al., 2007). Since then, there has been growing research into fatigue as a core outcome in RA, using a wide range of self-report measures. These include the multidimensional assessment of fatigue (MAF), the multidimensional fatigue inventory (MFI), a VAS, and the vitality subscale of the short form-36 (SF-36) scale (Hewlett et al., 2007; Smets et al., 1995; Tack, 1991; Ware & Sherbourne, 1992; Wolfe et al., 1996).

Depression

Poor mental health and psychological wellbeing is part of the extensive health burden in people with RA. People with RA are more likely to suffer with low self-esteem, anxiety, and depression than the general population (Gettings, 2010). Depression is characterised by persistent sadness and low mood, lack of interest and poor concentration (WHO, 2020b). In RA, depression is prevalent in up to 66% of individuals (Fiest et al., 2017).

Depression has been linked to RA-related inflammation, with recent studies finding that inflammatory mediators in the hypothalamic-pituitary axis may contribute towards the development of depression in people with RA (Margaretten et al., 2011; Nerurkar et al., 2019). It is also hypothesised that RA-related pain and fatigue play a role in the development of depressive symptoms in people with RA (Nerurkar et al., 2019). However, the exact cause of the poor mental health and psychological wellbeing present in many people with RA, is unknown; and consequently it is poorly understood, recognised and rarely treated in people with RA.

Depression in RA is associated with other OMERACT patient- and clinician-important health outcomes (e.g., increased pain and reduced functional ability and quality of life) (Margaretten et al.,

2011). It can also increase the risk of many comorbidities (e.g., CVD (Treharne et al., 2005)) and mortality. In addition, poorer clinical characteristics can contribute towards development of depressive symptoms (Margaretten et al., 2011). Furthermore, in people with RA, depression can independently contribute towards work disability, unemployment and increased healthcare costs. These numerous links highlight the need for further research into the prevalence and predictors of depression, and methods to manage and treat this debilitating symptom in the RA population (Fiest et al., 2017; Margaretten et al., 2011).

Measurement

Depression is rarely measured clinically as part of RA assessment. However, in research, self-report questionnaires, such as the hospital anxiety and depression scale (HADS) or the EuroQoL 5-Dimensional Descriptive System (EQ-5D) (Balestroni & Bertolotti, 2012), have been more frequently used. HADS and EQ-5D scores have been shown to be independently influenced by functional ability (measured using HAQ-DI), general VAS, pain VAS, disease activity (measured by DAS28), disease duration and worklessness (Hattori et al., 2018). These links demonstrate that assessing depression (and anxiety) is essential in people with RA (Hattori et al., 2018). The HADS and EQ-5D are widely validated in individuals with RA (Hurst et al., 1997; Linde et al., 2008; Zigmond & Snaith, 1983), and are being increasingly used as outcome measures in cross-sectional and intervention studies.

Anxiety

Although depression in people with RA has been studied, in comparison, anxiety has been somewhat overlooked (VanDyke et al., 2004). Many studies in people with RA examining anxiety have combined it with depression in outcome measures (Machin et al., 2020). However, anxiety is distinct from depression, and a recent meta-analysis has found that people with RA are at a significantly increased risk of developing anxiety than non-RA controls (Qiu et al., 2019). Further studies have found

anywhere between 20-70% of the RA population can be classified as anxious (Chandarana et al., 1987; Fiest et al., 2017; Ho et al., 2011).

Anxiety is related to other OMERACT patient- and clinician-important outcomes in people with RA, such as depression, pain, functional ability, tender joint counts, quality of life and disease activity (Ho et al., 2011; Matcham et al., 2016; Qiu et al., 2019; VanDyke et al., 2004). Studies in healthy individuals have shown anxiety and depression do not always co-exist (Machin et al., 2020). However, research into anxiety as a distinct outcome from depression is still scarce in individuals with RA, as demonstrated by the heterogeneity between anxiety prevalence reported in different studies.

Measurement

Although anxiety is a core OMERACT patient- and clinician-important outcome, and highly prevalent in people with RA, it is rarely assessed clinically or in research. Anxiety is most commonly measured alongside depression, as an overall indicator of "mental health", using questionnaires such as the HADS or EQ-5D (Balestroni & Bertolotti, 2012; Zigmond & Snaith, 1983). The HADS however can also specifically measure anxiety, as it is able to produce depression and anxiety sub-scores. Therefore, systematic reviews have recommended the use of this questionnaire, so that anxiety to be assessed as a separate outcome from depression in future studies (Machin et al., 2020).

Subjective Vitality

Although anxiety and depression provide a good indicator of negative mental health and psychological wellbeing, positive psychological states can also be explored and measured in RA to give a more comprehensive overview of wellbeing. Quality of life measures can assess positive mental wellbeing, however, these measurement tools focus on more hedonic (i.e., happiness or pleasure) aspects of wellbeing, and are not able to capture the eudaimonic features (i.e., meaning,

energy, spirit and optimal functioning) (Rouse et al., 2015; Ryan & Deci, 2001). In people with RA, eudaimonic wellbeing is particularly important as poor functional ability can be a debilitating symptom, and optimal functioning is a primary aim of RA treatment. Therefore, being able to quantify eudaimonic psychological wellbeing, and explore factors that may influence this outcome, can give a more relevant and all-encompassing view of wellbeing and its determinants in this population.

Subjective vitality is one such positive indicator of eudaimonic wellbeing and psychological functioning, and can be defined as feeling alive, vital and full of energy (Rouse et al., 2015; Ryan & Frederick, 1997). By assessing subjective vitality in people with RA, additional information can be obtained regarding the influence and relative contribution of RA disease on physical and psychological function (Rouse et al., 2015).

Measurement

Whilst hedonic wellbeing can be captured using quality of life measures (see below), the subjective vitality scale (SVS) was designed to specifically assess eudaimonic psychological functioning and wellbeing, and has been validated in multiple populations, including people with RA (Rouse et al., 2015; Ryan & Frederick, 1997). The SVS has been recommended for use in research and clinical settings, and is particularly relevant to people with RA, who may experience many of the negative symptoms assessed in the questionnaire (e.g., lack of energy) (Rouse et al., 2015).

Quality of Life

In comparison to subjective vitality, quality of life measures generally assess hedonic aspects of wellbeing. Quality of life is defined as "a broad ranging concept incorporating in a complex way the person's physical health, psychological state, level of independence, social relationships, person's beliefs and their relationship to salient features of the environment" by the World Health

Organisation (WHO) (WHO, 1995). Many core OMERACT outcomes, such as high disease activity, pain, fatigue, poor mental health and psychological wellbeing and impaired functional ability, are determinants of poor quality of life in people with RA (Matcham et al., 2014; Rosa-Gonçalves et al., 2018; Senra et al., 2017; Wan et al., 2016). In detail, increased disease activity is related to worsened pain, fatigue, impaired functional ability and poorer mental health and wellbeing, which, in turn, can all negatively impact quality of life (Matcham et al., 2014; Rosa-Gonçalves et al., 2018).

Measurement

Quality of life and hedonic wellbeing can be quantified using questionnaires such as the Quality of Life Scale (Burckhardt & Anderson, 2003), SF-36 (Ware & Sherbourne, 1992) or EQ-5D (Balestroni & Bertolotti, 2012), which are validated and reliable measures in patients with RA (Hurst et al., 1997; Linde et al., 2008). The SF-36 is a particularly comprehensive measure, which has been most widely used in people with RA (Kanecki et al., 2013; Rouse et al., 2015), and captures various components of positive psychological functioning. As quality of life has been deemed a key patient- and clinician-important outcome by OMERACT and COMET (Bartlett et al., 2012; Boers et al., 1994; Van Tuyl & Boers, 2015; Williamson et al., 2017), studies in RA patients frequently measure quality of life as an indicator of intervention efficacy (Barber et al., 2017).

Improving core OMERACT outcomes in RA: the role of Physical Activity and Sedentary Behaviour

Pharmacological interventions have been traditionally used to gain tight control over RA disease activity, in order to slow joint damage and reduce the experience of debilitating symptoms for people living with RA. For example, as high DAS28 scores can predict increased pain in people with RA (Walsh & McWilliams, 2012), targeting disease activity through biologic therapies and joint surgery have significantly impacted RA pain. Similarly, pharmacological therapies aim to improve RA-

related function through targeting disease activity, inflammation and minimising joint damage (Maini et al., 2004).

However, RA is complex, and the pharmacological control of disease activity does not always result in resolution of symptoms. For example, there is currently no effective pharmacological treatment for fatigue in RA, and many people with well controlled disease activity still experience symptoms of fatigue (Olsen et al., 2016). In addition, pharmacological therapies for depression can be less effective in people with RA, due to drug interactions (Fiest et al., 2017; Warner-Schmidt et al., 2011), and research suggests that RA patients with persistent anxiety may have reduced responses to pharmacological treatment (Matcham et al., 2016).

As conventional treatments generally only target disease symptoms and progression, a multidisciplinary approach, involving non-pharmacological therapies, has been recommended to manage RA symptoms (Gettings, 2010). Some non-pharmacological interventions have targeted increasing PA, and, more recently, reducing SB, and are increasingly recommended to target core OMERACT patient- and clinician-important outcomes (Fenton et al., 2018a; Hewlett et al., 2011; Katz, 2017a; O'Brien et al., 2021; Rongen-van Dartel et al., 2015).

The following sections will highlight role of PA and SB in the broad context of health, outline methods used to assess PA and SB in research, and detail studies of PA and SB in RA, utilising different PA and SB measurement approaches.

Physical Activity

PA is defined as any bodily movement produced by skeletal muscles that leads to an energy expenditure beyond the resting rate. Exercise is a subcategory of PA, which is planned, structured and repetitive, and done in order to maintain or improve physical fitness (Caspersen et al., 1985). The WHO recommends young and older adults, as well as many clinical populations, should take part in

at least 150-300 minutes of moderate intensity PA (MPA, [3.0-5.99 metabolic equivalent of task (METs)]), or 75-150 minutes vigorous intensity PA (VPA, [≥6.0 METs]) per week (WHO, 2020c), whereby 1 MET equals the amount of oxygen consumed at rest (i.e., 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$) (Ainsworth et al., 2011).

PA is encouraged for people of any age, healthy or diseased, for its physical and mental health benefits. Within the general population, PA is an effective non-pharmacological approach towards lower the risk of developing type 2 diabetes, cancer, hypertension and overall mortality (Helmrich et al., 1994; McTiernan et al., 2003; Schnohr et al., 2003; Wahid et al., 2016; Whelton et al., 2002). PA is often also prescribed to improve mental health, and systematic reviews have emphasised the efficacy of PA for improving quality of life in healthy adults (Bize et al., 2007).

PA is similarly advocated for older adults and people with musculoskeletal conditions to prolong life expectancy, reduce comorbidities and improve health outcomes. PA participation declines with age, and one such method to prolong healthy ageing is to increase PA engagement (Taylor et al., 2004). PA in older adults has benefits for cardiovascular, bone and muscle health (Taylor et al., 2004). Moreover, randomised controlled trials have reported a large antidepressant effect of exercise in older adults with clinical depression (Blumenthal et al., 1999; Singh et al., 2001; Taylor et al., 2004), highlighting the efficacy of PA also for mental health and wellbeing.

The majority of initial research assessing the links between PA with health, involved structured and organised exercise training or moderate to vigorous PA (MVPA). It is now well documented that MVPA is effective at preventing many morbidities (Amagasa et al., 2018; Lee et al., 2012). More recently, research has begun to explore the health benefits of light intensity PA (LPA [1.6-2.9 METs]), as an alternative to MVPA. LPA can involve structured exercise of a light intensity, such as slow walking or yoga, or incidental PA, such as cooking, laundry or dusting. According to data from the National Health and Nutrition Examination Survey (NHANES), adults in the United States spend on

average 7.8 hours per day engaging in LPA (Young et al., 2016). A meta-analysis has found independent associations between LPA with all-cause mortality, cardiometabolic risk factors and metabolic syndrome in healthy adults (Amagasa et al., 2018). Targeting and promoting LPA may therefore have additional health benefits, which may be particularly beneficial and more achievable in older and clinical populations with additional barriers to MVPA engagement (McMullan et al., 2020).

Another method to increase LPA participation can be through increasing engagement in overall PA, accrued through incorporating more activities into an individual's daily lifestyle. This approach of increasing overall PA can be referred to as increasing "lifestyle PA". Lifestyle PA has no formal definition, but it can encompass all PA which is accumulated as part of day-to-day life. Examples of lifestyle PA can include increasing total PA through: incidental PA (i.e., PA built up in small amounts over the course of a day, e.g., walking upstairs); increasing home-based PA (e.g., gardening, household chores); or increasing activities, such as walking, which may be higher intensity in nature (Katz et al., 2018). In addition, by decreasing SB, this can also result in heightened lifestyle PA, as less time spent sitting will assist in increasing an individual's total daily PA engagement.

Targeting and increasing lifestyle PA may be a more feasible and achievable alternative to MVPA to increase PA engagement for people who have additional barriers to being active (Veldhuijzen Van Zanten et al., 2015). In addition, increasing lifestyle PA is being progressively advocated as a cost-effective and clinically meaningful strategy to increase overall PA, and has demonstrated good acceptability and achievability in both healthy individuals and people with musculoskeletal diseases (Chmelo et al., 2013; Duvivier et al., 2013; Giraudet-le Quintrec et al., 2007; Swardh et al., 2020; Thomsen et al., 2017; Van Roie et al., 2010). However, lifestyle PA research is still a relatively novel concept, and few studies have assessed the links between lifestyle PA with health, and the efficacy of

lifestyle PA interventions in healthy participants, and clinical populations. Consequently, lifestyle PA, as an approach to increase overall PA engagement, will be a key focus of this thesis.

Sedentary Behaviour

SB is defined as "any waking behaviour characterised by an energy expenditure ≤1.5 METs while in a sitting or reclining posture" (Ainsworth et al., 2011; Sedentary Behaviour Research Network, 2012). Examples of common SBs include working at a computer, watching television or travelling in a vehicle; so SB can take place in most life domains (Ainsworth et al., 2011).

Previously, populations which did not meet WHO PA recommendations (of 150-300 minutes of MPA, or 75-150 minutes VPA per week, (WHO, 2020c)), were referred to as sedentary (Sedentary Behaviour Research Network, 2012). However, research now recognises SB as a distinct construct to physical inactivity (Fenton et al., 2018a; Owen et al., 2010). Therefore, someone can be classed as sedentary but also physically active if they spend large portions of their day sitting, but still engage in sufficient PA to meet the recommended guidelines (i.e., someone who has an office job but cycles to and from the office every day). Therefore, more recent activity-based studies have included both SB and PA as independent and distinct determinants and outcomes, and research is increasingly being conducted to investigate the independent contribution of SB to health in healthy and clinical populations.

SB is becoming increasingly prevalent in people of all ages and from all backgrounds. A British Heart Foundation report from 2017 estimated that that the average person in the UK spends 76 days per year sitting, and almost 30 hours per week watching TV (Foundation, 2017). In healthy young adults, SB is independently related to poor cardiovascular and cardiometabolic health, cancer, metabolic syndrome and premature mortality (Biswas et al., 2015; Carson et al., 2014; Ekelund et al., 2019; Green et al., 2014). Research in older adults has revealed independent associations between device-

assessed SB with functional ability, and fitness, which are key outcomes affecting quality of life (Rosenberg et al., 2016; Santos et al., 2012). Studies in older adults also suggest SB is independently related to other deleterious health outcomes, such as pain, fatigue, indicators of poor mental health and psychological wellbeing, as well as mortality (Balboa-Castillo et al., 2011; Chastin et al., 2014; Greenwood-Hickman et al., 2015; Okely et al., 2019; Park et al., 2018; Rosenberg et al., 2016; Saunders et al., 2020). These findings have been strengthened by results of 2 meta-analyses showing that ≥9.5 hours per day spent sedentary is linked to higher risk of death in older adults (Ekelund et al., 2019). This meta-analysis also reported that >3 hours per day watching TV was related to premature mortality, regardless of levels of PA engagement (Ekelund et al., 2016).

The Importance of Accurately Measuring Physical Activity and Sedentary Behaviour

There is a wealth of research investigating the relationships between PA and SB with health outcomes in various populations. However, in order to accurately measure engagement in these behaviours, make links with important health outcomes, and evaluate the effectiveness of interventions targeting PA and SB, free-living PA or SB must be accurately and reliably quantified. The following section will focus on the tools currently employed to measure PA and SB.

First, when assessing the usefulness of a tool for measuring PA or SB, it is important to appreciate the components of these behaviours. There are various intensities and types of PA which can have varied and differing associations with health outcomes. The acronym 'FITT' can be used to categorise the multidimensional elements of PA which should be considered when measuring PA. FITT stands for: frequency, intensity, time, and type of PA. Different measurement tools are able to specifically measure these different elements of PA with varying degrees of accuracy. Intensities can include LPA and MVPA, whilst types can include: non-exercise or leisure time PA, walking, occupational or transport-related PA. Different PA types have differing associations and relationships with various health outcomes (Fenton et al., 2018b). For example a systematic review and meta-analysis by

Samitz et al. (2011) has demonstrated there is an inverse relationship between domain specific PA with all-cause mortality in healthy adults, with particularly strong associations for exercise, leisure PA and activities of daily living compared to occupational and transport-related PA (Samitz et al., 2011). To add to this, a population-based study has found leisure-time PA was positively related to quality of life, whilst domestic and transport-related PA were inversely related to quality of life (Jurakić et al., 2010). Therefore, these highlight the need for accurate and reliable measurement tools to quantify different FITT elements of PA in research.

Regarding SB, it is not just overall sedentary or sitting time that has implications for health. There are many dimensions and types of SB which can also be quantified by measurement tools, which demonstrate independent associations with different aspects of health (Fenton et al., 2018a). Following on from the FITT acronym for PA, there exists a 'SITT' acronym (Tremblay et al., 2010). The components of SITT are:

- S – SB frequency (number of sedentary bouts)
- I – interruptions (number of breaks during sedentary time)
- T – time (duration of SB)
- T – type (context of SB)

As with FITT, different measurement tools are able to quantify different elements of SITT and SB. For example, typically questionnaires and device-based tools attempt to quantify duration of SB (i.e., sedentary time, $_{SI}T_{T}$). Device-based measures can also offer the ability to measure the number of sedentary bouts (S_{ITT}), average sedentary bout length, and the frequency of interruptions ($_{S}I_{TT}$) in sedentary time (Tremblay et al., 2010). On the other hand, self-report questionnaires and diaries may be able to assess type of SB ($_{SIT}T$), enabling understanding into the context of the behaviours performed. For example, SB types can be: TV viewing, occupational, transport and leisure time SB (Fenton et al., 2018a), and each of these types of SB could have different implications for health.

Physical Activity and Sedentary Behaviour Measurement Methods

PA and SB can be measured either subjectively, through questionnaires or self-report diaries, or using device-based measures. The following section will provide details of different methods of quantifying PA and SB, and **Table 1.2** summarises the different methods and their advantages and disadvantages for use in research.

Self-report Methods

Self-report methods used to assess PA and SB include questionnaires, such as the International Physical Activity Questionnaire (IPAQ), Sedentary Behaviour Questionnaire (SBQ), National Institutes of Health-American Association of Retired Persons (NIH-AARP) Diet and Health Study questionnaire and SIT-Q-7d Questionnaire (Cleland et al., 2018; DiPietro et al., 2018; Gierach et al., 2009; Rosenberg et al., 2010; Wijndaele et al., 2014). Questionnaires are the most extensively used self-report assessment tool to quantify free-living PA and sedentary time ($_{SI}T_T$) in large-scale observational and epidemiological studies (Healy et al., 2011; Sylvia et al., 2014). They are often designed to measure a particular type of PA or SB ($_{SIT}T$), or assess time spent in activities of SB over a particular time period (e.g., past week or month, $_{SI}T_T$) (Cleland et al., 2018). Questionnaires can also vary by how data is reported (i.e., time per day, calories, activity score), and obtained (i.e., computerised or written questionnaire) (Sylvia et al., 2014).

PA questionnaires have shown inconsistent results in validation studies when compared against doubly-labelled water (DLW) method, the gold standard measure of energy expenditure (Sylvia et al., 2014; Westerterp, 2009). Advantages and limitations of questionnaires are summarised in **Table 1.2**. Benefits include relative ease and low cost of administration, with minimal burden for participants (Healy et al., 2011; Sylvia et al., 2014). However, questionnaires asking about PA and SB over large timeframes are susceptible to recall bias (Dowd et al., 2018). People frequently over-report PA, under-report sedentary time and misinterpret questions, resulting in inaccurate answers about the

duration, frequency and intensity of PA or SB (Sylvia et al., 2014). Nevertheless, self-report measures have the ability to contextualise PA and SB ($_{SIT}$T), and they can perhaps be useful additional tools to device-based measures which do not have this capability.

PA diaries offer an alternative self-report method of PA quantification. For example, the Bouchard Physical Activity Record is a diary which asks participants to report the dominant activity done every 15 minutes, in real-time, for a period of 3 days (Bouchard et al., 1983; O'Brien et al., 2018b; Sylvia et al., 2014). Diaries has been widely used to provide context to PA and SB ($_{SIT}$T) in clinical research (O'Brien et al., 2018b). PA diaries are able to overcome some questionnaire limitations (**Table 1.2**): they can be completed multiple times a day, and have less potential for recall bias (Van Der Ploeg et al., 2010). However, for this reason, activity diaries can be more of a burden for participants, and are still a subjective assessment method (Rachele et al., 2012).

Device-based Methods

Device-based measures are another method of quantifying free-living PA and SB which have become more popular in recent times. Compared to self-report methods, device-based measures have demonstrated less variability and greater reliability at assessing PA and SB, particularly across different activity intensities (Dowd et al., 2018). Consequently, they are often preferred to self-report methods in small-scale studies and in some large-scale epidemiological research where financial resources allow (Brady et al., 2019; Matthews et al., 2008). PA can be measured using pedometers and accelerometers, whilst sedentary time ($_{SI}T_T$) and other elements of SB can be measured using accelerometers and posture sensors.

Pedometers

Pedometers can measure daily time in PA ($_{FI}T_T$) by recording the number of steps a person takes, and are typically worn attached to clothing or a belt on the hip. Steps are measured via a horizontal lever arm which gets deflected when a person's hip displaces vertically during walking. Each time the lever

arm moves past a pre-specified threshold, a step is counted. As step count accumulates, this provides an estimate of PA over the time period of interest (Liu et al., 2015; Sylvia et al., 2014).

A typical experimental protocol involves participants wearing pedometers during waking hours for 4-7 days, to give a comprehensive overview and reliable estimate of average daily step count and PA (Brady et al., 2019). Advantages of pedometers are summarised in **Table 1.2**, and include that they are small, inexpensive and easy to wear (Brady et al., 2019; Sylvia et al., 2014). They can pick up PA of short durations, which are typically missed by self-report methods (Sylvia et al., 2014). Pedometers are highly validated and reliable measures of PA in healthy adults (Tudor-Locke et al., 2005), demonstrating ±3%, and ±10% error compared to direct observation at counting steps in controlled laboratory conditions and free-living settings, respectively (Liu et al., 2015). However, they are unable to assess non-ambulatory activities (e.g., cycling or swimming) and cannot quantify the frequency (F_{ITT}), type ($_{FIT}T$) or intensity ($_{F}I_{TT}$) of PA. This can be problematic in clinical research investigating relationships between different types of PA with health (Brady et al., 2019). Moreover, pedometers can underestimate steps at slow walking speeds; demonstrating 56% error in participants with a slow step cadence of 50 steps per minute (Liu et al., 2015). This lack of sensitivity at assessing PA in people with slow step cadence, such as clinical or older, less mobile populations (Martin et al., 2012), means pedometers are not recommended for use in research involving these groups of participants.

Accelerometers

Accelerometers offer an alternative to pedometers for assessing free-living PA, and are also frequently used to quantify SB in research, particularly in clinical and older adult populations. They are small, lightweight devices which measure activity via acceleration of body movements of the area of the body they are attached to. Accelerometers can be used to assess frequency (F_{ITT}), intensity ($_{F}I_{TT}$) and duration ($_{FI}T_{T}$) of PA, as well as different features of SB. Specifically, in reference to the SITT

acronym, accelerometers can measure sedentary time ($_{SI}T_T$), sedentary bouts (S_{ITT}) and sedentary breaks ($_{SI}I_{TT}$).

Accelerometers record bodily accelerations as an analogue voltage created by a piezoelectric instrument which can sense compression from 1 (vertical (Y)) or 3 (the vertical (Y), antero-posterior (Z), and medio-lateral (X)) axes (Vanhelst et al., 2012). Accelerometers whereby the piezoelectric instrument is only sensitive to compression in a vertical direction, that record movement from vertical axis only, are called uniaxial accelerometers. Those recording accelerations from 3 axes are triaxial accelerometers. Triaxial accelerometers are able to pick up a greater variety of body movements and activities which require movement in a antero-posterior (Z) or medio-lateral (X) direction (Vanhelst et al., 2012). The acceleration data from the 3 axes are used to calculate vector magnitude (VM) ($VM = \sqrt{(axisY^2 + axisZ^2 + axisX^2)}$).

Acceleration data can be downloaded as raw data (i.e., gravity-based acceleration (g) units, where $1g = 9.81ms^{-1}$) or as "count-based" data. Manufacturer software can be used to compress and time-stamp raw accelerations to create an "accelerometer count" per unit of time, called an epoch. The higher the count within each epoch, the higher the intensity of PA performed (Vanhelst et al., 2012). Either raw or count-based data can be used and interpreted to quantify PA. Furthermore, the raw or count-based output can be calibrated against energy expenditure (as measured by indirect calorimetry) to identify PA "thresholds", or "cut-points", which correspond to different intensities of PA (e.g., LPA, MVPA), or sedentary time (Freedson et al., 1998; Matthews, 2005; Migueles et al., 2017). Different cut-points have been developed specifically for different makes and models of accelerometers (Matthews, 2005), as well as being population and placement site specific (Migueles et al., 2017).

Typically, accelerometers are worn on the hip, but can also be worn on the wrist, ankle or foot (Rhudy et al., 2020). Experimental protocols typically require participants to wear accelerometers for

7 days during waking hours, in order to reliability assess free-living PA and SB (Dowd et al., 2018). Advantages and limitations of using accelerometers in research to quantify free-living PA and SB are summarised in **Table 1.2**. Benefits include their ability to assess multiple components of FITT and SITT. Specifically they can measure the duration, intensity and frequency of activity. In addition, accelerometry is precise, reproducible, and cause minimal interferences with normal daily living (Vanhelst et al., 2012). Various makes and models of accelerometer have been validated for measuring different intensities of PA and SB amongst both healthy and clinical populations in laboratory and free-living settings (Calabro et al., 2014; Dowd et al., 2018; Feito et al., 2012; Kelly et al., 2013). In detail, for treadmill walking in a laboratory setting, correlations of r=.88 were demonstrated between the counts per minute (cpm) of a widely used accelerometer model with the gold standard, indirect calorimetry (Kelly et al., 2013). In free-living settings, accelerometers have displayed good inter-monitor reliability for estimates of LPA and MPA (Calabro et al., 2014). Further studies have confirmed these findings, and the superior validity of accelerometers compared to self-report measures has led to them being more commonly used in large-scale epidemiological research to quantify free-living PA and SB (Matthews et al., 2008; Troiano et al., 2008).

However, using accelerometers in research can be complex, and has its limitations (**Table 1.2**). For example, analysing and interpreting accelerometer data requires researchers to make their own decisions regarding data capture methods (i.e., cut-points, valid wear days, epoch length). As a result, analytical approaches are observed to be heterogeneous across studies (Migueles et al., 2017), and there becomes an inherent degree of subjectivity in data processing and analysis. In addition, accelerometers are comparatively more expensive than many self-report measures, and despite causing minimal interference in daily activities, many studies have reported poor wear compliance of hip-worn accelerometers. One cohort study in adults demonstrated only 78% adherence over 3 to 6 days of hip-worn accelerometer wear (Evenson et al., 2015). Finally, accelerometers cannot accurately assess activities whereby there is little bodily movement, but require energy expenditure,

such as weight-lifting or cycling, particularly when they are worn on an area of the body not involved in the movement (**Table 1.2**) (Brady et al., 2019). They are also unable to measure the type or context of PA ($_{FIT}$T) or SB ($_{SIT}$T), which are important dimensions to understand the types of activities different populations partake in (Tremblay et al., 2010). However, these activities can be captured via other means, such as questionnaires, so many studies employ multiple methods to give a greater overview of total PA. Despite these limitations, accelerometers remain accurate and reliable research-grade measurement tools for assessing free-living PA and SB, frequently employed in large-scale epidemiological and intervention studies (Matthews et al., 2008; Troiano et al., 2008).

ActiGraph Accelerometer

The ActiGraph GT3X triaxial accelerometer is one of the most frequently used research-grade accelerometers (Duncan et al., 2018). The most common wear site for the GT3X is attached to an elasticated belt, on the hip at the anterior axillary line (**Figure 1.1**). When worn on the hip, the GT3X gives reliable and valid measurements of sedentary time and PA in healthy adults (Aadland & Ylvisaker, 2015; Clevenger et al., 2020a; Santos-Lozano et al., 2013).

Figure 1.1: ActiGraph GT3X+ and example of accelerometer wear on elasticated band on the right

The ActiGraph GT9X is a more recent triaxial ActiGraph model. It calculates accelerations and activity counts using the same method as the GT3X, with additional features: a display screen, gyroscope and secondary accelerometer to measure rotation, movement and body position (Duncan et al., 2018; Ekelund et al., 2020; Loprinzi & Smith, 2017). When worn on the hip (**Figure 1.2**), the GT9X reliably assesses laboratory and free-living PA in healthy young and older adults, and has been validated

against previous ActiGraph models, demonstrating good to excellent inter-monitor reliability (Clevenger et al., 2020a; Clevenger et al., 2020b; Montoye et al., 2018b). Due to being a relatively new model, the GT9X has not been widely validated against gold standard measures for measurement of sedentary time.

Figure 1.2: ActiGraph GT9X and example of accelerometer wear on elasticated band on the right hip and attached to a watchstrap on the wrist

As with other triaxial accelerometers, the ActiGraph GT3X and GT9X capture movement on 3 axes, which can be used to determine VM. These accelerations, along with a user-defined epoch length, can be compressed into activity counts using ActiGraph-specific software, ActiLife (ActiGraph, LLC., Pensacola, Florida, USA). ActiLife software can also classify data into "wear time" and "non-wear time" using detection algorithms. Non-wear time is defined as a certain number of consecutive 0 activity counts, and can include time intervals when participants remove accelerometers such as to sleep and water-based activities. It is critical to determine non-wear time so these periods are not incorrectly classified as sedentary time or other activity intensities (Choi et al., 2011; Mâsse et al., 2005).

Automated algorithms within ActiLife use criteria to detect and remove non-wear periods. The most widely used algorithm was developed by Troiano (2007), and has subsequently been used in the NHANES large-scale population-based dataset (Matthews et al., 2008). Non-wear was defined as ≥60 consecutive minutes of zero VM counts, with a spike tolerance of 2 minutes of counts between 0-100 counts (Troiano, 2007). Once non-wear has been determined, researchers must decide on how much accelerometer wear determines a "valid day" and "valid week". From this, researchers can then

determine if a participant has worn the accelerometer for a sufficient amount of time to be included in further analysis. Typical studies in healthy adults use criteria of ≥4 valid days/week, including ≥1 weekend day, to comprise valid wear time (Troiano et al., 2008). Participants log books or activity diaries are frequently checked to help determine wear and non-wear.

Once non-wear data has been removed, the next stage of analysis is to apply cut-points to activity counts (i.e., count-based cut-points), to quantify time spent in different activity intensities (e.g., sedentary time, LPA, MVPA etc.). Although, there are other approaches to process and analyse ActiGraph data (e.g., using intensity gradients and average acceleration to capture the volume and intensity of PA (Rowlands et al., 2018; Rowlands et al., 2019b), cut-points remain the most commonplace and easy to use method (Montoye et al., 2018a; Rowlands et al., 2019a). For this reason, this thesis will focus on the cut-point method of PA quantification by accelerometers. The most commonly used ActiGraph uniaxial cut-points (calculated using accelerations from the vertical (Y) axis only) were defined by Troiano et al. (2008), as sedentary time = <100 cpm, LPA = 100-2019 cpm and MVPA = ≥2020 cpm, and were developed in a study using the uniaxial 7164 ActiGraph model. The cut-points were defined from weighted average calculations from results of multiple calibration studies assessing walking and running activities only (Brage et al., 2003; Freedson et al., 1998; Leenders et al., 2001; Yngve et al., 2003). Since then, these cut-points have been widely employed in epidemiological datasets (Matthews et al., 2008). However, they have demonstrated poor validity at assessing PA when compared to the "gold standard" PA measure of indirect calorimetry, and poor validity at estimating sedentary time compared to the activPAL™ ("gold standard" free-living measure of sedentary time) (Crouter et al., 2013; Koster et al., 2016; Watson et al., 2014). In spite of this, Troiano cut-points remain the most frequently used in accelerometer research for measurement of PA and SB, even though they are uniaxial, and were developed specifically for older uniaxial ActiGraph models (Koster et al., 2016).

The most common placement site of ActiGraph accelerometers is on the hip, although research has found that this can result in poor wear compliance (Rhudy et al., 2020; Troiano et al., 2014). A popular alternative with research participants (Rhudy et al., 2020; Troiano et al., 2014), and the general population (Montoye et al., 2020), is to wear accelerometers on the wrist. Wrist-worn accelerometers may be able to capture non-ambulatory activities (e.g., ironing, dusting, upper body resistance exercises), and obtain PA and SB information from people with atypical gait (Diaz et al., 2018; Troiano et al., 2008). Some epidemiological studies now employ wrist-worn accelerometers to measure PA for these reasons (Troiano et al., 2014). The ActiGraph GT3X and GT9X have the option of being worn, attached to a watchstrap, on the wrist (**Figure 1.2**).

Growing research has indicated that accelerometer cut-points should be placement site and model specific (Migueles et al., 2017; Rhudy et al., 2020). Consequently, newer cut-points have been developed specifically for the hip- and wrist-worn triaxial ActiGraph models (i.e., GT3X, GT9X). For example, Sasaki et al. (2011) triaxial cut-points were developed specifically for hip-worn ActiGraph GT3X. They define PA and sedentary time as: sedentary time= ≤150 cpm, LPA= 151–2690 cpm, MPA= 2691–6166 cpm, VPA= 6167–9642 cpm, and very vigorous intensity PA (VVPA)= >9642 cpm (Sasaki et al., 2011). In addition, Montoye et al. (2020) have developed triaxial cut-points for the wrist-worn ActiGraph GT9X as: sedentary time = <2860 cpm, LPA= 2860−3940 cpm and MVPA= ≥3941 cpm.

As triaxial cut-points use data from 3 axes, research has suggested that they may give a more comprehensive measure of different intensities of PA and sedentary time compared to uniaxial cut-points (Kozey-Keadle et al., 2014; Leeger-Aschmann et al., 2019). However, these triaxial placement-specific cut-points must be compared against widely used uniaxial cut-points (e.g., Troiano et al. (2008)), in order to determine the extent to which PA and sedentary time estimates differ when different cut-points are employed to the same data. This will elucidate the extent to which studies utilising different devices (uniaxial vs triaxial), placement sites (wrist vs hip), and analytical methods

(e.g., the cut-points selected to quantify PA and sedentary time) when employing accelerometers, may influence PA and SB outcomes reported in research. As such, one of the aims of this thesis is to compare the estimates of LPA, MVPA and sedentary time produced by newer triaxial wrist-worn GT9X-specific cut-points (developed by Montoye et al. (2020)) and triaxial hip-worn GT9X-specific cut-points (developed by Sasaki et al. (2011)) with uniaxial hip-worn cut-points (developed by Troiano et al. (2008)) (***Chapter 4***). Results of this study will enable comparisons across studies, and aid in the movement towards more homogenous analytical approaches and measurement methods for quantification of PA and SB in research.

The activPAL™

The activPAL™ is a postural classification device, typically worn on the thigh in the mid-anterior position (**Figure 1.3**) (Edwardson et al., 2017). Its major differences from previously mentioned accelerometers is that it incorporates gravity into its readings, allowing for postural allocation to determine free-living sedentary (sitting/lying), upright (standing) and ambulatory (walking) activities (Chan et al., 2017). Advantages and limitations of using the activPAL in clinical research are summarised in **Table 1.2**. In brief, the activPAL is small, lightweight and easy to wear, due to its thigh placement. The activPAL has the ability to measure many components of SITT, including sedentary bouts (S_{ITT}), interruptions ($_{S}I_{TT}$) and sedentary time ($_{SI}T_{T}$). However, the activPAL is relatively expensive compared to self-report methods. In addition, the activPAL is limited in its ability to classify lower limb movement and to detect steps of a slow cadence (≤0.47 m/s) (Taraldsen et al., 2011), and may therefore inaccurately quantify step count in elderly or clinical populations (Chan et al., 2017; Grant et al., 2010).

A growing number of studies have determined the validity of the activPAL in for measuring components of PA and SB different populations (Dowd et al., 2012; Edwardson et al., 2017; Kozey-Keadle et al., 2011; Sellers et al., 2016). Consequently, the activPAL is considered the “gold standard”

device-based measure for the assessment of free-living sedentary time, and is used as a criterion when assessing the validity or reliability of new accelerometers or cut-points (Edwardson et al., 2017; O'Brien et al., 2020). In accordance with this, one of the aims of this thesis is to assess the validity of triaxial wrist-worn GT9X-specific cut-points (developed by Montoye et al. (2020)), triaxial hip-worn GT9X-specific cut-points (developed by Sasaki et al. (2011)), and uniaxial hip-worn cut-points (developed by Troiano et al. (2008)) compared to the activPAL for measurement of free-living sedentary time in healthy adults (***Chapter 4***).

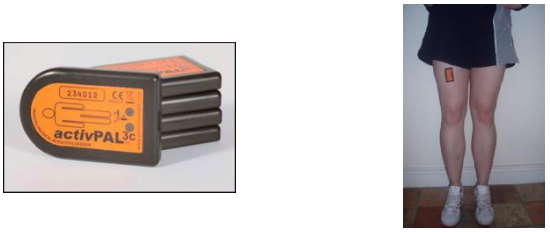


Figure 1.3: The activPAL and example of wear on the thigh in the mid-anterior position

Table 1.2: Measurement methods of physical activity and sedentary behaviours in clinical research

Method of assessment	*Self-report or Device-based*	*Examples*	*Advantages*	*Disadvantages*
Questionnaires	Self-report	IPAQ SBQ NIH-AARP Diet and Health Study questionnaire SIT-Q-7d Questionnaire	• Easy for participants and researchers • Low cost of administration • Minimal burden for participants • Can provide context to activities • Useful in large-scale research	• Recall bias- participants frequently over-report PA, under-report SB and misinterpret questions, resulting in inaccurate answers about the duration, frequency and intensity of PA or SB • Inconsistent results in validation studies when compared against DLW method • Subjective method
Activity diaries	Self-report	Bouchard Physical Activity Record	• Less potential for recall bias due to real-time completion • Easy for participants and researchers • Low cost of administration	• Increased participant burden due to having to complete multiple times a day • Memory bias may occur when diaries not completed in real-time • Subjective method
Pedometers	Device-based	Yamax pedometer DigiWalker	• Small • Inexpensive • Easy to wear • Can pick up PA of short durations • Highly validated and reliable measure of PA in healthy adults	• Unable to assess non-ambulatory activities • Cannot quantify the frequency, duration or intensity of PA • Underestimate steps at slow walking speeds • Lack of sensitivity when assessing PA in clinical populations
Accelerometers	Device-based	ActiGraph GT3X ActiGraph GT9X	• Can assess PA duration, intensity and frequency • Precise and reproducible • Minimal interference with daily living • Widely validated and reliable for measuring different intensities of PA • Commonly used in large epidemiological studies	• Data analysis and interpretation involves skill and expertise • Approaches can be heterogeneous across studies • Poor wear compliance of hip-worn accelerometers • Cannot detect posture • Cannot accurately assess activities with little bodily movement, but require energy expenditure • More expensive than self-report methods
activPAL™	Device-based	activPAL™	• "Gold standard" measure of sedentary time • Small and lightweight • Easy to wear with high wear compliance • Can detect posture	• More expensive than self-report methods • Limited ability to classify lower limb movement • Limited ability to detect steps of slow cadence

Note: IPAQ= International Physical Activity Questionnaire, SBQ= Sedentary Behaviour Questionnaire, NIH-AARP= National Institutes of Health-American Association of Retired Persons, PA= Physical Activity, SB= Sedentary Behaviour, DLW= Doubly Labelled Water.

Physical Activity and Sedentary Behaviour in Rheumatoid Arthritis

The incidence of PA and SB, and relationships with health are well documented in healthy adults. In comparison, fewer studies have investigated PA and SB, and their associations with health outcomes in people living with RA. The following section will describe measurement methods, and prevalence of PA and SB in individuals with RA, and the links between PA and SB with health, with a focus on core OMERACT patient- and clinician-important outcomes.

Measurement of Physical Activity in Rheumatoid Arthritis

In order to accurately assess PA, researchers must weigh up the benefits and limitations of measurement tools to decide which is most suitable for the population of interest. In people with RA, the most common method to assess PA is self-report measures, such as the IPAQ (Cleland et al., 2018). However, similarly to healthy individuals, patients with RA have the tendency to over-report PA levels (Yu et al., 2015). A study assessing the agreement between the IPAQ with accelerometer-measured PA in patients with RA found participants reported less time spent in LPA and more MVPA compared to accelerometer (ActiGraph GT3X) measurements (Yu et al., 2015). In addition, only self-reported MPA and MVPA intensities were correlated with VO_2 max (assessment of cardiorespiratory fitness) (Yu et al., 2015). Nevertheless, self-report questionnaires may have a place in large-scale and epidemiological studies in people with RA, where device-based measurement of PA is not feasible.

Alternatively, there is a growing body of research that has used device-based measures to quantify free-living PA in people with RA (Fenton et al., 2017; Fenton et al., 2018b). Reliability and validity in the RA population specifically must be determined in order for these devices to be deemed accurate for measuring PA in high-quality research. Associations between accelerometer (ActiGraph GT3X) assessed PA with VO_2 max have been investigated in patients with RA, and results found significant correlations between the two measures for all PA intensities (Yu et al., 2015). More recently, a

validation study by O'Brien et al. (2020) created RA-specific ActiGraph GT3X triaxial cut-points. These cut-points were: LPA= 245–2501 cpm; MPA= ≥2502 cpm; and showed good validity when compared against indirect calorimetry for assessments of all PA intensities (O'Brien et al., 2020). From this, future studies are able to apply these RA-specific triaxial cut-points to ActiGraph GT3X data from RA patients to give more valid and reliable measurement of free-living PA, than has been possible previously.

Levels of Physical Activity in Rheumatoid Arthritis

PA guidelines for people with RA are not different to the WHO recommendations advocated for the general population. That is, at least 150 minutes MPA or 75 minutes VPA, as well as muscle strengthening activities are recommended at least twice per week (Metsios et al., 2015). However, research has shown that only a small proportion of people with RA meet these PA recommendations (Iversen et al., 2017), and many people living with RA are less active than the general population. Cross-sectional studies have reported that 70% people with RA do not partake in regular PA (Metsios et al., 2015; Sokka et al., 2008; Tierney et al., 2012), and a systematic review found lower PA levels in the RA population, compared to healthy controls (Tierney et al., 2012). However, the majority of studies included in the review used self-report measures to assess free-living PA, and people have previously shown that they are likely to overestimate engagement in self-reported PA compared to device-based measures (Hagstromer et al., 2010; Yu et al., 2015). Therefore, levels of PA are probably lower than those reported in these self-report studies, and results should be interpreted with caution (Tierney et al., 2012).

As well as investigating total PA, it is important to understand the composition of PA by assessing different intensities of PA. Khoja et al. (2016) conducted a cross-sectional study using accelerometry, and estimated that participants with RA spent 3.5 hours/day in very light intensity PA [1.1 − 1.9 METs], 2.1 hours/day in LPA and 35 minutes/day in MPA, with only 17% participants meeting PA

recommendations (Khoja et al., 2016). Another cross-sectional study assessing accelerometer-measured PA was conducted by Fenton et al. (2020b), who reported RA participants spent over 4 hours/day in LPA and 18 minutes/day in MVPA. These findings indicate that the majority of daily PA among people living with RA may be accumulated through LPA. However, few studies have used accelerometry to estimate overall PA and time spent in different PA intensities in individuals with RA, and longitudinal research is needed to strengthen and confirm these findings.

Physical Activity and Health in Rheumatoid Arthritis

Traditionally, rheumatologists would advise people with RA to restrict exercise and PA, due to fears that it may exacerbate symptoms such as pain and fatigue (Cooney et al., 2011; Metsios & Kitas, 2018). However, substantial evidence now indicates PA can significantly improve health outcomes, and is not accompanied by aggravation of RA symptoms or any joint damage (Cooney et al., 2011; Metsios & Kitas, 2018; Van Den Ende et al., 1998; Veldhuijzen Van Zanten et al., 2015). As a result, the ACR and EULAR now recommend PA as an effective non-pharmacological method to improve RA disease activity (Rausch-Osthoff et al., 2018). The following section will evaluate evidence for the associations between PA with core OMERACT patient- and clinician-important outcomes in people with RA. It will concentrate on results of cross-sectional and intervention studies, summarise the relevant limitations and conclusions of existing studies, and identify areas of future research.

Physical Activity and Pain

Pain is frequently reported as a reason for not engaging in PA by people with RA (Law et al., 2013). Previous beliefs and physician recommendations were that PA may aggravate joint inflammation and pain, which have since been disproven (Hernández-Hernández & Díaz-González, 2017). Cross-sectional studies assessing associations between PA and pain in people with RA are limited. However, Haider et al. (2020) found no link between accelerometer-measured MVPA with pain intensity (measured using a VAS). The DAS28 measure of disease activity includes an assessment of number of

painful or tender joints, and a study by Hernández-Hernández et al. (2014) discovered that DAS28 was negatively correlated with IPAQ measured PA. They also demonstrated a trend towards a negative association with accelerometry measured PA. However, painful joint count demonstrated no links with IPAQ or accelerometer-measured PA (Hernández-Hernández et al., 2014). Finally, Prioreschi et al. (2013) found self-reported pain (assessed using SF-36) was not correlated with accelerometer-measured PA. All these studies assessed pain using self-report or single-item methods, and therefore, results should be interpreted with caution. These methods may not be able to measure the fluctuating and multidimensional nature of pain, which is particularly prevalent in people with RA.

Nevertheless, multimodal management, including PA and exercise interventions, are recommended to reduce pain associated with RA (Walsh & McWilliams, 2014) and results of these intervention studies have supported these recommendations (Li et al., 2020; Lofgren et al., 2018a; Neuberger et al., 1997; Nordgren et al., 2015; Stenstrom, 1994). For example, a 12 week home-based exercise intervention demonstrated improvements in pain in patients with RA (Stenstrom, 1994). In addition, Neuberger et al. (1997) found that participation in their 12 week PA program resulted in reduced pain perception, and Lofgren et al. (2018a) reported reductions in global pain at 2 year follow-up after a health-enhancing PA program. Nevertheless, other PA interventions have displayed no effects on pain (Brodin et al., 2008; Cramp et al., 2020; Feldthusen et al., 2016; Gilbert et al., 2018; Katz et al., 2018). Furthermore, narrative and systematic reviews have displayed varied results regarding the efficacy of PA for improving RA-related pain (Pedersen & Saltin, 2006; Plasqui, 2008; Stenström & Minor, 2003).

The mixed results of cross-sectional and intervention studies suggest that mechanisms behind PA-induced analgesia may not be completely understood. Furthermore, it is thought that the application of QST to quantify the multidimensional aspects and mechanisms of RA-related pain in clinical

populations may aid the development of more appropriate and targeted interventions for pain management (Moloney et al., 2012).

Physical Activity and Disease Activity

Disease activity has been shown to be a determinant of participation in PA, as research has found that high disease activity interferes with time spent in MVPA among RA patients (Hernández-Hernández et al., 2014). Observational and intervention studies have found that regular PA in people with RA is associated with lower levels of CRP and ESR inflammatory markers (Hakkinen et al., 2001; Metsios et al., 2009; Stavropoulos-Kalinoglou et al., 2013). In detail, DAS28 (using CRP) has demonstrated significant correlations with IPAQ-measured PA in one cross-sectional study (Hernández-Hernández et al., 2014). Khoja et al. (2016) also found very light intensity PA, LPA and MPA (measured using accelerometry) were significantly correlated with DAS28, but associations were only observed in regression models for very light intensity PA. Conversely, another study found no difference in accelerometer-assessed MVPA between RA patients with low, moderate and high disease activity (measured using CDAI) (Haider et al., 2020).

Further PA interventions have demonstrated mixed results regarding efficacy at improving disease activity. Katz et al. (2018) and Lange et al. (2020) demonstrated reductions in disease activity, with contrasting results reported from other PA interventions (Knittle et al., 2015; Veldhuijzen Van Zanten et al., 2021). However, interventions were heterogeneous in terms of the type and intensity of PA targeted.

Physical Activity and Functional Ability

The links between PA and functional ability have been investigated by cross-sectional studies. One study found accelerometer-measured MVPA was related to HAQ-DI-measured functional ability. However, when models were adjusted for age, sex and education level, these associations did not continue (Haider et al., 2020). Contrastingly, Hernández-Hernández et al. (2014) found HAQ was

negatively correlated with IPAQ and accelerometer-measured PA. In addition, Prioreschi et al. (2013) reported the HAQ-DI was negatively associated with Actical accelerometer-measured PA, even when age and disease duration were adjusted for (r= .34, p=.03). Another study showed accelerometer-measured very light intensity PA, LPA and MPA were inversely associated with HAQ scores, with regression analysis demonstrating that a greater amount of time spent in MPA was associated with higher HAQ scores ($R^2\Delta$= .11, p<.05), and smaller associations with very light intensity PA and LPA ($R^2\Delta$= .05, p<.05) (Khoja et al., 2016). This greater link between higher intensity PA with functional ability has also been reported by Fenton et al. (2017), who found accelerometer-measured MVPA, but not LPA, was correlated with HAQ-DI. Together, this cross-sectional evidence suggests that different intensities of PA have individual and differing links with functional ability.

In addition, there is growing confirmation that PA or exercise training can improve disease activity, fitness, strength and pain in people with RA, which may have downstream beneficial effects on functional ability (Hurkmans et al., 2009; Metsios & Kitas, 2018). The majority of interventions in people with RA have focused on more structured and organised exercise, as opposed to overall PA. One randomised controlled trial revealed that participating in a 2 year high intensity aerobic and strength training exercise program significantly improved functional ability (De Jong et al., 2003), agreeing with findings from other exercise interventions (Hakkinen et al., 2001). When functional ability has been targeted by interventions targeting overall PA, not just exercise, results have been less concrete. Some PA interventions have demonstrated improvements in functional ability (Katz et al., 2018; Nordgren et al., 2015; Van Den Berg et al., 2006; Veldhuijzen Van Zanten et al., 2021), whereas an intervention by Knittle et al. (2015) led to increased PA but no improvements in functional ability.

Physical Activity and Fatigue

Fatigue is commonly reported as one of the major disease-related barriers for not engaging in PA by people with RA (Law et al., 2013; Veldhuijzen Van Zanten et al., 2015). Despite this, there are a limited number of cross-sectional studies that have investigated the links between PA and fatigue in individuals with RA. Hernández-Hernández et al. (2014) reported fatigue (measured using Functional Assessment of Chronic Illness Therapy) showed a nearly significant negative relationship with self-reported PA (measured using IPAQ). Furthermore, associations have been observed between achieving an activity goal and decreased fatigue in people with RA (Weinstein et al., 2009).

Nonetheless, randomised controlled trials have been conducted to assess the effects of organised exercise training of fatigue in people with RA. For example, a 12 week PA program involving regular exercise sessions demonstrated reductions in fatigue (Neuberger et al., 2007). Walking interventions and interventions targeting overall PA in people with RA have also shown positive reductions in fatigue (Feldthusen et al., 2016; Katz et al., 2018; Knittle et al., 2015). Consequently, systematic reviews have found that exercise and PA may be beneficial for managing fatigue in this population (Balsamo et al., 2014; Cramp et al., 2013). Results of one meta-analysis in people with RA revealed that exercising for 12 weeks resulted in a significant reduction in fatigue, compared to non-exercising controls (Rongen-van Dartel et al., 2015). Review findings also included that although land-based aerobic exercise programs positively affected fatigue, effects were not sustained at 24 week follow-up (Rongen-van Dartel et al., 2015). However, this review included aerobic exercise interventions only, and people with RA may have additional disease-related barriers to exercise of a higher intensity, such as pain, fatigue, reduced functional ability and stiffness (Veldhuijzen Van Zanten et al., 2015).

Physical Activity and Mental Health and Psychological Wellbeing

Early studies have described the presence of a relationship between PA and quality of life in people with RA (Chang et al., 2009; Van Lankveld et al., 2000). However, a more recent cross-sectional study by Hernández-Hernández et al. (2014) gave contrasting results, finding no relationship between IPAQ and accelerometer-measured PA with SF-36-measured quality of life in regression analyses. In addition, Prioreschi et al. (2013) assessed correlations between each subscale of the SF-36 with accelerometer-measured daily activity counts, and only found significant correlations for the "composite physical health" subscale.

Intervention studies in people with RA have demonstrated statistically and clinically significant reductions in depressive and anxious symptoms (Feldthusen et al., 2016; Kelley et al., 2015; McKenna et al., 2021). To confirm this, a meta-analysis has shown that aerobic and strength training interventions in the United States caused significant reductions in depressive and anxious symptoms and improvements in quality of life (Kelley et al., 2015). However, this meta-analysis was conducted in patients with different types of arthritis, and not RA specifically. Contrastingly, some interventions targeting overall PA have displayed no significant improvements in psychological health outcomes in individuals with RA (Knittle et al., 2015; Veldhuijzen Van Zanten et al., 2021). For example, a 3 month self-determination theory-based exercise intervention did display enhanced cardiorespiratory fitness, however this was not accompanied by improvements in depression, vitality, quality of life or anxiety. (Veldhuijzen Van Zanten et al., 2021)

Physical Activity and Health in Rheumatoid Arthritis: Limitations and Conclusions

The existing evidence regarding associations between PA with core OMERACT patient- and clinician-important outcomes described above is varied within and between outcomes. For example, particularly for functional ability, fatigue and mental health and wellbeing outcomes, cross-sectional findings were heterogeneous. Furthermore, very few studies explored the links between different

types and intensities of PA with OMERACT outcomes, and the limited findings available require reinforcing. In addition, observational studies using measures which assess multiple dimensions of outcomes, such as QST to assess pain, were scarce. Together, this suggests there is a need for further large-scale observational studies to confirm the presence of the link between PA with core OMERACT outcomes in RA. Specifically, by investigating different types and intensities of PA using reliable device-based measures, this may provide more evidence into the precise types of PA to target in future interventions. Moreover, by implementing multidimensional measures to quantify outcomes, this may give greater insight into the mechanisms behind associations.

As well as the need for further observational studies, the evidence regarding the efficacy of current PA interventions on OMERACT outcomes requires summarising and evaluating. Researchers have highlighted the need for a thorough systematic review and meta-analysis, to draw definitive conclusions about the effectiveness of existing PA interventions (Metsios & Kitas, 2018). From this, further more targeted interventions assessing the effects of PA on these core OMERACT patient- and clinician-important health outcomes in people with RA can be developed. This thesis will address some of these limitations and conclusions, and add to the existing evidence database exploring the relationships between PA and core OMERACT patient- and clinician-important outcomes in people with RA.

Challenges to Increasing Physical Activity in Rheumatoid Arthritis

The majority of previous observational studies in people with RA have assessed the relationships between OMERACT outcomes with prescribed, organised, structured and repetitive exercise and MVPA (Metsios & Kitas, 2018). Although the benefits of partaking in regular MVPA are numerous, people with RA report additional barriers to leading an active lifestyle. Qualitative studies have found that RA patients report RA-related symptoms, such and pain and fatigue, act as additional barriers to PA participation (Larkin et al., 2017). In addition, those patients with highest disease activity, poorest

functional ability, or those who experience regular RA disease "flares", report having more barriers to being active (Khoja et al., 2016; Larkin et al., 2017; Veldhuijzen Van Zanten et al., 2015). The RA population may, therefore, find the high intensity requirement of MVPA not feasible, and so rely on LPA to reach PA recommendations.

In RA, research investigating the role of LPA is scarce. However, early studies have suggested that increased LPA engagement is associated with lower risk of CVD and disease activity, greater functional ability and improved psychological wellbeing (Fenton et al., 2018b; Khoja et al., 2016). This indicates that PA, at any intensity, may be sufficient for people with RA to achieve health benefits (Khoja et al., 2016). LPA may potentially be a more acceptable and achievable approach than MVPA to encourage less mobile populations, such as people with RA, to meet PA guidelines (Buman et al., 2010; Manns et al., 2012). In addition, increasing lifestyle PA (accrued through all activities of daily living) and reducing SB, may be another more feasible alternative for people with RA. However, limited cross-sectional, longitudinal and intervention studies have been conducted exploring the relative contributions of different intensities (e.g., LPA) and types (e.g., lifestyle PA) of PA and SB on OMERACT patient- and clinician-important outcomes in people with RA. Nonetheless, before SB can be targeted through interventions, accurate measurement within research is essential.

Measurement of Sedentary Behaviour in Rheumatoid Arthritis

Until recently, self-report measures were the preferred research method to quantify sedentary time and other elements of SB in people with RA due to being cheap and easy, requiring minimal participant and researcher burden (**Table 1.2**) (Fenton et al., 2018a). Most studies have used sedentary time (minutes/day) as a SB outcome, and have not assessed the other elements and dimensions of SB (i.e., SITT). As with healthy individuals, studies have demonstrated that people with RA often under-report their sedentary time when using self-report measures compared to accelerometry (Gilbert et al., 2016; Tudor-Locke & Myers, 2001; Yu et al., 2015). One study found

that RA participants self-reported significantly less sedentary time (minutes/day) through completion of the IPAQ, compared to estimates from the ActiGraph GT3X accelerometer (Yu et al., 2015).

Consequently, newer studies have employed device-based measures, such as accelerometers (e.g., ActiGraph) or posture sensors (e.g., activPAL) to quantify different elements of SB (i.e., sedentary bouts, interruptions, and sedentary time) in people with RA. However, most commonly accelerometers have been used to quantify total daily sedentary time. In comparison to self-report measures, accelerometers are more expensive and burdensome for participants and researchers (Fenton et al., 2018a). Further advantages and limitations of these measures are outlined in **Table 1.2**.

In terms of accelerometry, the ActiGraph is a widely used accelerometer that has been validated for sedentary time quantification in people with RA (O'Brien et al., 2020). Yu et al. (2015) have demonstrated significant positive correlations between ActiGraph GT3X-measured sedentary time (minutes/day) with VO_2 max, with no associations found for self-reported sedentary time (measured using IPAQ). Interestingly, O'Brien et al., (2020) reported that the sedentary time accelerometer cut-point of <100 cpm (used widely in observational, epidemiological and interventional research (Matthews et al., 2008; Troiano et al., 2008)) showed poor validity when used in studies of RA, when compared to the activPAL, the "gold standard" measure of free-living sedentary time. This is perhaps not surprising given this <100 cpm cut-point was derived from calibration studies in healthy adults, using uniaxial accelerometer data (Brage et al., 2003; Freedson et al., 1998; Leenders et al., 2001; Yngve et al., 2003). In contrast, O'Brien et al. (2020) developed a new RA-specific triaxial sedentary time cut-point of ≤244 cpm. Despite this development in the RA literature, many studies continue to use the <100 cpm cut-point in people with RA (Fenton et al., 2020b; Pinto et al., 2020a; Summers et al., 2019).

One major limitation of using accelerometers to measure SB is their inability to detect changes in posture. That is, accelerometers define SB based on the absence of acceleration, rather than a sitting posture – where sitting is a core facet of the definition of SB. As a result, some standing activities which require little acceleration, such as washing up or washing laundry, may be misclassified as SB (Kozey-Keadle et al., 2011). The activPAL posture sensor overcomes this limitation, as it is able to detect posture, and, as stated above, it is now considered the "gold standard" measure of sedentary time in research. A total of 2 studies have validated the activPAL for sedentary time measurement in people with RA (Larkin et al., 2016; O'Brien et al., 2020). Larkin et al. (2016) showed no significant differences, and a strong relationship between activPAL vs direct observation estimates of sedentary, standing and stepping time. O'Brien et al. (2020) further displayed that the activPAL accurately quantified free-living standing, sitting and stepping time >98% of the time in people with RA. These findings suggest that the activPAL is a valid tool to quantify free-living sedentary time in people with RA (O'Brien et al., 2020), and could be used as a criterion in order to validate other measurement tools in this population.

Levels of Sedentary Behaviour in Rheumatoid Arthritis

SB is particularly prevalent in people with clinical conditions that can hinder lifestyle behaviours and movement, leading to functional limitation, such as individuals with RA. The majority of studies which have used self-report measures in people with RA have reported 4-6 hours sitting per day (Semanik et al., 2004; Yu et al., 2015). However, Gilbert et al. (2016) found that people with RA self-reported spending approximately 13 hours per day sitting. The divergent results may result from heterogeneity in populations (e.g., Semanik et al. (2004) included only female participants whilst Gilbert et al. (2016) included males and females), in measurement methods of sedentary time (e.g., Yu et al. (2015) used the IPAQ, whilst Gilbert et al. (2016) used the Yale Physical Activity Survey (YPAS) to quantify total sedentary time), and time periods covered (e.g., IPAQ used by Yu et al. (2015)

assesses sedentary time over previous week, and YPAS used by Gilbert et al. (2016) asks about total sedentary time over previous month). Few self-report studies have investigated the prevalence of the different elements of the SITT acronym, with most studies solely investigating total sedentary time ($_{SI}T_{T}$). However, Kramer et al. (2012) and Giles et al. (2008) reported people with RA spend 2 hours per day watching TV, a particularly deleterious type ($_{SIT}T$) of SB.

Device-based observational studies have shown that people with RA accumulate approximately 8−9 hours of sedentary time per day, equivalent of spending between 60−70% of the day sedentary (Fenton et al., 2017; Gilbert et al., 2016; Hammam et al., 2019; Yu et al., 2015). This is similar to device-based studies in older adults who report 9−11 hours of daily sedentary time (Hajna et al., 2018; Harvey et al., 2015). However, other studies using devices to quantify sedentary time in people with RA, report up to 19 hours per day spent sedentary (Huffman et al., 2014; Paul et al., 2014). Variability in sedentary time estimates may be due to differences in study populations, measurement tools and data capture methods used, which can be highly variable between studies.

Sedentary Behaviour and Health in Rheumatoid Arthritis

Cross-sectional research in RA suggests the presence of relationships between device-based or self-report different elements and dimensions of SB with pain, disease activity, functional ability, fatigue and indicators of mental health and wellbeing (Fenton et al., 2017; Greene et al., 2006; Huffman et al., 2014; Khoja et al., 2016; O'Leary et al., 2021). The following sections will review the current cross-sectional, longitudinal and intervention studies demonstrating links between SB and these core OMERACT patient- and clinician-important outcomes in people with RA, and summarise the limitations and conclusions of these studies.

Sedentary Behaviour and Pain

As pain is the most commonly reported symptom in people with RA, and PA interventions have shown promising results at reducing RA-related pain, there is growing research into SB as a target to minimise pain in this population. Cross-sectional studies have indicated that a relationship exists between SB and pain in people with RA (Greene et al., 2006; O'Leary et al., 2021). In detail, Greene et al. (2006) found higher self-reported sedentary time was related to greater pain intensity. To strengthen these findings, a study using the activPAL to assess daily sedentary time demonstrated positive associations between sedentary time with pain intensity and number of painful joints (O'Leary et al., 2021). However, multivariable analysis indicated that these associations were not independent, suggesting pain may not have a significant influence on sedentary time in people with RA (O'Leary et al., 2021). In addition, Huffman et al. (2014) found no association between accelerometer-measured sedentary time with pain.

To my knowledge, only 1 longitudinal study has been conducted investigating the link between SB with pain in people with RA. O'Brien et al. (2021) conducted longitudinal correlations, regressions and path analysis to examine the relationship between sedentary time and pain in RA, and reported change in sedentary time was significantly positively associated with change in pain. Path models further showed sedentary time had a significant positive bi-directional relationship with pain. These findings indicate pain may be both a determinant and consequence of sedentary time in people with RA (O'Brien et al., 2021).

Sedentary Behaviour and Disease Activity

Research has suggested that links exist between SB with disease activity in people with RA. For example, Khoja et al. (2016) reported positive associations between accelerometer-assessed daily sedentary time with disease activity (assessed using DAS28) in correlation and regression analysis. O'Leary et al. (2021) also reported that device-measured sedentary time was inversely linked to

disease activity in participants with RA. Indeed, findings also indicated that disease activity was indirectly associated with sedentary time, mediated by pain intensity (O'Leary et al., 2021). However, due to the cross sectional nature of these studies, we cannot infer that associations refer to causation and therefore, SB may be a cause or a consequence of increased disease activity.

Longitudinal associations between SB and disease activity have been investigated by Prioreschi et al. (2014). Their findings included that, following DMARD therapy, there were improvements in disease activity and functional ability in parallel with decreases in accelerometer-measured sedentary time (Prioreschi et al., 2014).

Sedentary Behaviour and Functional Ability

Regarding the links between SB with functional ability, Greene et al. (2006) reported higher self-reported sedentary time was associated with functional disability. Another cross-sectional study by Khoja et al. (2016), demonstrated significant positive correlations between accelerometer-measured sedentary time with HAQ-measured functional ability ($r= .43$, $p<.05$). These associations remained in regression analysis, when models were adjusted for age and gender ($R^2\Delta= .16$, $p<.001$) (Khoja et al., 2016). Conversely, Fenton et al. (2017) found no relationship between ActiGraph GT3X-measured sedentary time with functional ability in correlation analysis and regression models. These contrasting findings may be due to these studies using different definitions of SB (≤1.5 METs by Fenton et al. (2017) and <1 METs by Khoja et al. (2016)). In addition, TV viewing, a particularly deleterious type of SB, has shown negative associations with functional ability, whereby increasing TV viewing by 1 hour per day was related to a 0.09 unit increase in the HAQ score (Giles et al., 2008; Greene et al., 2006).

Sedentary Behaviour and Fatigue

Despite the fact that fatigue is one of the major reported symptoms by people with RA, few studies have been conducted investigating the relationships between SB with fatigue, and findings have

been varied. O'Leary et al. (2021) found no association between activPAL-assessed sedentary time with overall fatigue severity in patients with RA. Nevertheless, O'Brien et al. (2021) conducted a longitudinal study assessing relationships between SB with fatigue in people with RA. Path analysis models showed a bi-directional relationship between changes in activPAL-measured sedentary time with change in fatigue. Therefore, results indicate variability in fatigue may be a determinant and consequence of sedentary time in the RA population (O'Brien et al., 2021).

Sedentary Behaviour and Mental Health and Psychological Wellbeing

The link between SB with depression and anxiety has been investigated by O'Leary et al. (2021), who found significant correlations between activPAL-measured daily sedentary time with depression ($r=.28$, $p<.05$) and anxiety ($r=.31$, $p<.05$) sub-scores of HADS. To my knowledge, no comparable studies exploring links between SB and indicators of mental health and wellbeing have been conducted in people with RA.

Sedentary Behaviour and Health in Rheumatoid Arthritis: Limitations and Conclusions

As with PA, the links between SB with core OMERACT health outcomes in people with RA are varied and inconsistent between outcomes. Relative to PA, SB research in RA is still in its infancy, with a limited number of large-scale cross-sectional studies conducted investigating the different dimensions and elements of SB (i.e., SITT). Further cross-sectional studies are needed, investigating SB as an independent behaviour to physical inactivity, to investigate the extent of the reported associations, to confirm findings and determine the direction of relationships. Through using device-based and reliable measurement tools, this may enable the assessment of different dimensions and elements (i.e., sedentary bouts, interruptions, sedentary time and context) of SB (i.e., SITT) and outcomes, in order to further understand the relationships.

To my knowledge, only 1 SB intervention has been conducted in people with RA. The 16 week randomised controlled trial involved motivational counselling sessions and text messages sent to

participants with the aim to reduce SB (Thomsen et al., 2017). Results demonstrated significant post-intervention reductions in activPAL-measured daily sitting time, increased standing and stepping time, reduced pain (measured via VAS), functional ability (measured via the HAQ) and fatigue (measured by the MFI), and increased quality of life (measured using SF-36) (Thomsen et al., 2017). Effects were sustained at 18 month follow-up assessment for all these outcomes (Thomsen et al., 2020). These results were comparable to those of PA interventions in people with RA, and indicate that there may be substantial health benefits accrued from and sustained by replacing sitting time with standing and/or stepping (Thomsen et al., 2020).

Nonetheless, this intervention had its limitations. For example, there were significant baseline differences between the intervention and control group in daily sitting time, pain and fatigue. The intervention group displayed significantly higher pain and fatigue, and more time spent sitting than the control group. Therefore, results for these outcomes in particular should be interpreted with caution. In addition, the majority of outcome measures were self-reported, and did not assess multidimensional aspects of outcomes. For example, a VAS was used to assess pain, which only measures overall pain intensity. Perhaps by using QST to quantify multiple mechanisms involved in RA-related pain in future SB interventions, it may give a more comprehensive overview of the multidimensionality of pain processing present in people with RA, in relation to their SB. Furthermore, this intervention did not assess positive and negative aspects of mental health and wellbeing, such as depression, anxiety and subjective vitality. Future interventions should seek to include these as separate and distinct outcomes, due to their potential for having differing relationships with SB in people with RA. Finally, this intervention was perhaps a little premature, as research establishing links between the different SITT elements and dimensions of SB with OMERACT health outcomes in people with RA is still in its infancy. Thus, only once there is a solid evidence base establishing the role of SB for different core OMERACT outcomes, should SB interventions be conducted targeting those outcomes identified as being likely to benefit from such interventions.

Challenges to Reducing Sedentary Behaviour in Rheumatoid Arthritis

Qualitative research has indicated that people with RA require significant prioritisation, planning and self-management in order to minimise their time spent in SB (Thomsen et al., 2015). In addition, due to constant disease fluctuations and periods where patients experience significant pain and fatigue, there are times whereby SB may predominate life. Therefore, although SB may benefit health in people with RA, and many of the RA population are aware of this, the additional barriers they face mean that they require an individually tailored approach to encourage sustainable long-term change in SB (Thomsen et al., 2015).

Research to date indicates that high levels of SB in people with RA may contribute towards the high disease burden (Fenton et al., 2018a). However further, more rigorous research is required to establish the link between SB and core OMERACT patient- and clinician-important outcomes in RA. For example, past research has primarily been conducted in small cohorts, assessing cross-sectional associations between SB with RA outcomes. A limited number of longitudinal studies and interventions targeting SB have been conducted in RA participants. Although O'Brien et al. (2021) demonstrated longitudinal bi-directional relationships between change in sedentary time with pain and fatigue in people with RA, investigation into other factors and outcomes that may be influenced by changes in SB is required.

Research investigating the relationships between different elements and dimensions of PA and SB (i.e., FITT and SITT) with core OMERACT patient- and clinician-important health outcomes is lacking. Particularly important is the need for high-quality studies using reliable, device-based measures of behaviours and multidimensional and reliable measures of outcomes. Only once these high-quality observational studies have been conducted can we identify the modifiable determinants, and inform which outcomes and behaviours to focus on in future PA and SB interventions.

Contribution of thesis and aims: pre-COVID-19

Prior to COVID-19, the primary aims of this thesis were to:

1. Evaluate the current evidence regarding the efficacy of existing lifestyle PA and SB interventions at improving core OMERACT patient- and clinician-important outcomes in people with RA, through conducting a systematic review and meta-analysis ***(Chapter 2)***.
2. Advance measurement of pain through psychophysical multidimensional measurement methods, by examining the inter-rater and test-retest reliability of QST modalities for assessing pain in populations of healthy participants, people with LBP and RA patients ***(Chapter 3)***.
3. Assess the comparability of different ActiGraph models, cut-points and placement sites for assessing free-living PA, and determine the validity of these models, cut-points and placement sites at quantifying free-living sedentary time compared to the activPAL ***(Chapter 4)***.
4. Use the tools (i.e., QST) and devices (i.e., accelerometers, activPAL) tested in these studies to inform the methods and design of a large-scale longitudinal study assessing relationships between different intensities and types of PA (i.e., LPA, MVPA, lifestyle PA) and SB with OMERACT patient- and clinician-important health outcomes in people with RA ***(Chapter 5)***. As part of this study, data-prompted interviews with RA patients were intended to provide novel qualitative information about determinants, barriers and facilitators to PA and SB, to inform intervention design.
5. To use the quantitative and qualitative data from the longitudinal study, to design and compare the initial feasibility of two interventions targeting: 1) increasing MVPA through structured exercise vs 2) increasing total PA through increasing lifestyle PA and reducing SB.

Methodological aims

Both reliability and validity are important aspects in order to quantitatively assess the rigour and quality of research or measures (Heale & Twycross, 2015; Kimberlin & Winterstein, 2008). Aims 2 and 3 of this thesis focus on advancing the measurement methods used in research of pain and PA and SB. As part of these aims, in ***Chapters 3 and 4***, I aim to assess the reliability and validity of new and existing measures, so they can be subsequently used in other studies.

Validity

Validity can be defined as the extent to which a measurement tool can accurately measure the concept of interest (Heale & Twycross, 2015). There are 4 main types of validity and these are content (i.e., does the tool accurately measure all aspects of a construct), construct (i.e., does the tool measure what it's intended to measure) and criterion (i.e., is the tool related to other instruments that measure the same outcome) validity. A subset of content validity is face validity, which refers to whether the method or tool used looks like it measures the concept intended (Heale & Twycross, 2015).

Reliability

Reliability describes the extent to which a research tool consistently measures the same results when used on repeated occasions in the same environment (Heale & Twycross, 2015), i.e., the stability of a measure. Key attributes of reliability are homogeneity, stability and equivalence. Homogeneity (i.e., internal consistency) is measured using Cronbach α. Stability can be tested using test-retest reliability (level of agreement between measures when they are taken at a different point in time), and equivalence can be assessed through inter-rater reliability (level of agreement between 2 independent observers). Together these can provide an indication of the reliability of a measure (Heale & Twycross, 2015).

Depending on the outcome measures and populations, different types of validity and reliability assessments are warranted. For example, in this thesis, I will explore the test-retest and inter-rater reliability of some QST modalities. In addition, I will investigate the inter-monitor reliability of some accelerometers worn at different attachment sites, with different cut-points applied for measurement of free-living PA. Regarding validity, the face validity of QST measures and criterion validity of accelerometer cut-points compared to the activPAL for measurement of free-living sedentary time will be explored within this thesis. Once appropriate reliability and validity have been established, it can be decided if these measures (accelerometers and QST modalities) can be implemented in subsequent research.

COVID-19 pandemic

In January 2020, the WHO declared SARS-CoV-2 (known as COVID-19) a public health emergency and by March 11, 2020 the outbreak was further declared a pandemic. Since the first confirmed case, there have been over 500 million COVID-19 infections, and over 6 million deaths worldwide (as of May 2022) (WHO, 2022). Unprecedented nationwide restrictions were enforced, with the UK initially going into national lockdown on March 23, 2020. Restrictions required people to only leave their homes for some basic necessities (i.e., medical treatment, food shopping), essential work that could not be carried out at home, and for daily exercise. COVID-19 restrictions in the UK were ongoing throughout 2020 and 2021. As a consequence of this, all NHS-based non-COVID clinical research was immediately halted in March 2020. As a result, it was not possible to achieve thesis aims 4 and 5 as set out above.

Nevertheless, the COVID-19 pandemic and UK lockdown was a unique worldwide event which had a significant impact on behaviours of daily life, including peoples engagement in PA and SB. Therefore, the initial aims of this thesis were revised as described below.

Contribution of thesis and aims: post-COVID-19

Early research conducted during the COVID-19 pandemic found poor mental health and psychological wellbeing reported in the general population (Harper et al., 2020; Huang & Zhao, 2020; Lai et al., 2020; Rodriguez-Rey et al., 2020). The pandemic displayed particularly deleterious effects in people with rheumatic diseases (Michaud et al., 2020), with worsened mental health present in 73% of patients (Ziade et al., 2020). However, little was known about the relative risk of COVID-19 infection on people with auto-immune diseases, such as RA. In addition as people with RA already report high levels of poor mental health (Fiest et al., 2017), they may have been at higher risk of adverse mental health and psychological wellbeing during the pandemic (Veldhuijzen Van Zanten et al., 2020). The OMERACT patient- and clinician-important outcomes highlighted previously all had the potential to be impacted further during the pandemic.

As COVID-19 lockdowns restricted opportunities to be active, large-scale population-based studies reported less PA engagement and greater screen time in in healthy adults (Castañeda-Babarro et al., 2020; Meyer et al., 2020; Pépin et al., 2020). Reduced PA and increased prolonged sitting were also reported in people with RA (Pinto et al., 2020b), with another study observing significantly lower self-reported PA participation (measured via IPAQ) in people with RA compared to healthy controls (1160 minutes/week for RA and 2940 minutes/week for non-RA, $p<.001$) (Balchin et al., 2021). Barriers to PA participation included limited access to gym facilities and equipment, and this reduced PA was accompanied by low mental wellbeing in participants with RA only (Balchin et al., 2021). As research into PA and SB emerged during the COVID-19 pandemic, studies showed links between PA and SB with health in some populations (Maugeri et al., 2020; Puccinelli et al., 2021; Robinson et al., 2021). In response to the COVID-19 pandemic, the amended final aim of this thesis (in place of aims 4 and 5 above) was:

6. To conduct a large-scale online study with the objective to investigate associations between different types and intensities of PA (non-exercise light intensity PA, exercise and walking) and SB with core OMERACT patient- and clinician-important health outcomes in people with RA during COVID-19 (***Chapter 6***).

CHAPTER 2: EFFECTS OF LIFESTYLE PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR INTERVENTIONS ON HEALTH OUTCOMES IN RHEUMATOID ARTHRITIS: A SYSTEMATIC REVIEW WITH META-ANALYSIS

Abstract

Lifestyle physical activity (PA) is defined as increasing any type of PA as part of daily life and can include engagement in activities of daily living, incidental PA, walking or reducing sedentary behaviours (SB). People with Rheumatoid Arthritis (RA) experience impaired disease activity, leading to worsened health. PA is recommended to improve disease activity and other health outcomes in this population. The aims of this systematic review and meta-analysis were to evaluate the effectiveness of interventions targeting lifestyle PA and/or SB on 1) disease activity; 2) PA, SB and other core patient- and clinician-important outcomes in people with RA. Risk of bias for each intervention was assessed. Eight databases were searched from inception until June 2021. Inclusion criteria included lifestyle PA and/or SB interventions conducted in adults with RA. Of 880 relevant articles, 16 lifestyle PA and/or SB interventions met inclusion criteria. Meta-analyses showed significant effects of lifestyle PA interventions on disease activity, moderate to vigorous PA, light/leisure PA, steps, functional ability, depression and fatigue. There were significant positive SB intervention effects on sedentary time, leisure/light intensity PA, functional ability, pain, fatigue, and quality of life. Most interventions displaying improvements in lifestyle PA and SB also had improvements in other secondary health outcomes. This is evidence that lifestyle PA and SB interventions are promising approaches to increase PA, reduce SB and improve patient- and clinician-important outcomes in people with RA. More high-quality SB interventions are needed to determine their effectiveness at producing clinical benefits in these health outcomes.

Introduction

Rheumatoid Arthritis (RA) is a chronic inflammatory autoimmune condition, characterised by high levels of pain and fatigue (Lee & Weinblatt, 2001; Smolen et al., 2016). Consequently, people with RA frequently report low levels of physical activity (PA), with a significant proportion of daily life engaged in sedentary behaviours (SB) (Sokka et al., 2008; Tierney et al., 2012; Yu et al., 2015). SB is defined as any waking activity expending energy ≤1.5 metabolic equivalents (METs) whilst in a sitting/reclining/lying posture (Sedentary Behaviour Research Network, 2012). In people with RA, participating in PA has shown improvements in disease activity, inflammation, functional ability, pain, fatigue, depression and anxiety (Kelley et al., 2015; Metsios et al., 2020; Metsios et al., 2015; Nordgren et al., 2015; Rongen-van Dartel et al., 2015). Therefore, regular PA, as well as self-management, is recommended as a non-pharmacological approach in RA (Nikiphorou et al., 2021). Furthermore, recent evidence has shown that high levels of SB are independently related to increased disease activity, reduced functional ability and pain in people with RA (Fenton et al., 2018a; Khoja et al., 2016; O'Leary et al., 2021). Together, the independent health benefits of PA and SB emphasise the need for behavioural interventions to encourage PA and/or reduce SB in people with RA.

Previously, the most commonplace non-pharmacological interventions in RA involved structured, supervised, and purposeful exercise, targeting moderate to vigorous PA (MVPA) (i.e., behaviour ≥3 METs) (Fenton et al., 2020a; Metsios et al., 2015). Despite the well-known benefits of MVPA, exercise training can be difficult for people with RA, especially in those with high disease activity (Khoja et al., 2016), who experience additional barriers to being active (Veldhuijzen Van Zanten et al., 2015). Consequently, alternative therapeutic approaches that focus on increasing overall PA, through incorporating more PA into an individual's daily lifestyle, are increasingly advocated (Swardh et al.,

2020). This approach of increasing "lifestyle PA", may be perceived as more feasible, achievable, and sustainable for people with RA (Brady et al., 2021).

Although there is no formal definition for lifestyle PA, it comprises increasing any type of PA as part of day-to-day life. This can include increasing engagement in activities of daily living (e.g., gardening, housework, walking to work), increasing incidental PA (i.e., PA built up in small amounts over the day), as well as increasing engagement in activities such as walking, which may be higher-intensity in nature (e.g., walking at a moderate to vigorous intensity) (Katz et al., 2018). Reducing SB is also an avenue to increasing lifestyle PA, as sitting less will assist in increasing an individual's total daily PA, irrespective of intensity (Thomsen et al., 2017). In healthy individuals and amongst those living with other musculoskeletal conditions, emerging evidence has suggested that engagement in lifestyle PA is a clinically meaningful and cost-effective strategy to increase PA and improve health outcomes, with good compliance and high acceptability (Chmelo et al., 2013; Duvivier et al., 2013; Giraudet-le Quintrec et al., 2007; Thomsen et al., 2017; Van Roie et al., 2010).

There is little summarised information regarding the effectiveness of lifestyle PA and SB interventions in people with RA, particularly related to improving core patient- and clinician-important outcomes (i.e., outlined by Outcome Measures in Rheumatoid Arthritis Clinical Trials, OMERACT), and particularly disease activity. Disease activity is associated with disease progression, severity, hospitalisation and comorbidities in RA (Metsios et al., 2015; Metsios et al., 2011). There is substantial evidence that exercise interventions can improve disease activity (Rausch-Osthoff et al., 2018). However, to my knowledge, no systematic review has assessed the effectiveness of lifestyle PA and SB interventions at improving disease activity in the RA population. To understand the value of lifestyle interventions to promote PA or reduce SB for improving health outcomes in RA, it is important to examine and appraise the current evidence. The aim of this systematic review and meta-analysis was to evaluate the effectiveness of lifestyle PA and SB (both individually and

collectively) interventions on disease activity, and other core OMERACT patient- and clinician-important outcomes in people with RA (Bartlett et al., 2015; Boers et al., 2014). Additional aims included assessing the quality of these existing interventions, using risk of bias and quality assessments.

Methods

Registration

This systematic review was registered in the International Prospective Register of Systematic Review database (PROSPERO, CRD42020149345) and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) was followed (Moher et al., 2015).

Electronic Data Sources and Literature Searches

Following PRISMA guidelines (Moher et al., 2015) and the Cochrane Handbook (Higgins et al., 2021b), a literature search strategy was designed, through consultations with research librarians and members of the research team. The PICO method was used to assist search strategy creation, and the search strategy was adapted for each database.

Eight databases [Medline, Cochrane Library CENTRAL, Web of Science, PsychINFO, Cumulative Index to Nursing & Allied Health Literature (CINAHL), Scopus, Excerpta Medica database (EMBASE) and Physiotherapy Evidence Database (PEDro)] were searched from inception to June 2021 to identify relevant publications. The search algorithms used in each database can be found in **Supplementary Table 2.1**.

Study Selection and Inclusion Criteria

In order to be considered for inclusion, studies needed to: 1) be conducted in adults (≥18 years) with self- or physician-diagnosis of RA; 2) include an intervention which is directly or indirectly targeting

lifestyle PA and/or SB; and 3) include assessments of core patient- (i.e., functional ability, pain, fatigue, depression, anxiety, vitality, quality of life) and/or clinician- (i.e., disease activity, functional ability) important outcomes, as defined by OMERACT (Bartlett et al., 2012; Boers et al., 1994; Van Tuyl & Boers, 2015). Selection of eligible interventions was based on the nature of outcome measure (i.e., an outcome measure quantifying lifestyle PA participation), rather than the nature of the intervention. For example, if an exercise intervention had an outcome measure of total PA engagement, this indicates that an aim of the intervention was to explore the effectiveness of the intervention to change PA as part of daily life (i.e., promoting lifestyle PA), and therefore this intervention would be eligible for inclusion in this systematic review. Publications were also required to be in English, with no restrictions on length of follow-up or geographic location. Randomised controlled trials (RCTs), quasi-randomised and single-arm trials were included in this review. Studies involving participants with various diagnoses of arthritis, whereby results of RA participants could not be distinguished from other cohorts (e.g., osteoarthritis), were excluded. Additional exclusion criteria were if interventions were multicomponent, and the effects of lifestyle PA and/or SB could not be separated from other components (e.g., cognitive behavioural therapy, education, and diet). Protocols, review articles, poster presentations and conference proceedings were also excluded.

The primary outcome in this review was disease activity, as it is a OMERACT patient- and clinician-important outcome, a key clinical target for treatment and management of RA, and a predictor of health, disease severity and hospitalisation (Arts et al., 2015; Metsios et al., 2015; Metsios et al., 2011). Secondary outcomes consisted of lifestyle PA and SB (including, total PA, steps, MVPA, and leisure/light intensity PA and sedentary time) and other core OMERACT outcomes relevant to RA (pain, functional ability, fatigue, anxiety, depression and quality of life) (Bartlett et al., 2015; Boers et al., 2014).

Data Extraction and Risk of Bias Assessment

Data were extracted from all included studies. Details of each study were collected and characterised by author, date of publication, sample size, participant characteristics (i.e., age, gender, disease duration, and disease activity), intervention characteristics (i.e., description of intervention, assessment timepoints and intervention length), methods of outcome assessment and results.

Study risk of bias was appraised using the Cochrane Risk of Bias 2 (RoB2) tool for randomised trials. The National Institute of health (NIH) National Heart Lung and Blood Institute study quality assessment tool for before-after (pre-post) studies with no control group, was used where intervention studies: 1) had no control group (i.e., single-arm trials) [n=2], or 2) did not measure any of the primary or secondary outcomes of this review [n=2] (Higgins et al., 2021a; National Heart, 2019; Sterne et al., 2019). Two reviewers independently graded the risk of bias for each study, and any inconsistencies were discussed, and resolved with an additional third reviewer, if required. The RoB2 was individually scored for 5 domains, as outlined in **Figures 2.3a, 2.3b and 2.4**. To assess the outcome bias domain, we used the most consistently reported outcomes across studies (disease activity and functional ability) (Higgins et al., 2021a; Sterne et al., 2019). An overall risk of bias was calculated, as “low risk”, “some concerns” or “high risk”, for each study. For the four studies which we used the NIH tool, overall risk of bias was assessed by answering 12 questions, and studies were scored as “good”, “fair” or “poor” (National Heart, 2019).

Quality of evidence was assessed using Grading of Recommendations Assessment Development and Evaluation (GRADE) analysis, with overall GRADE quality of evidence rated as high, moderate, low or very low quality (**Table 2.2**).

Data Synthesis and Analysis

For studies that provided suitable data for a meta-analysis, we extracted and collated data into relevant outcomes. Where similar outcomes measures were assessed in different studies, these were grouped appropriately using continuous, inverse variance, random effects models. Where data was not reported by studies, efforts were made to contact authors [n=10] to obtain additional data (i.e., e-mails sent, with follow up 2 weeks later), and if data could still not be obtained, reviewers imputed means and standard deviations [for n=5 interventions], where possible, using the Cochrane Handbook recommended methods (Higgins et al., 2021b).

Mean differences (MD) (for outcomes containing studies that used the same measurement scales) and standardised mean differences (SMD) (for outcomes containing studies that used different measurement scales) were tested between experimental groups and control groups (or pre- and post-intervention data, n=2 single-arm studies (Cramp et al., 2020; Nordgren et al., 2015)). As some studies only reported non-normally distributed data for each outcome, normally distributed values were logarithmically transformed to non-normal values, so all studies included in one outcome meta-analysis contained non- normally distributed data (Feng et al., 2014; Higgins et al., 2008). Where this was not possible (for functional ability and depression outcomes), normal and non-normally distributed data were analysed separately. We evaluated the 95% confidence intervals (CI) and heterogeneity between studies using the I^2 statistic, which indicates the variability of the intervention effect due to heterogeneity. A result was considered statistically significant if $p<.05$, and interpretation of I^2 value was made based on Cochrane recommendations, whereby, 0−40%= not important; 30−60%= moderate heterogeneity; 50−90%= substantial heterogeneity; and 75−100%= considerable heterogeneity (Higgins et al., 2021b). Review Manager 5.4.1 was used to conduct meta-analyses. Subgroup analysis was conducted to compare the similarity of findings between different types of interventions. Subgroup analysis focused on 1) target of intervention, i.e., intervention

primarily targeting PA or SB, and 2) outcome assessment timepoint, i.e., during/immediately post-intervention or follow-up.

Results

Searching and Selection Procedure Results

The search procedure is described **Figure 2.1** (PRISMA flowchart). Initial database searches identified 1156 relevant articles, with a total of 880 articles when duplicates (n= 276) were removed. Titles and abstracts for all articles were retrieved and reviewed by two independent reviewers. Where title and abstract did not provide sufficient information regarding the intervention, full texts were examined. Reference lists of included articles were manually examined to supplement searches and identify further relevant studies. Full texts (n=120) were retained and reviewed against inclusion and exclusion criteria by 2 independent reviewers. Where reviewers disagreed, discrepancies were discussed and a third reviewer was involved to make final inclusion decisions. In total, 17 studies provided sufficient data to be included in this systematic review and meta-analysis, with 1 study providing insufficient information (Van Den Berg et al., 2006).

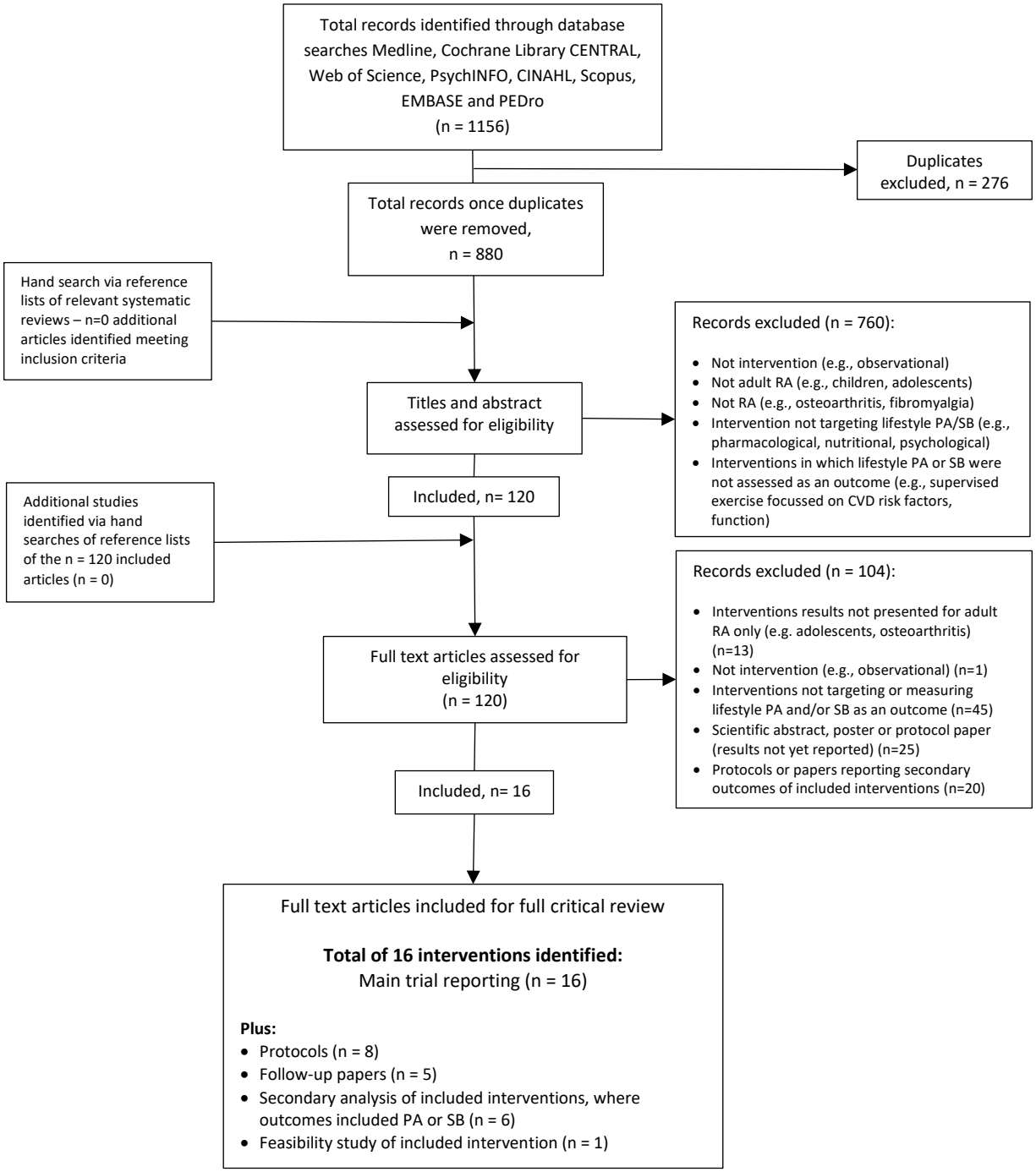


Figure 2.1: PRISMA diagram of the literature search results.

Note: PA= Physical Activity, SB= Sedentary Behaviour, CVD= Cardiovascular Disease, CINAHL= Cumulative Index to Nursing & Allied Health Literature, EMBASE= Excerpta Medica database, PEDro= Physiotherapy Evidence Database, PRISMA= Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Table 2.1: Summary of Findings

Author, year and Country of publication	***Characteristic:*** *Sample size (n), Age (M ± SD), Gender (% female)*	***IG and CG: design and content***	***Duration & timepoints***	***Assessment of PA and/or SB***	***Disease activity Results***	***PA/SB Results***	***Secondary Outcomes Results***
Brodin et al., 2008 Sweden	IG: 94 54 ± 14.0 72 CG: 134 56 ± 13.9 75	**IG:** Individual coaching program aimed to implement healthy PA. Telephone support given after 1 week, then monthly. 3 monthly function tests to support adherence **CG:** Ordinary physical therapy	1 year baseline DI: 3 months DI: 6 months DI: 9 months PI: 1 year FU: NR	SR: 3 questions- intensity of low, moderate and high intensity PA not validated in RA	DAS28 (ESR): IG: ~ , CG: ~ BGD not assessed	IG: n=26 (34%) ↑, n=19 (20%) ↓ CG: n=23 (20%) ↑, n=31 (23%) ↓ no BGD in number increasing PA	FA (HAQ): IG: ~ , CG: ~ , no BGD QoL: IG: ↑* , CG: ~ , sig. BGD Pain: IG: ~ , CG: ~ , no BGD
Feldthusen et al., 2016 Sweden	IG: 36 54.2 ± 8.5 88.9 CG: 34 52.7 ± 10.9 88.2	**IG:** Develop self-care plan focussing on tailoring health enhancing PA (reaching adult PA guidelines- i.e., aerobic MPA>30min, 5d/week; aerobic VPA >20min, 3d/week; combination of the 2)) and balancing life activities Follow-up support meetings and telephone calls conducted by specialised physical therapists. Frequency of follow-up was individualised. **CG**: Usual care and activities only	12 weeks baseline PI: 12 week FU: 6 months	SR: LTPAI not validated in RA	DAS28 (ESR): IG: ↓ (at post-test and follow-up) , CG: ~ , no BGD	LTPAI: IG:↑ , CG: ~ sig. BGD between at PI and FU	Fatigue: IG: ↓* , CG: ↓* at PI and FU, no BGD Pain: IG: ~ , CG: ~ , no BGD Anxiety: IG: ↓* , CG: ~ , sig. BGD at PI and FU Depression: IG: ↓* , CG: ~ , no BGD QoL: IG: ↑* , CG: ~ , sig. BGD at FU
Gilbert et al., 2018 USA	IG: 93 55.0 ± 13.8 82.8 CG: 92 54.7 ± 13.7 84.8	**IG:** Minimum 3-monthly motivational interviews with HCP (in person/telephone)- , individual goal setting, tailored strategies for increasing PA and monitoring progress Progress evaluated in subsequent interviews and further goals set **CG**: Brief PA counselling - physician advice only	24 months baseline DI: 3 months DI: 6 months DI: 12 months PI: 24 months FU: NR	DB: GT1M ActiGraph SR: Yale physical activity scale		Total PA (mins/day): IG: ~ , CG: ~ , no BGD MVPA (mins/day): IG: ~ , CG: ~ , no BGD	FA (HAQ): IG: ~ , CG: ~ , no BGD QoL- Physical: IG: ~ , CG: ~ , no BGD QoL- Mental: IG: ~ , CG: ~ , sig. BGD at follow-up Pain: IG: ~ , CG: ~ , no BGD

Knittle et al., 2015 Netherlands	IG: 38 60.7 ± 11.9 79* CG: 40 64.7 ± 11.5 55*	**IG:** Small group patient education sessions delivered by physical therapist- and one to one motivational interviews and self-regulation coaching FU telephone self-regulation coaching sessions **CG:** Group based patient education session	5 weeks baseline PI: 6 weeks FU: 32 weeks	SR: SQUASH	RADAI: IG: ~ , CG: ~ , sig. BGD at FU in favour of CG	Leisure time PA (mins/week): IG: ↑ , CG: ~ , sig. BGD at FU Number active days (days/week): IG: ↑ , CG: ~ , sig. BGD at PI and FU	FA (HAQ)): IG: ~ , CG: ~ , no BGD Depression: IG: ~ , CG: ~ , no BGD Fatigue: IG: ~ , CG: ~ , no BGD
Giraudet-Le Quintrec et al., 2008 France	IG: 104 55.3 ± 11.8 86.4 CG: 104 54.3 ± 14.4 85.4	**IG:** multidisciplinary educational intervention- home based exercise prescription and recommendations for leisure PA 8 group weekly face to face, 5-hour education program sessions on RA management and physical program, OT, physical therapist, aquatic or relaxation training **CG:** Usual medical care and information booklets with PA recommendations and exercises	12 months baseline DI: 6 months PI: 12 months FU: NR	SR: Baeke questionnaire (assessed leisure time PA (sports + hobbies)) not validated in RA	DAS28 : IG: ~ , CG: ~ , no BGD	Leisure PA score: IG: ↓ , CG: ↓ , no BGD	FA (HAQ): IG: ~ , CG: ~ , no BGD Anxiety: IG: ~ , CG: ~ , no BGD Depression: IG: ~ , CG: ~ , no BGD QoL: IG: ~ , CG: ~ , no BGD Fatigue: IG: ~ , CG: ~ , no BGD
Thomsen et al., 2017 Denmark	IG: 75 59.7 ± 10.7 81 CG: 75 59.5 ± 12.7 80	**IG:** 1: 3x individual motivational counselling sessions - individual goal setting and self-efficacy, set behavioural goals to reduce sitting, motivation and confidence to encourage behaviour change. Booklets given containing key messages 2: SMS reminders- based on goals (frequency is individualised) **CG:** Current lifestyle	16 weeks baseline PI: 16 weeks- FU: 6 months FU: 22 months	DB: activPAL™ SR: PAS 2.1	DAS28 (CRP): IG: ↓ , CG: ↓ , no BGD (assessed at FU only)	DB sitting time (hr/day): IG: ↓ , CG: ↑ , sig. BGD at PI and FU DB standing time (hr/day): IG: ↑ , CG: ↓ , sig. BGD at PI and FU DB stepping time (hr/day): IG: ↑ , CG: ↓ , sig. BGD at PI and FU SR sitting at work (hr/day): IG: ↓ , CG: ~ , sig. BGD at PI and FU SR sitting in leisure (hr/day): IG: ↓ , CG: ↑ , sig. BGD at PI and FU	FA (HAQ): IG: ↓* , CG: ↑* , sig. BGD at post-test and follow-up QoL: IG: ↑* , CG: ↓* , sig. BGD at post-test and follow-up Pain: IG: ↓*, CG: ↑* , sig. BGD at post-test and follow-up Fatigue: IG: ↓*, CG: ↑* , sig. BGD at post-test and follow-up

Van den Berg et al., 2006 Netherlands	IG: 82 49.5 (12.9) median (IQR) 76 CG: 78 49.8 (13.9) median (IQR) 77	**IG:** Internet based PA programme (performed 5x/week)- Individual PA guidance, bicycle ergometer. Participants advised to do other forms of PA as well. Weekly email supervision with physical therapist 3-monthly group meetings - demonstrated new exercises, exchange of experiences. Tailored self-management strategies addressed during meeting **CG:** Internet based general PA training advice	12 months baseline DI: 3 months DI: 6 months DI: 9 months PI: 12 months FU: NR	SR: Questionnaire (number meeting MPA and VPA recommendations) DB: Actilog 3	DAS28 (ESR): IG: ↓ , CG: ↓ , no BGD	MPA: IG: ↑ , CG: ↑ , sig. BGD at 6 and 9 months VPA: IG: ↑ , CG: ↑ , sig. BGD at 6, 9 and 12 months DB PA score: IG: ↓ , CG: ↓ (at 6 months), no BGD DB Peak amplitude: IG: ~ , CG: ~ , no BGD DB No. peaks: IG: ~ , CG: ~ , no BGD	FA (HAQ): IG: ↓* , CG: ~ , sig. BGD at 12 months only QoL : IG: ↑* , CG: ↑* , sig. BGD at 9 and 12 months
Veldhuijzen et al., 2021 England	IG: 43 55.4 ± 12.1 63 CG: 45 54.5 ± 13.0 69	**IG:** 3-month exercise program and SDT-based psychological intervention One to one consultations with BC counsellor: to support autonomous motivation for PA RA tailored exercise program: 3x30min/wk independent exercise sessions at gym (x2) and home (x1), semi-supervised **CG:** RA tailored exercise program	3 months baseline PI: 3 months FU: 6 months FU: 12 months	SR: IPAQ	DAS28: IG: ~ , CG: ~ , no BGD	MVPA (mins/week): IG: ~ , CG: ↓ , sig. BGD at 3, 6 and 12 months	FA (HAQ): IG: ↓* , CG: ↑*, sig. BGD at 6 and 12 months QoL: IG: ~ , CG: ~ , no BGD Depression: IG: ~ , CG: ~ , no BGD Anxiety: IG: ~ , CG: ~ , no BGD Fatigue: IG: ~ , CG: ~ , no BGD
Li et al., 2020 Canada	IG: 43 54.8 ± 15.4 88.4 CG: 43 55.3 ± 11.5 93	**IG:** 1. in person group education session and individual counselling. 2. Wear Fitbit Flex 2 and given PA goals 3. biweekly phone calls from physical therapist trained in motivational interviewing- reviewed PA goals **CG**: Routine activities weeks 1-9, did intervention weeks 10-18 (delay group)	8 weeks baseline PI: Week 9 (post-test IG) PI: Week 18 (post-test CG) FU: Week 27	DB: Sensewear accelerometer		MVPA (mins/day): IG: ↑ , CG: ~ , no BGD Purposeful activity (mins): IG: ~ , CG: ~ , no BGD Steps (no./day): IG: ~ , CG: ~ , no BGD Sedentary time (mins): IG: ~ , CG: ~ , no BGD	Depression: IG: ~ , CG: ~ , no BGD Pain: IG: ↓* (9 weeks) , CG: ~ , sig. BGD at 9 weeks Fatigue: IG: ~ , CG: ~ , no BGD

Katz et al., 2018 USA	PED+: 34 50.2 ± 14.1 88.2 PED: 34 55.9 ± 12.4 88.2 CG: 28 59.1 ± 12.5* 85.7	**IG:** 1. PED+: individualized step-count goals + pedometer + step-monitoring diary: booklet and discussion, pedometer, step diary and individualised daily step targets. Follow-up- target review phone call every 2 weeks 2. PED: pedometer + diary, NO targets: booklet and discussion, pedometer and diary to record daily pedometer steps. Follow-up- step count recorded via phone call every 2 weeks **CG:** education booklet and discussion on PA benefits	21 weeks baseline DI: 10 weeks PI: 21 weeks FU: NR	DB: Jawbone pedometer DB: Fitbit	RADAI (1-10) : PED+: ↓ , PED: ↓ , CG: ↑ (at week 21) , sig. BGD (lower in PED and PED+ than CG)	Steps (no./day): PED+: ↑ , PED: ↑ , CG: ~ , sig. BGD (changes within PED and PED+ differed from CG) % sedentary participants: PED+: ↓ , PED: ↓ , CG: ↑, sig. BGD % achieving healthy PA: PED+: ↑ , PED: ↑ , CG: ~ , no BGD	FA (HAQ): PED+: ↓* , PED: ↓* , CG: ~ , sig. BGD in PED+ vs CG at 21 weeks Pain: PED+: ↓* , PED: ↓* , CG: ~ , no BGD Fatigue: PED+: ↓* , PED: ↓* . CG: ~ , no BGD Depression: PED+: ↓* , PED: ↓* , CG: ~ , no BGD
Nordgren et al., 2015 Sweden	IG: 220 59 ± 8.8 81	**IG:** Health enhancing PA **(**HEPA) programme 1. 30+mins MPA on most days- given pedometer and access to webpage for step registration to encourage daily PA 2. 2x circuit training sessions/week in gym 3. biweekly support group meetings by PTs Alternative types of HEPA were encouraged- competitions, monitor aerobic capacity, weekly texts Expert lectures **CG:** No control, single-arm trial	2 years baseline DI: 3 months DI: 6 months DI: 12 months FU: NR	SR: IPAQ-SF SR: modified ESAI		% meeting current HEPA: IG: ↑ (at 1 year), ↓ from year 1 to year 2 (82% to 75%) % maintained (>6 months) HEPA: IG: ↑ 0 to 37% (at 1 year), ↓ from year 1 to year 2 (841% to 27%)	FA (HAQ): IG: ↓* QoL: IG: ↑* Pain: IG: ↓* Fatigue: IG: ~
Lange et al., 2020 Sweden	IG: 24 73.5 ± 2.7 75.0 CG: 23 74.0 ± 2.1 78.3	**IG:** Moderate-high intensity, aerobic and resistance exercise with person-centred guidance 3 sessions/week tailored gym based exercise: semi-supervised. Home based exercise: LPA 5 days/week and home exercises 2x/week Telephone support 7 months post intervention **CG:** Encouraged to perform home-based light intensity exercise	20 weeks baseline FU: 4 years	SR: LTPAI SR: ESAI	DAS28 (ESR): IG: ~ , CG: ↑ , sig. BGD	LTPAI: IG: ↑ , CG: ~ , no BGD ESAI- current HEPA: IG: 33% , CG: 26% , no BGD ESAI- maintained HEPA: IG: 25% , CG: 17% , no BGD	FA (HAQ)I: IG: ~ , CG: ~ , no BGD QoL: IG: ~ , CG: ↓* , sig. BGD Pain: IG: ~ , CG: ↑* , no BGD Fatigue: IG: ~ , CG: ↑* , no BGD

John et al., 2012 England	IG: 52 62.2 ± 10.6 71 CG: 58 60.8 ± 10.7 74	**IG:** Cognitive behavioural education intervention 3x interactive small group meetings by HCPs The important role of lifestyle modifications discussed, and individuals challenged to (using probing behavioural techniques), and commit to, a specific behaviour change Weekly progress reviews encouraged to self-monitor **CG:** Information leaflet	8 weeks baseline PI: 8 weeks FU: 6 months	SR: IPAQ		MET PA (mins/week): no BGD (WGD not assessed)	
Garner et al., 2018 Canada	IG: 14 45 ± 10 93 CG: 14 49 ± 14 71	**IG:** Individualised counselling intervention on PA and dietary intake 3x individualized visits to review strategies on: 1. Nutrition: with dietician, food questionnaire, reviewed diet recommendations, asked questions about diet. 2. PA: with rheumatology PT. Reviewed current PA and fitness tests results, instructions on PA guidelines, exercises to improve fitness. **CG:** Standard care	6 months baseline PI: 6 months FU: NR	DB: Pedometer	DAS28: IG: ↓ , CG: ↓ , no BGD	Steps (no./week): IG: ↑ +9,583 steps , CG: ↑ +6,696 steps, no BGD	FA (HAQ): no within group data reported, no BGD
Cramp et al., 2020 England	IG: 12 58 (range: 23-79) 75	**IG:** 4x group sessions: set goals, autonomy support, facilitate relatedness, group discussion, action plans tailored, individualised, to promote intrinsic motivation, peer support, self-monitoring (daily diaries and pedometers to take home) incorporated to promote self-efficacy and BC. One to one session: individual support to meet specific needs- discussion of individual PA barriers, strategies to overcome these **CG**: No control, single-arm trial	12 weeks baseline PI: 12 weeks FU: NR	SR: IPAQ-SF		IPAQ PA: IG: ~ (1 = ↑ , 1 = ↓)	FA (modified HAQ): IG: 3 = ↑ QoL: IG: ~ Pain: IG: 4 = ↑ Fatigue: IG: 6 = ↑ (better) , 3 = ↓ (worse) , 2 = ~ didn't test for significance
McKenna et al., 2021 Ireland	IG: 10 58 ± 7.4 100 CG: 10 56 ± 7.9 100	**IG:** Walking based exercise intervention based on ACSM, WHO and EULAR guidelines Sessions increased in length, intensity and duration each week from 2 to 5 sessions by week 8. Incrementally longer walks and more challenging targets. Progress self-monitored. Unsupervised sessions performed at time and location of choice **CG:** verbal and written instructions about benefits of exercise in RA	8 weeks baseline PI: Week 9 FU: NR	DB: activPAL™	CDAI : IG: ↓ (-0.7) , CG: ↑ (+0.7) (didn't test for significance, BGD not assessed)	MVPA (mins/day): IG: ↑ , CG: ~ (BGD not assessed)	FA (HAQ): IG:↓ (-0.6), CG: ↑ (+0.14) QoL: IG: ↑ (+10.4) , CG: ↑ (+0.3) Pain: IG: ↓ , CG: ~ Fatigue: IG: ↓ (-11) , CG: ↑ (+1) (didn't test for significance, BGD not assessed)

Note: WGD= within group difference, BGD= between group difference, ~= no significant change, ↑= increase (not significant), ↓= decrease (not significant), ↑*= increase (significant), ↓*= decrease (significant).

USA= United States of America, M ±SD = mean ± standard deviation, IG= intervention group, CG= control group, DI= during intervention, PI= post-intervention, FU= follow-up, NR= not reported, DB= device-based, SR= self-report, PA= physical activity, MVPA= moderate to vigorous physical activity, MPA= moderate physical activity, VPA= vigorous physical activity, LTPAI= Leisure time PA index, PAS 2.1= Physical Activity Scale 2.1, HEPA= health enhancing physical activity SQUASH= Short Questionnaire to Assess Health-Enhancing Physical Activity, IPAQ-SF= International Physical Activity Questionnaire- short form, ESAI= Exercise Stage Assessment Instrument, DAS28= Disease activity score- 28, ESR= erythrocyte sedimentation rate, CRP= C-reactive protein, CDAI= clinical disease activity index, RADAI= Rheumatoid Arthritis Disease Activity Index, FA= functional ability, HAQ= health assessment questionnaire, QoL= quality of life.

Characteristics of Included Studies

This review describes 13 interventions targeting and assessing lifestyle PA only, 1 intervention with an exclusive focus on SB (Thomsen et al., 2017), and 2 interventions targeting both lifestyle PA and SB (Katz et al., 2018; Li et al., 2020). In total, 14 studies were RCTs, and 2 observational cohort interventions (i.e., single-arm trials, with no control group) (Cramp et al., 2020; Nordgren et al., 2015). A total of 12 studies were conducted in Europe, 2 studies in Canada, and 2 studies in the United States. Intervention duration varied from 5 weeks to 24 months, with an average length of approximately 6 months. Study inclusion criteria generally required participants to have established RA, with only 1 study conducted in newly diagnosed RA (Garner et al., 2018). Most participants had low disease activity and few/no severe disabilities. Further characteristics of the included studies can be found in **Table 2.1**.

Effect of Interventions

Primary Outcome

Measurement tools and intervention results regarding disease activity are reported in **Table 2.1**. In brief, disease activity was reported by 11 studies, with some heterogeneity in the measurement tools. In total, 8 studies used the disease activity score 28 (DAS28) (Prevoo et al., 1995), 2 used the Rheumatoid Arthritis Disease Activity Index (RADAI) (Fransen et al., 2000), and 1 used the Clinical Disease Activity Index (CDAI) (Smolen et al., 2003). All measures of disease activity were based on patient or clinician assessment, with only the DAS28 having a serological marker of inflammation included as an objective element. Only 2 interventions demonstrated significantly greater improvements in disease activity in the intervention group compared to the control group (Katz et al., 2018; Lange et al., 2020), with an additional 5 studies displaying intervention group improvements, but no between group differences (Feldthusen et al., 2016; Garner et al., 2018; McKenna et al., 2021; Thomsen et al., 2017; Van Den Berg et al., 2006).

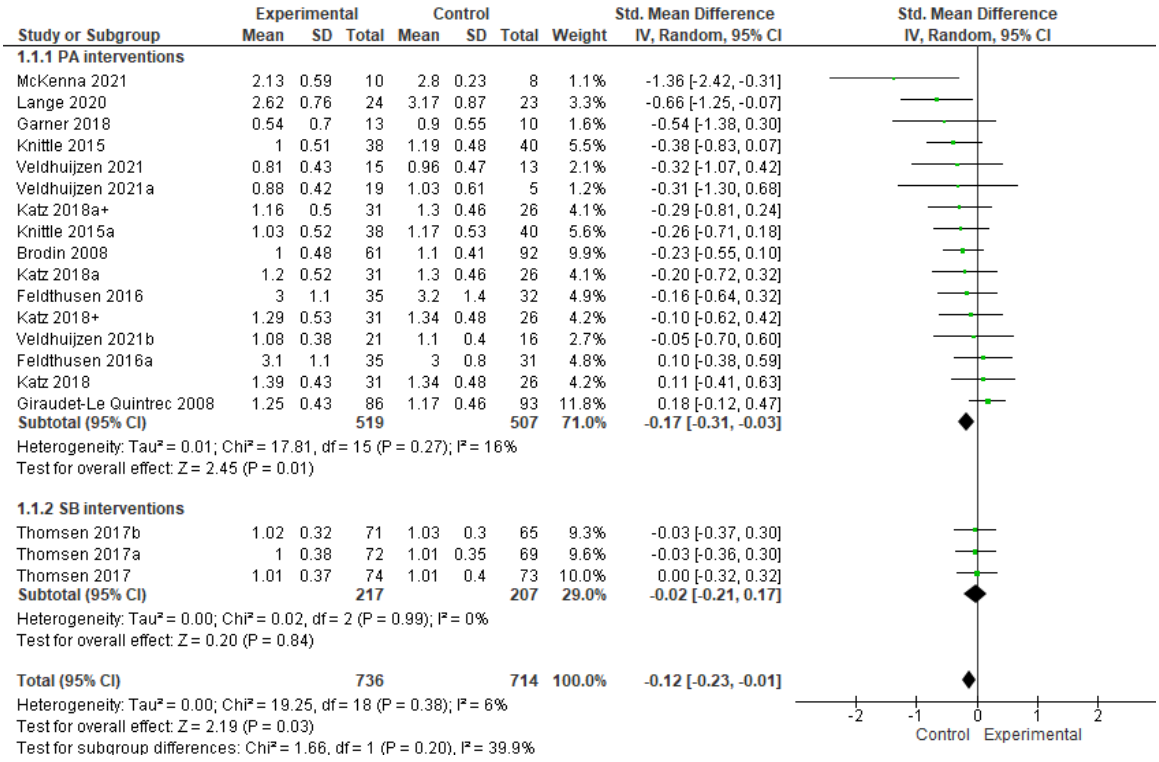


Figure 2.2a: The effects of lifestyle physical activity and sedentary behaviour interventions on disease activity in people with rheumatoid arthritis, with physical activity vs sedentary behaviour intervention subgroup analysis

Note: Where studies reported data from multiple post-intervention timepoints, these were included as separate studies in each meta-analysis (e.g., Thomsen 2017= 16-week timepoint, Thomsen 2017a= 10 month timepoint), SD= standard deviation, 95% CI= 95% confidence interval

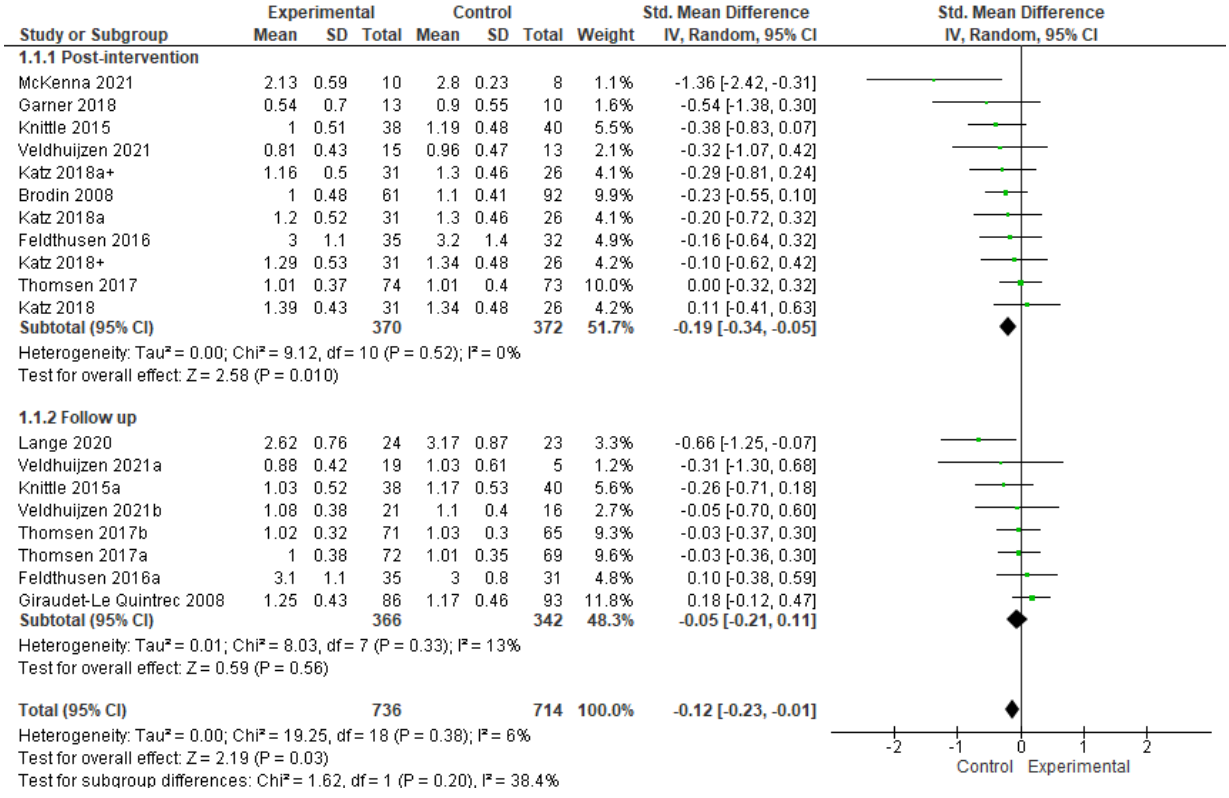


Figure 2.2b: The effects of lifestyle physical activity and sedentary behaviour interventions on disease activity in people with rheumatoid arthritis with post-intervention vs follow-up subgroup analysis

Note: Where studies reported data from multiple post-intervention timepoints, these were included as separate studies in each meta-analysis (e.g., Thomsen 2017= 16-week timepoint, Thomsen 2017a= 10 month timepoint), SD= standard deviation, 95% CI= 95% confidence interval

The meta-analysis included data from 10 studies (including 5 studies with multiple timepoints), with a total of 19 sets of data included in the meta-analysis, comprising n=1450 participants (n=736 in intervention groups, n=714 in control groups). Results showed a significant positive effect of lifestyle PA and SB interventions on reducing disease activity compared to the control group, with SMD of -0.12 (95% CI -0.23 to -0.01, I^2=6%, z=2.19, p=.03) (**Figure 2.2a**). GRADE analysis (**Table 2.2**) revealed results were not affected by the inclusion of studies with varied risk of bias, with moderate quality of evidence for this outcome due to studies being varied in their primary aims.

Subgroup analysis: Subgroup analyses showed that only lifestyle PA interventions, but not SB interventions, had significant effects on disease activity (**Figure 2.2a**). PA interventions (n=16, 1026 participants) demonstrated an SMD of -0.31 (95% CI -0.31 to -0.03, I^2=16%, z=2.45, p=.01), whilst SB interventions (n=3, 424 participants) an SMD of -0.02 (95% CI -0.21 to 0.17, I^2=0%, z=0.20, p=.84). When examining changes relative to different assessment timepoints, whilst lifestyle PA interventions showed significant during or immediately post-intervention effects on disease activity, no intervention effects were demonstrated at follow-up (**Figure 2.2b**). It was not possible to do this subgroup analysis on SB interventions due to insufficient data.

Secondary Outcomes

Lifestyle PA and SB: In total, 11 studies employed self-report methods to assess lifestyle PA and SB outcomes (sedentary time, steps, MVPA, total PA and leisure/light intensity PA), and 7 studies used device-based assessments [i.e., pedometers (Garner et al., 2018; Katz et al., 2018) and accelerometers (Gilbert et al., 2018; Li et al., 2020; McKenna et al., 2021; Thomsen et al., 2017; Van Den Berg et al., 2006)]. Only 2 interventions used both self-report and device-based measures (Thomsen et al., 2017; Van Den Berg et al., 2006). Overall, 7 studies reported increased post-intervention PA in the intervention group compared to the control group (Feldthusen et al., 2016; Giraudet-le Quintrec et al., 2007; Katz et al., 2018; Knittle et al., 2015; McKenna et al., 2021;

Nordgren et al., 2015; Van Den Berg et al., 2006). In addition, 2 studies reported significant intervention effects on SB compared to the control group (Katz et al., 2018; Thomsen et al., 2017). Method of assessment demonstrated little effect on intervention success at changing lifestyle PA or SB, however interventions targeting a particular dimension of lifestyle PA or SB, such as SB, steps, MVPA and leisure/light intensity PA, were more effective than those targeting total PA.

These findings are strengthened by meta-analysis results (**Supplementary Figures 2.9-2.13**), which revealed significant intervention effects on: sedentary time with an MD of -66.01 minutes/day (95% CI -102.28 to -29.74, I^2=70%, z=3.57, p<.001), steps with an MD of 617.17 steps/day (95% CI 30.34 to 1204.00, I^2=0%, z=2.06, p=.04), MVPA with an SMD of 0.99 (95% CI 0.37 to 1.6, I^2=96%, z=3.16, p=.002) and leisure/light intensity PA with an SMD of 0.55 (95% CI 0.41 to 0.70, I^2=32%, z=7.36, p<.001), with no intervention effects reported for total PA. Subgroup analysis revealed significant PA intervention effects on steps, MVPA and leisure/light intensity PA, and SB intervention effects on sedentary time and leisure/light intensity PA (**Supplementary Figures 2.9-2.13**). Post-intervention effects were demonstrated for steps, MVPA and leisure/light intensity PA, with significant follow-up effects displayed for sedentary time, MVPA and leisure/light intensity PA (**Supplementary Figures 2.22-2.26**).

OMERACT outcomes: A total of 5 interventions displayed significant between- and within-group improvements in functional ability (42%) (Katz et al., 2018; Nordgren et al., 2015; Thomsen et al., 2017; Van Den Berg et al., 2006; Veldhuijzen Van Zanten et al., 2021). A total of 3 interventions displayed between- and within-group reductions in pain (30%) (Li et al., 2020; Nordgren et al., 2015; Thomsen et al., 2017), with a further 1 study displaying reductions in pain, with no between-group difference (Katz et al., 2018). For fatigue, 3 interventions demonstrated reductions in intervention groups compared to control groups (27%). Finally, 7 studies demonstrated significant intervention effects on mental health, psychological wellbeing or quality of life outcomes (Brodin et al., 2008;

Feldthusen et al., 2016; Gilbert et al., 2018; Lange et al., 2020; Nordgren et al., 2015; Thomsen et al., 2017; Van Den Berg et al., 2006).

Meta-analyses reported significant intervention effects on: measures of functional ability (normal) with MD of -0.16 (95% CI -0.27 to -0.06, I^2=87%, z=3.06, p=.02), fatigue with a SMD of -0.35 (95% CI -0.5 to -0.19, I^2=60%, z=4.43 p<.0001) and depression (non-normally distributed data) with MD of -1.21 (95% CI -2.11 to -0.31, I^2=0%, z=2.64 p=.008). These effects all demonstrated improvements in outcomes. No other significant results were observed for OMERACT outcomes (**Supplementary Figures 2.1-2.8**). Subgroup analysis showed significant lifestyle PA intervention effects on increasing functional ability (normally distributed data), decreasing fatigue, and depression (non-normally distributed data). In addition, there were significant SB intervention effects on increasing functional ability (normally distributed data), decreasing pain and fatigue, and increasing quality of life (**Supplementary Figures 2.1-2.8**). Furthermore, immediate positive post-intervention effects were seen for functional ability (normally distributed data) and fatigue, whilst effects at follow-up were demonstrated for reducing pain, fatigue and depression (non-normally distributed data), and improving quality of life (**Supplementary Figures 2.14-2.21**).

Changes in lifestyle PA and SB in the context of OMERACT: Of the 2 studies demonstrating significant between- and within-group improvements in disease activity, both also displayed increases in intervention group leisure/light intensity PA (Katz et al., 2018; Lange et al., 2020). Regarding secondary outcomes, all studies reporting functional ability improvements also displayed intervention effects for lifestyle PA and/or SB (Katz et al., 2018; Nordgren et al., 2015; Thomsen et al., 2017; Van Den Berg et al., 2006; Veldhuijzen Van Zanten et al., 2021). Of the 4 studies reporting reductions in pain (Katz et al., 2018; Li et al., 2020; Nordgren et al., 2015; Thomsen et al., 2017), 3 also reported significant reductions in SB, and increased steps and leisure/light intensity PA (Katz et al., 2018; Nordgren et al., 2015; Thomsen et al., 2017). For fatigue, 2 of the 3 studies demonstrating

reductions in fatigue post-intervention also observed significant decreases in SB, and increases in steps and leisure/light intensity PA (Katz et al., 2018; Thomsen et al., 2017). Finally, 4 of the 7 studies reporting improvements in mental health, psychological wellbeing or quality of life following intervention, also demonstrated significantly increased lifestyle PA and/or reduced SB (Feldthusen et al., 2016; Nordgren et al., 2015; Thomsen et al., 2017; Van Den Berg et al., 2006).

Table 2.2: GRADE analysis for Disease Activity and secondary outcomes

Summary of findings table according to GRADE analysis							*Evaluation components to lower quality*						*Evaluation components to higher quality*		
Outcome	*Intervention Effects (SMD/MD)*	*No. studies*	*Participants IG*	*Participants CG*	*GRADE*	*Comments*	*Methodological design start point*	*Risk of bias*	*Inconsistency of results*	*Indirectness*	*Imprecision*	*Publication bias*	*Large effect*	*Dose response*	*Confounding*
Disease Activity	SMD= -0.12 [-0.23, -0.01]	19	736	714	Moderate ⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Mixture of RCTs and non-RCTs: High quality	74% studies had moderate RoB, 24% had high Rob: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 6%, no heterogeneity. No downgrade	Very few studies with disease activity as primary aim. Downgrade 1 level	N=1450 sample size, very large so unlikely to be imprecise. No downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). Funnel plot is asymmetrical, downgrade 1 level	z score = 2.19, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
Functional Ability (normal)	MD= -0.16 [-0.27, -0.06]	13	764	794	Low ⊕⊕◯◯	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.	Mixture of RCTs, non-RCTs and observational cohort studies: Moderate quality	61% studies had moderate RoB, 31% had high RoB: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 87%, considerable heterogeneity, downgrade 1 level	Studies highly varied in primary aim, with very few with function as primary aim. Downgrade 1 level	n=1558, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z score = 3.06, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
Functional Ability (non-normal)	MD= -0.00 [-0.03, 0.02]	9	521	516	High ⊕⊕⊕⊕	We are very confident that the true effect lies close to that of the estimate of the effect.	Mixture of RCTs and non-RCTs: High quality	100% studies had moderate RoB: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 0%, no heterogeneity	most studies primary aim is function, No downgrade	n=1037, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z score = 0.07, no effect. No upgrade	N/A	We found no confounding factors that indicate upgrading
Pain	SMD= -0.10 [-0.59, 0.39]	19	1098	1117	Very Low ⊕◯◯◯	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect	Mixture of RCTs, non-RCTs and observational cohort studies: Moderate quality	85% studies moderate, 10% high RoB: no downgrade (2 studies with high RoB small sample)	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 97%, considerable heterogeneity, downgrade 1 level	Studies varied in primary aim, with 6 with pain as primary aim (<50%). Downgrade 1 level	n>2000, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). Funnel plot is asymmetrical, downgrade 1 level	z= 0.4, no effect. No upgrade	N/A	We found no confounding factors that indicate upgrading

Fatigue	SMD = -0.35 [-0.50, -0.19]	22	980	961	Moderate ⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Mixture of RCTs, non-RCTs and observational cohort studies: Moderate quality	73% moderate, 23% high RoB: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 60%, not considerable (<75%), no downgrade	studies had varied primary aims, downgrade 1 level	n=2000, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 4.43, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
Anxiety	SMD= -0.18 [-0.47, 0.12]	8	259	229	Moderate ⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Mixture of RCTs and non-RCTs: High quality	75% moderate, 25% high risk: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 55%, not considerable (<75%), no downgrade	<50% studies had anxiety in their primary aim, downgrade 1 level	n= 488, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 1.18, little effect. No upgrade	N/A	We found no confounding factors that indicate upgrading
Depression (non-normal)	MD= -1.21 [-2.11, -0.31]	5	144	121	Moderate ⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Mixture of RCTs and non-RCTs: High quality	100% studies had moderate RoB: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 0%, no heterogeneity	<50% studies had depression in their primary aim, downgrade 1 level	n= 265, small sample so likely to be imprecise, downgrade 1 level	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 2.64, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
Depression (normal)	SMD= -0.05 [-0.32, 0.21]	9	345	326	Moderate ⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Mixture of RCTs and non-RCTs: High quality	56% studies had moderate Rob, 33% high risk: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 62%, not considerable (<75%), no downgrade	<50% studies had depression in their primary aim, downgrade 1 level	n= 671, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 0.39, no effect. No upgrade	N/A	We found no confounding factors that indicate upgrading
Quality of Life	SMD= -0.37 [-0.98, 0.25]	24	1587	1580	Very Low ⊕◯◯◯	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect	Mixture of RCTs, non-RCTs and observational cohort studies: Moderate quality	92% studies had moderate RoB, 8% high risk: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 98%, considerable heterogeneity, downgrade 1 level	<50% studies had QoL in their primary aim, downgrade 1 level	n>3000, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). Funnel plot is asymmetrical, downgrade 1 level	z= 1.17, little effect. No upgrade	N/A	We found no confounding factors that indicate upgrading

Sedentary Time	MD= -66.01 [-102.28, -29.74]	6	369	369	High ⊕⊕⊕⊕	We are very confident that the true effect lies close to that of the estimate of the effect.	Mixture of RCTs and non-RCTs: High quality	50% studies had moderate RoB, 50% high risk: downgrade 1 level	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 70% (<75%), no downgrade	>50% studies primary aim was to target SB, no downgrade	n=738, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 3.57, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
Steps	MD= 617.17 [30.34, 1204.00]	3	231	227	High ⊕⊕⊕⊕	We are very confident that the true effect lies close to that of the estimate of the effect.	Mixture of RCTs and non-RCTs: High quality	33% studies had moderate RoB, 67% high risk: downgrade 1 level	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 0%, no heterogeneity, no downgrade	All studies outcomes were PA/SB measure so these are sufficiently similar, no downgrade	n=458, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 2.06, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
MVPA		17	800	759	High ⊕⊕⊕⊕	We are very confident that the true effect lies close to that of the estimate of the effect.	Mixture of RCTs and non-RCTs: High quality	65% studies moderate risk, 35% high risk: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 96%, considerable heterogeneity, downgrade 1 level	>50% studies primary aim was to target MVPA, no downgrade	n=1559, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 3.16, large effect. Upgrade 1 level	N/A	We found no confounding factors that indicate upgrading
Total PA	SMD= -0.04 [-0.34, 0.27]	11	620	585	Moderate ⊕⊕⊕◯	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Mixture of RCTs and non-RCTs: High quality	73% studies had moderate RoB, 27% high risk: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 85%, considerable heterogeneity, downgrade 1 level	>50% studies primary aim was to target PA, no downgrade	n=1205, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 0.24, no effect. No upgrade	N/A	We found no confounding factors that indicate upgrading
Light/leisure PA	SMD= 0.55 [0.41, 0.70]	12	602	560	High ⊕⊕⊕⊕	We are very confident that the true effect lies close to that of the estimate of the effect.	Mixture of RCTs and non-RCTs: High quality	83% studies had moderate RoB, 17% high risk: no downgrade	Even though we used a random effect model meta-analysis, we consider heterogeneity as an index of inconsistency. I2 = 32%, no heterogeneity	>50% studies primary aim was to target some form of PA or SB, no downgrade	n=1162, unlikely to be imprecise, no downgrade	We used an exhaustive searching approach (i.e. scientific databases, grey literature, scientific organizations). No major bias in the funnel plots. No downgrade	z= 7.36, large effect. Upgrade 2 levels	N/A	We found no confounding factors that indicate upgrading

Note: an overall quality score is obtained using the assessments of risk of bias, inconsistency, indirectness, imprecision, publication bias, large effect, dose response and confounding factors for all outcomes.

IG: intervention group, CG= control group, FA= functional ability, MVPA= moderate to vigorous physical activity, PA= physical activity, SMD= standardised mean difference, MD= mean difference, RCT= randomised controlled trial, RoB= risk of bias.

Risk of Bias Assessment Results

A summary of the RoB2 assessment with disease activity and functional ability as outcomes is illustrated in **Figures 2.3a and 2.4**, respectively. To summarise, for disease activity no studies had a low risk of bias, 7 had some concerns (Brodin et al., 2008; Feldthusen et al., 2016; Katz et al., 2018; Lange et al., 2020; Thomsen et al., 2017; Van Den Berg et al., 2006; Veldhuijzen Van Zanten et al., 2021), and 4 had high risk of bias (Garner et al., 2018; Giraudet-le Quintrec et al., 2007; Knittle et al., 2015; McKenna et al., 2021). For functional ability outcome no studies were low risk, 8 had some concerns (Brodin et al., 2008; Gilbert et al., 2018; Giraudet-le Quintrec et al., 2007; Katz et al., 2018; Lange et al., 2020; Thomsen et al., 2017; Van Den Berg et al., 2006; Veldhuijzen Van Zanten et al., 2021), and 3 had high risk of bias (Garner et al., 2018; Knittle et al., 2015; McKenna et al., 2021). Full domain results of RoB2 analysis for disease activity can be visualised in **Figure 2.3b**.

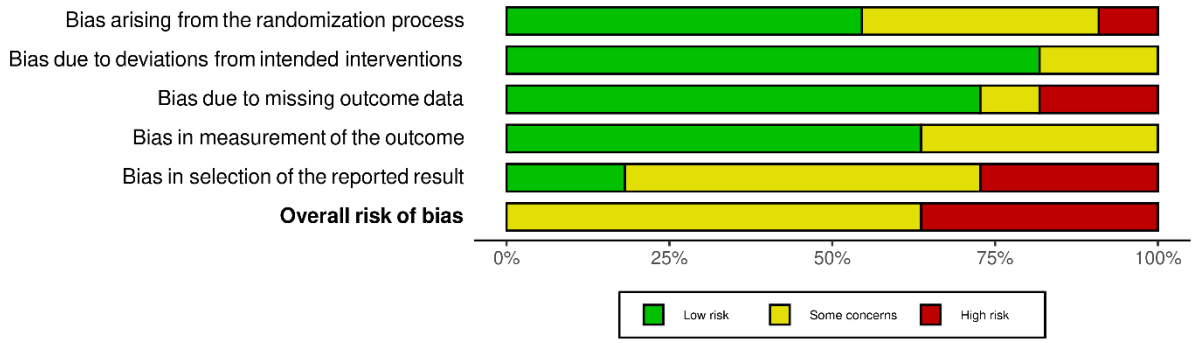


Figure 2.3a: Summary Risk of bias assessment for Disease Activity

Note: ROB domains include; (1) Bias arising from the randomization process; (2) Bias due to deviations from intended interventions; (3) Bias due to missing outcome data; (4) Bias in measurement of the outcome; and (5) Bias in selection of the reported result.

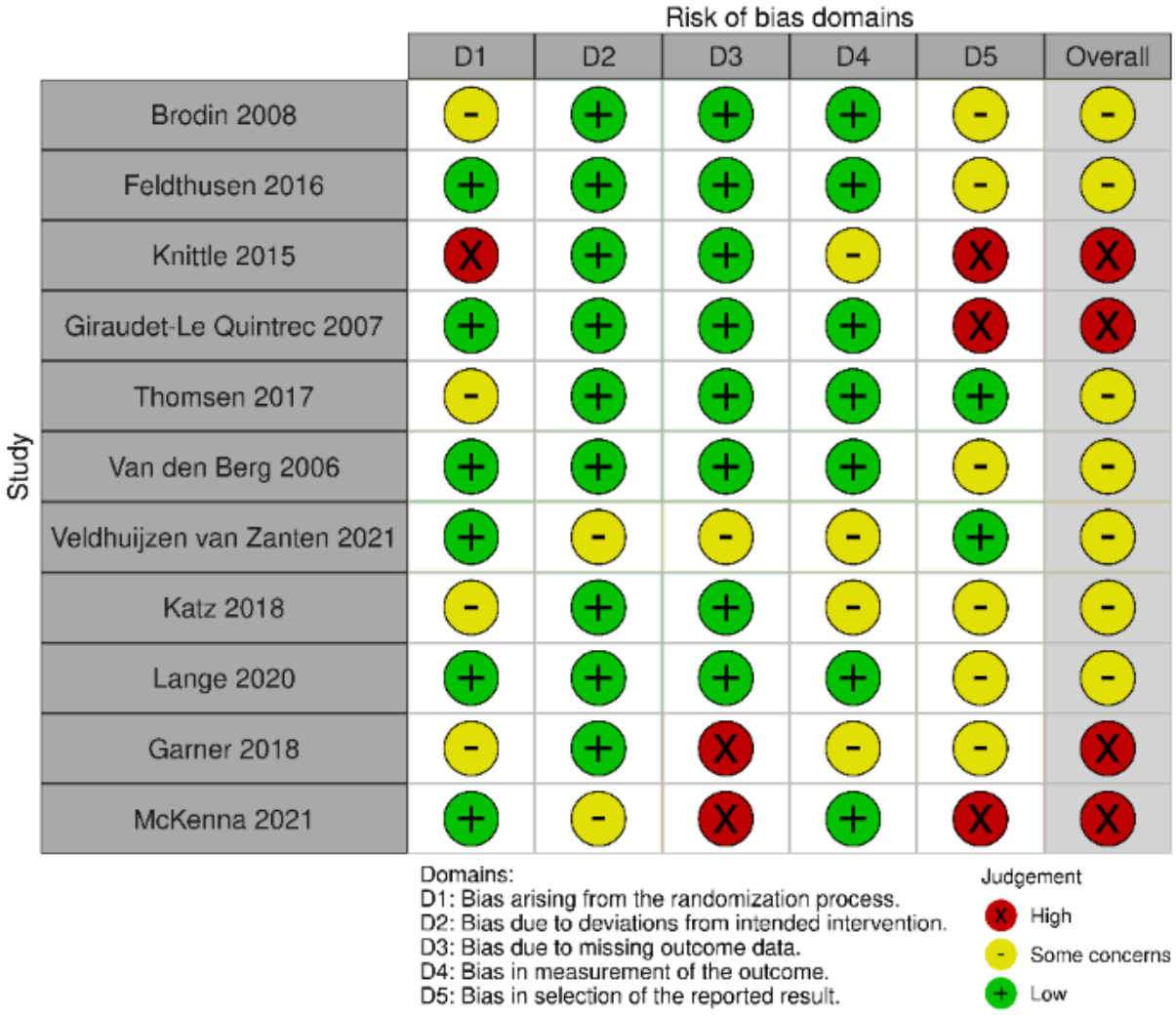


Figure 2.3b: Risk of bias assessment for Disease Activity

Note: With disease activity as the outcome of interest: 55% studies showed low risk of bias, 36% showed some concerns and 9% had high risk of bias arising from the randomisation process, due to insufficient information about blinding in the randomisation process. In "deviations from intended interventions", 82% studies displayed low risk of bias, and only 18% had some concerns, indicating that few studies appeared to deviate from their protocol or methods.73% included studies demonstrated low risk, 9% had some concerns and 18% had high risk of bias due to missing outcome data, as some studies were feasibility studies, with small sample sizes. For the "bias in measurement of the outcome" domain, 55% studies demonstrated low risk and the remaining 45% displayed some concerns. This domain was mostly low risk due to the disease activity measures being valid and partially objective in nature. For "bias in selection of the reported result", 18% studies showed low risk, with 55% showing some concerns and 27% with high risk of bias, due to missing data at some pre-specified timepoints

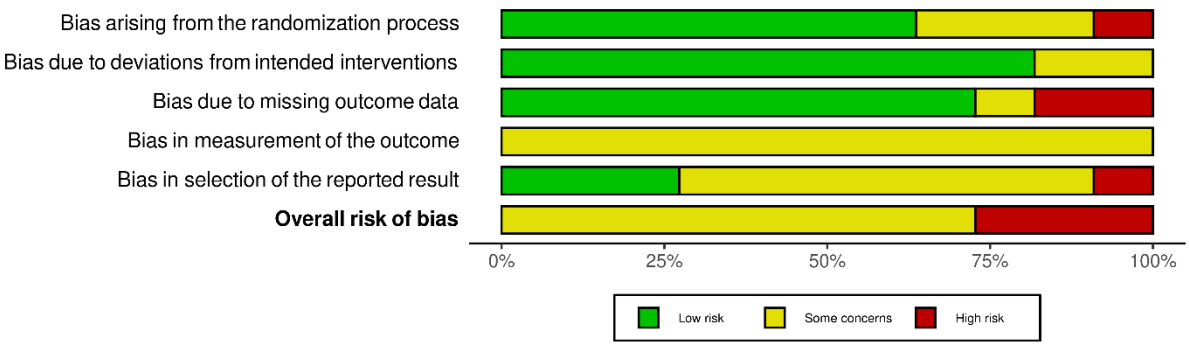


Figure 2.4: Summary Risk of bias assessment for Functional Ability

Note**:** ROB domains include; (1) Bias arising from the randomization process; (2) Bias due to deviations from intended interventions; (3) Bias due to missing outcome data; (4) Bias in measurement of the outcome; and (5) Bias in selection of the reported result.

Discussion

This systematic review and meta-analysis identified 16 lifestyle PA and SB interventions in RA, and aimed to evaluate their effect on disease activity, lifestyle PA and SB, and OMERACT patient- and clinician-important outcomes in people with RA.

Overview of Main Outcomes

Lifestyle PA interventions demonstrated significant effects on reducing disease activity in individuals with RA. Statistically significant effects were also observed for sedentary time, steps, leisure/light intensity PA and MVPA. The majority of interventions which displayed increased PA and reductions in SB also revealed improvements in other OMERACT outcomes. Specifically, lifestyle PA interventions were effective at improving functional ability, fatigue and depression, and the 1 SB intervention reported significant effects on all secondary outcomes assessed in their study (functional ability, pain, fatigue, quality of life). Despite this, findings also revealed lifestyle PA and SB interventions were unsuccessful at targeting total PA, pain, anxiety and quality of life in people with RA.

Completeness and Applicability of Evidence

Our analysis showed that lifestyle PA interventions may be beneficial to treat disease activity in RA, supporting findings from observational studies (Metsios & Kitas, 2018; Metsios et al., 2015). Those interventions demonstrating efficacy in improving disease activity displayed similar characteristics: long in duration (approximately 20 weeks), with a primary focus on promoting light-to-moderate intensity PA or walking (Katz et al., 2018; Lange et al., 2020). These results add to emerging evidence which suggests that light intensity PA is linked with disease activity and inflammation in people with RA (Fenton et al., 2017; Khoja et al., 2016). Together, a longer length of intervention, targeting light intensity PA may be required for detectable changes in disease activity in people with RA. The only SB intervention included in this review showed no effects on disease activity, which highlights the

need for more experimental studies investigating the role of SB for disease activity in RA (Fenton et al., 2017).

We provide evidence that lifestyle PA and SB interventions are effective at decreasing sedentary time and increasing leisure/light intensity PA, MVPA and steps in people with RA. Interventions demonstrated a reduction in sedentary time by 66 minutes/day, and an increase in 617 steps/day. Previous research in healthy populations has indicated that an increase in steps by 1000 to 2000 steps/day is sufficient to be deemed clinically meaningful (Dwyer et al., 2011; Thomson et al., 2012). However, other studies of older adults and people with RA have shown that an increase in steps of any amount is sufficient to give some clinical benefits due to its high correlations with mortality and morbidity outcomes (Ewald et al., 2014; Summers et al., 2019). For sedentary time, O'Brien et al. (2021) previously found a reduction in sedentary time by 33 minutes/day was sufficient to display clinically significant reductions in pain and fatigue. Fenton et al. (2017) also revealed that reducing sedentary time by 68 minutes/day equated to a significant 5.5% reduction in cardiovascular disease risk. Together, this suggests that my findings of a 66 minutes/day reduction in sedentary time resulting from lifestyle PA and SB intervention is clinically significant.

Still, it is interesting that previous research has also reported that SB interventions were more effective than PA-only or combined PA+SB interventions at reducing SB (Nguyen et al., 2020). However, this review was not conducted in people with RA. This meta-analysis included only 1 SB intervention, limiting the ability to conduct comparable analyses in the case of RA. Because of this, effects have the potential to be underpowered. Further SB interventions are needed to elucidate if targeting and reducing SB is sufficient to improve disease activity and other core patient- and clinician-important outcomes in people with RA. The dimensions of PA or SB assessed (i.e., SITT and FITT elements), and measurement tools used to quantify these outcomes (i.e., self-report, device-based, validated vs not validated) varied between studies. Although, all interventions that targeted

SB and steps used device-based measures, and all demonstrated significant reductions in sedentary time and increased steps. This suggests device-based measurement may act as aids or reminders to be active and minimise sitting. Therefore, future behavioural interventions should employ validated device-based measures to target SB and steps in people with RA.

Effective lifestyle PA and SB interventions may play a role in improving OMERACT outcomes. My findings agree with results of previous systematic and narrative reviews highlighting the effects of general PA and exercise training on health outcomes, in people living with RA (Larkin & Kennedy, 2014; Metsios & Kitas, 2018; Metsios et al., 2015). The lack of beneficial effect of interventions on some secondary outcomes may be due to heterogeneity between the interventions, in terms of intervention length, content and method of outcome assessment (**Table 2.1**). This was indicated by large I^2 statistic for these outcomes (I^2= 0-98%). This highlights the need for a consensus on optimal measurement methods and reporting for these OMERACT outcomes (e.g., MD, rather than SMD), in order for interventions effects on outcomes to be reliably and accurately assessed in future meta-analyses.

Findings from subgroup analyses revealed post-intervention effects of lifestyle PA and SB interventions on steps, MVPA, disease activity and functional ability, however, these were not sustained at follow-up. Whereas, no post-intervention effects were observed for sedentary time, depression, quality of life and pain, although follow-up effects on these outcomes were seen. Both post-intervention and follow-up effects were demonstrated for light/leisure PA and fatigue. The varied results regarding intervention efficacy at different assessment timepoints may be due to follow-up periods being particularly heterogeneous between studies. For example, follow-up periods ranged from 6 months to 4 years post-intervention between studies. A more consistent approach between interventions would give a greater insight into the long-term effectiveness of these interventions. Many interventions (n= 9) included in this review did not conduct follow-up

assessments, therefore it is not surprising that little is known regarding the effectiveness of interventions to promote long-term adherence to PA and SB. By necessitating that follow-up assessments are done, this ensures interventions are targeting sustained clinical benefits (Metsios & Kitas, 2018; Veldhuijzen Van Zanten et al., 2015). In addition, interventions which demonstrate beneficial effects at follow-up reflect a lifestyle change, whereby adoption becomes maintenance (Lange et al., 2020). Previous reviews and qualitative findings have reported that a main challenge of an intervention program is to assess and ensure beneficial effects post-intervention (Metsios & Kitas, 2018; Nguyen et al., 2016; Swardh et al., 2020). Therefore, future interventions should conduct regular follow-up assessments over long periods, to assess their long-term clinical efficacy.

Interventions where the primary focus was on promoting PA or reducing SB, were more successful in terms of number of observed significant improvement in behaviours (increased PA and or reduced SB) and outcomes. These interventions were often more personalised and tailored to individuals' abilities and had good adherence (Katz et al., 2018; McKenna et al., 2021; Thomsen et al., 2017). These more focused interventions may be deemed more feasible by people with RA, who have additional disease-related barriers to PA (Veldhuijzen Van Zanten et al., 2015), leading to more successful implementation. This supports findings of a previous meta-analysis conducted in healthy adults (Martin et al., 2015). By contrast, multi-component interventions (e.g., including counselling, education, nutrition advice or self-management components), targeting multiple health behaviours appeared to be less effective, with fewer significant improvements in outcomes, increases in PA and/or reductions in SB.

Interestingly, successful interventions included regular support, most commonly in the form of regular phone calls at a frequency of every 1-2 weeks (Katz et al., 2018; Li et al., 2020; Nordgren et al., 2015), or individualised based on goals (Feldthusen et al., 2016; Thomsen et al., 2017). However frequency and type of support varied across studies. Future research could explore what mode and

frequency of support is likely to be optimal for this patient group. Successful interventions were also more likely to be delivered in accessible settings, rather than a specified facility (e.g., public training centre, gym), which has previously shown to be an obstacle for intervention adherence (Cramp et al., 2020; Katz et al., 2018; Nordgren et al., 2015). These resource intensive interventions were generally more multicomponent in nature, and required travel to other settings for the other components of their interventions (Nordgren et al., 2015). Perhaps this complex nature, and focus on multiple health behaviours of some multicomponent interventions, diluted down the key message of lifestyle PA and SB interventions, to move more.

The efficacy of interventions differed according to the specific dimension and element of PA and SB targeted (i.e., FITT, SITT), with effects reported only for interventions targeting MVPA, steps, sedentary time and leisure/light intensity PA. Interventions which focus on these dimensions of PA and/or SB may be deemed more feasible, achievable or preferable to participants with RA, confirming the findings of previous research (Manns et al., 2012; Thomsen et al., 2017).

Interestingly, the majority of measurement tools used to assess PA were various self-report questionnaires. Whereas interventions assessing MVPA, steps and SB in particular more consistently used device-based measures. As interventions which targeted and measured MVPA, steps and SB were generally more effective than those targeting total PA, the validity and reliability of measurement tools used may have impacted if effects were seen. For example, self-report methods are subject to recall bias, and this may explain why no effects were observed for total PA. Therefore, results for total PA and other PA behaviours (e.g., light/leisure PA) should be interpreted with caution (Sylvia et al., 2014). Nevertheless, device-based and self-report measures of PA and SB are not conceptually equivalent, producing different outputs for each other, and are thought to be different approaches to measure PA and SB (Troiano et al., 2014). Therefore, interventions which use device-based vs self-report measures of PA or SB perhaps shouldn't be grouped together in meta-

analyses, and future reviews should analyse the effectiveness of interventions using device-based vs self-report measures separately.

Strengths and Limitations

Strengths include the use of transparent methods, clear inclusion criteria and a robust search strategy; and therefore, results and conclusions are likely to be valid and can be replicated in future reviews. The subgroup analysis allowed for the exploration of moderating variables, to give more investigative interpretation of results. Lastly, my choice of OMERACT outcomes to describe RA-related health helped to identify gaps in current research, which should be addressed in future interventions.

In meta-analyses, functional ability and depression outcomes could not be successfully transformed, so were split into normal and non-normal outcomes which gave different results. Therefore, findings regarding these outcomes should be interpreted with caution. Moreover, no subgroup analyses were undertaken for mode of intervention delivery (e.g., individual, group, internet, app-based), dimension of lifestyle PA/SB targeted, nature of the comparison group (e.g., placebo, no intervention, advice only), and whether or not the intervention had a theoretical basis. This was due to heterogeneity between studies; meaning we were unable to confidently group studies into these categories.

Potential Implications

Future interventions should be clearer and more specific in describing subgroups in order for future meta-analyses to be able to assess their efficacy at improving OMERACT patient- and clinician-important outcomes in people with RA. In addition, the majority of studies scored poorly in risk of bias assessments and GRADE analysis (Sterne et al., 2019). Therefore, future studies should publish trial registrations or protocols, provide information about participant and personnel blinding, and use

validated measures to assess outcomes to ensure transparent reporting of results. Moreover, some small-scale feasibility interventions were included in this review which were not adequately powered to detect significant changes in outcomes. These were, consequently, graded as high risk of bias. Nevertheless, conducting feasibility studies shows good research practice, and future large-scale interventions using identical study designs and methods are welcomed to confirm and strengthen their findings. The choice of outcomes were varied and inconsistent between studies, showing little consideration of OMERACT guidelines (Boers et al., 2014). There was also little consistency between outcome measurement methods, as demonstrated by the high I^2 statistic results for many meta-analyses. Accordingly, researchers need to provide a consensus on the optimal methods and outcomes to reliably assess the efficacy of lifestyle PA and SB interventions in the RA population. Finally, as studies consistently displayed moderate to high risk of bias selection of reported result and measurement of outcome domains (Figures 2.3a and 2.4), future investigations should seek to provide more detailed explanations of study design and methods to enable further researchers to replicate these and strengthen findings.

Conclusions

The lifestyle PA and SB interventions included in this systematic review demonstrated significant increases in PA and reductions in SB, as well as improvements in disease activity and other core OMERACT outcomes in people with RA. PA and SB interventions differed in effectiveness at targeting different outcomes, due to differences in content, structure and focus of intervention. In addition, due to varied follow-up assessment periods, intervention benefits on OMERACT outcomes at post-intervention and follow-up were inconsistent. Future research in this area should seek to standardise PA, SB and outcome measures and measurement tools across studies, and employ regular/consistent follow-up periods to allow long-term clinical benefit of interventions to be assessed. More studies are also required to explore the value of interventions targeting SB for improving health in RA.

Declarations

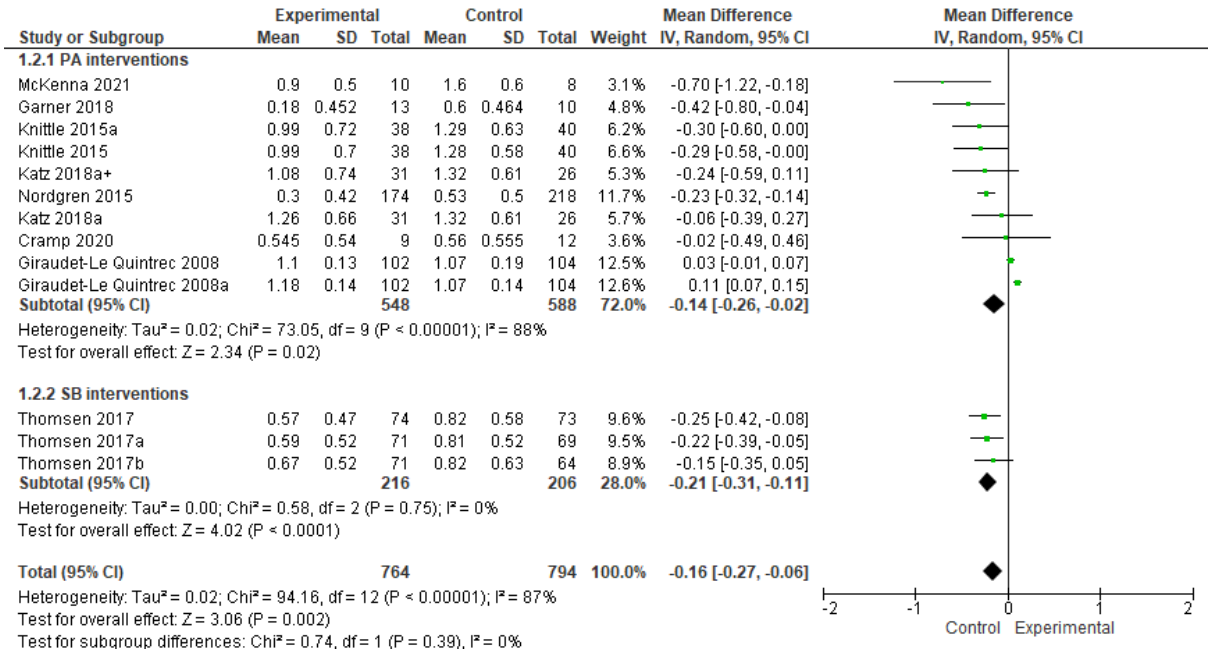
AUTHOR CONTRIBUTORSHIP: Conception and design of the study: Sophia Brady, Sally Fenton, Jet Veldhuijzen van Zanten, George Metsios. Data acquisition: Sophia Brady, Saleh Elmsmari. Data analysis: Sophia Brady, Saleh Elmsmari, Thomas Nightingale, Petros Dinas. Data interpretations and drafting of manuscript: all authors. Final approval of manuscript: all authors.

Supplementary Table 2.1: Search Strategies for 8 databases

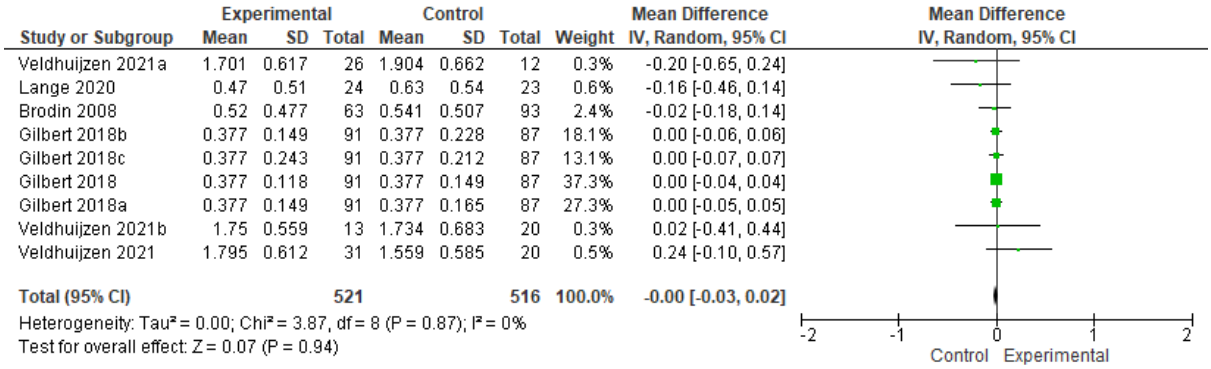
Database	*Search Strategy*
Cochrane Library	("rheumatoid arthritis") AND ("physical activity" OR Physical Fitness OR "physical function" OR "Activities of Daily Living" OR "lifestyle physical activity" OR "life-style physical activity" OR Motor Activity OR "lifestyle behavior" OR "lifestyle behaviour" OR sedentary OR "Sedentary Behavior" OR "sedentary lifestyle" OR "sedentary behaviour" OR sitting OR "sitting time" OR "sitting behavior" OR screen time) AND (randomized controlled trial OR controlled clinical trial OR Random Allocation OR "randomised controlled trial" OR randomized OR single-blind method OR Double-Blind Method OR trial OR groups OR intervention OR placebo) AND (promot* OR educat* OR uptake OR start OR increase OR program*)
CINAHL Plus	("rheumatoid arthritis") AND ("physical activity" OR "Physical Fitness" OR "physical function" OR "Activities of Daily Living" OR "lifestyle physical activity" OR "life-style physical activity" OR "Motor Activity" OR "lifestyle behavior" OR "lifestyle behaviour" OR "sedentary" OR "Sedentary Behavior" OR "sedentary lifestyle" OR "sedentary behaviour" OR "sitting" OR "sitting time" OR "sitting behavior" OR "screen time") AND ("randomized controlled trial" OR "controlled clinical trial" OR "Random Allocation" OR "randomised controlled trial" OR "randomized" OR "randomised" OR "single-blind method" OR "Double-Blind Method" OR "trial" OR "groups" OR "intervention" OR "placebo") AND ("promot*" OR "educat*" OR "uptake" OR "start" OR "increase" OR "program*")
Scopus	ABS ("rheumatoid arthritis") AND ("physical activity" OR physical AND fitness OR "physical function" OR "activities of daily living" OR "lifestyle physical activity" OR "life-style physical activity" OR motor AND activity OR "lifestyle behavior" OR "lifestyle behaviour" OR sedentary OR sedentary AND behavior OR sedentary AND lifestyle OR "sedentary behaviour" OR sitting OR "sitting time" OR "sitting behavior" OR screen AND time) AND (randomized AND controlled AND trial OR controlled AND clinical AND trial OR random AND allocation OR "randomised controlled trial" OR randomized OR single-blind AND method OR double-blind AND method OR trial OR groups OR intervention OR placebo) AND (promot* OR educat* OR uptake OR start OR increase OR program*) AND (LIMIT-TO (DOCTYPE , "ar"))
PEDro	"rheumatoid arthritis" filter by clinical trials AND health promotion/fitness training/behaviour modification/education
PsychINFO, Medline, EMBASE	"rheumatoid arthritis" "physical activity" OR Physical Fitness OR "physical function" OR "Activities of Daily Living" OR "lifestyle physical activity" OR "life-style physical activity" OR Motor Activity OR "lifestyle behavior" OR "lifestyle behaviour" OR sedentary OR "Sedentary Behavior" OR "sedentary lifestyle" OR "sedentary behaviour" OR sitting OR "sitting time" OR "sitting behavior" OR screen time randomized controlled trial OR controlled clinical trial OR Random Allocation OR "randomised controlled trial" OR randomized OR single-blind method OR Double-Blind Method OR trial OR groups OR intervention OR placebo promot* OR educat* OR uptake OR start OR increase OR program*
Web of Science	TS= ("rheumatoid arthritis") AND TS=("physical activity" OR "Physical Fitness" OR "physical function" OR "Activities of Daily Living" OR "lifestyle physical activity" OR "life-style physical activity" OR "Motor Activity" OR "lifestyle behavior" OR "lifestyle behaviour" OR "sedentary" OR "Sedentary Behavior" OR "sedentary lifestyle" OR "sedentary behaviour" OR "sitting" OR "sitting time" OR "sitting behavior" OR "screen time") AND TS=("randomized controlled trial" OR "controlled clinical trial" OR "Random Allocation" OR "randomised controlled trial" OR "randomized" OR "single-blind method" OR "Double-Blind Method" OR "trial" OR "groups" OR "intervention" OR "placebo") AND TS=("promot*" OR "educat*" OR "uptake" OR "start" OR "increase" OR "program*")

Note: CINAHL= Cumulative Index to Nursing & Allied Health Literature, EMBASE= Excerpta Medica database, PEDro= Physiotherapy Evidence Database

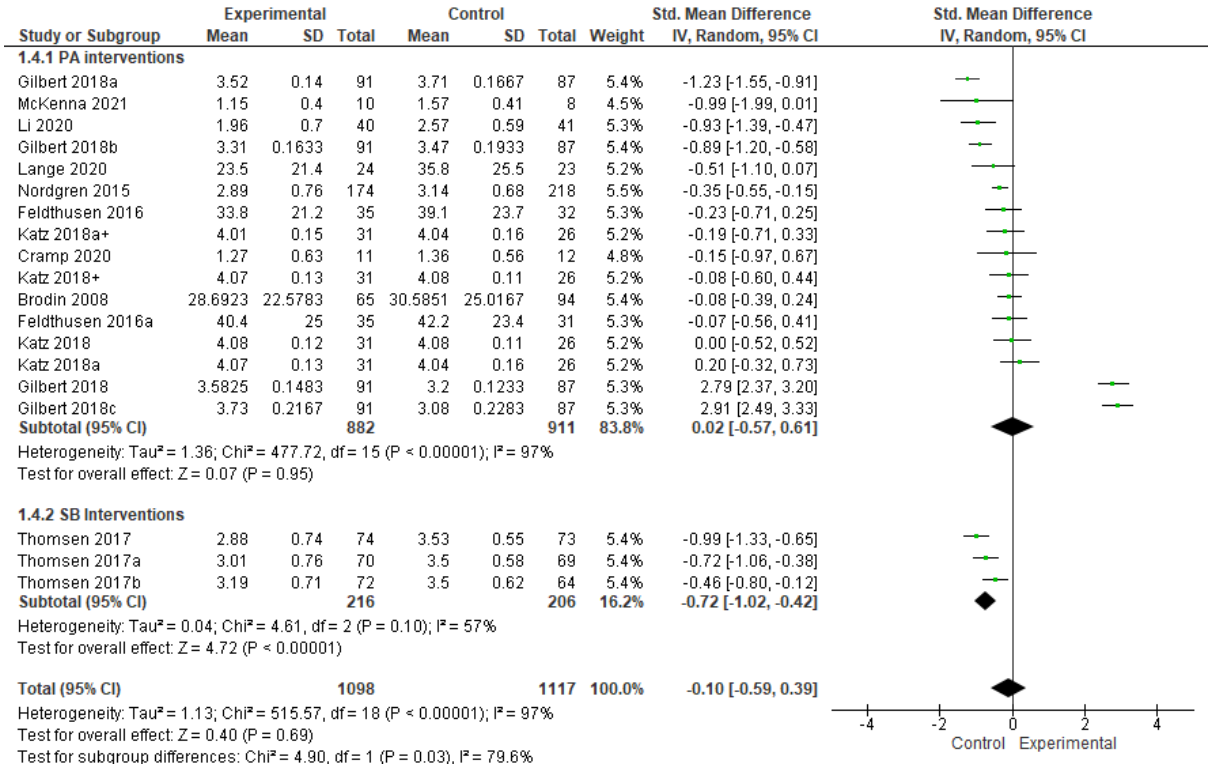
Supplementary Figures 2.1-2.13: Forest plots for secondary outcomes- Physical Activity vs Sedentary Behaviour interventions



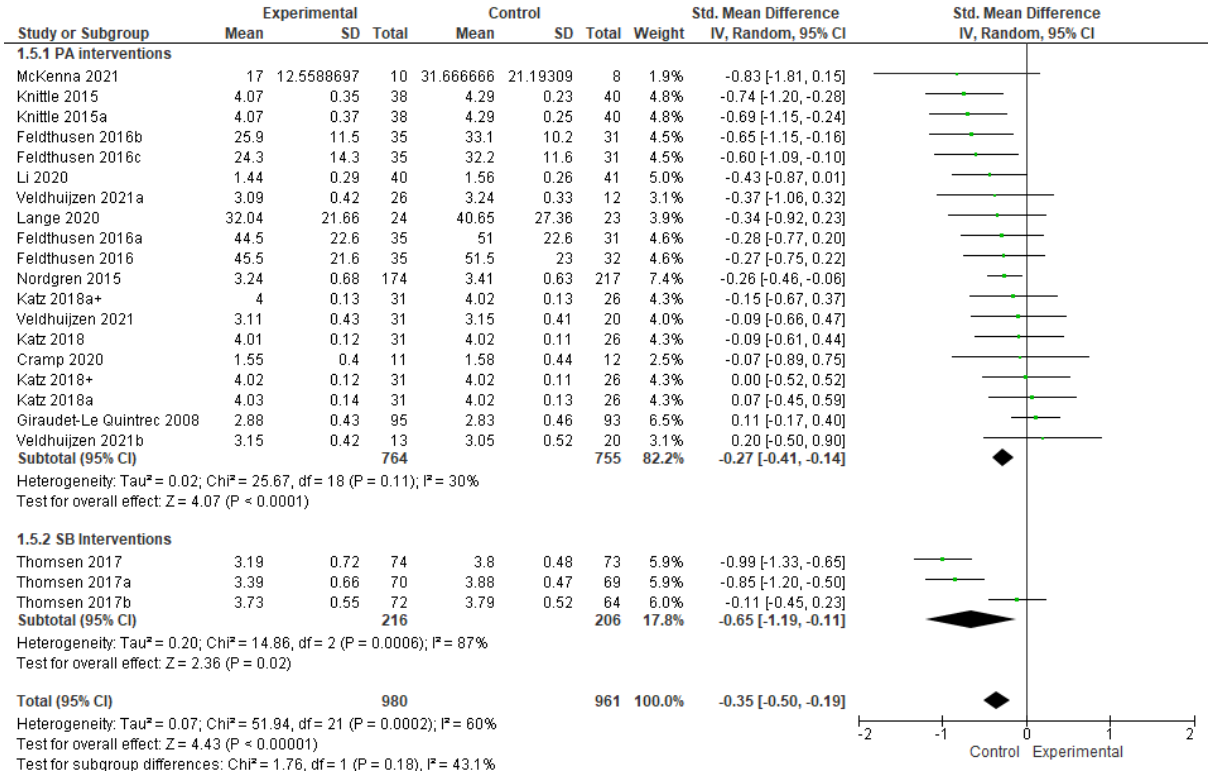
Supplementary Figure 2.1: The effects of interventions on functional ability (normally distributed). SD= standard deviation, 95% CI= 95% confidence interval.



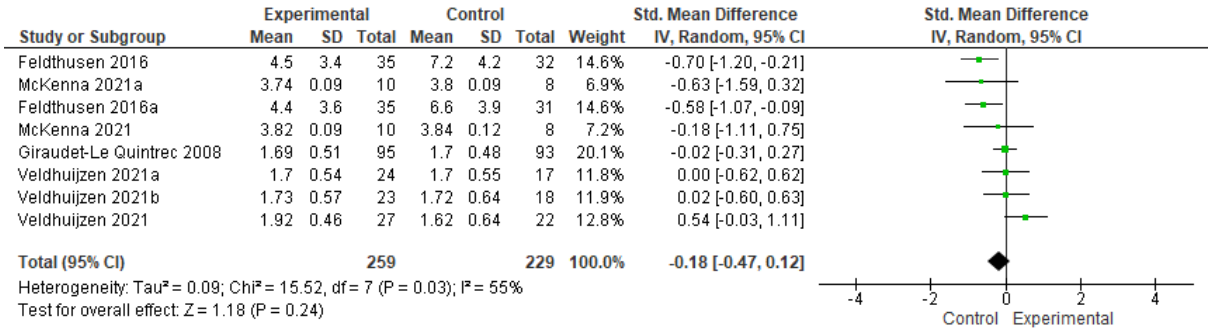
Supplementary Figure 2.2: The effects of interventions on functional ability (non-normally distributed). SD= standard deviation, 95% CI= 95% confidence interval.



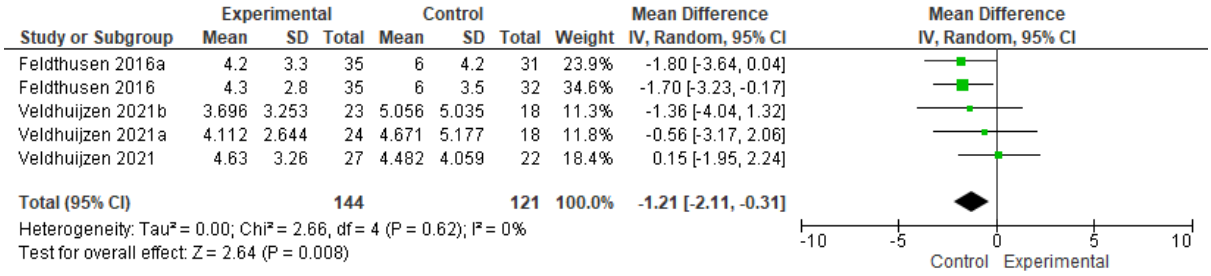
Supplementary Figure 2.3: The effects of interventions on pain. SD= standard deviation, 95% CI= 95% confidence interval.



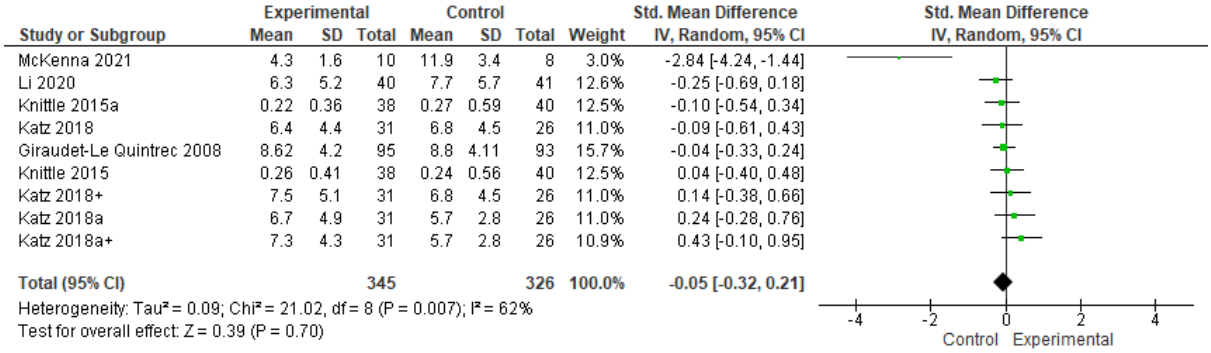
Supplementary Figure 2.4: The effects of interventions on fatigue. SD= standard deviation, 95% CI= 95% confidence interval.



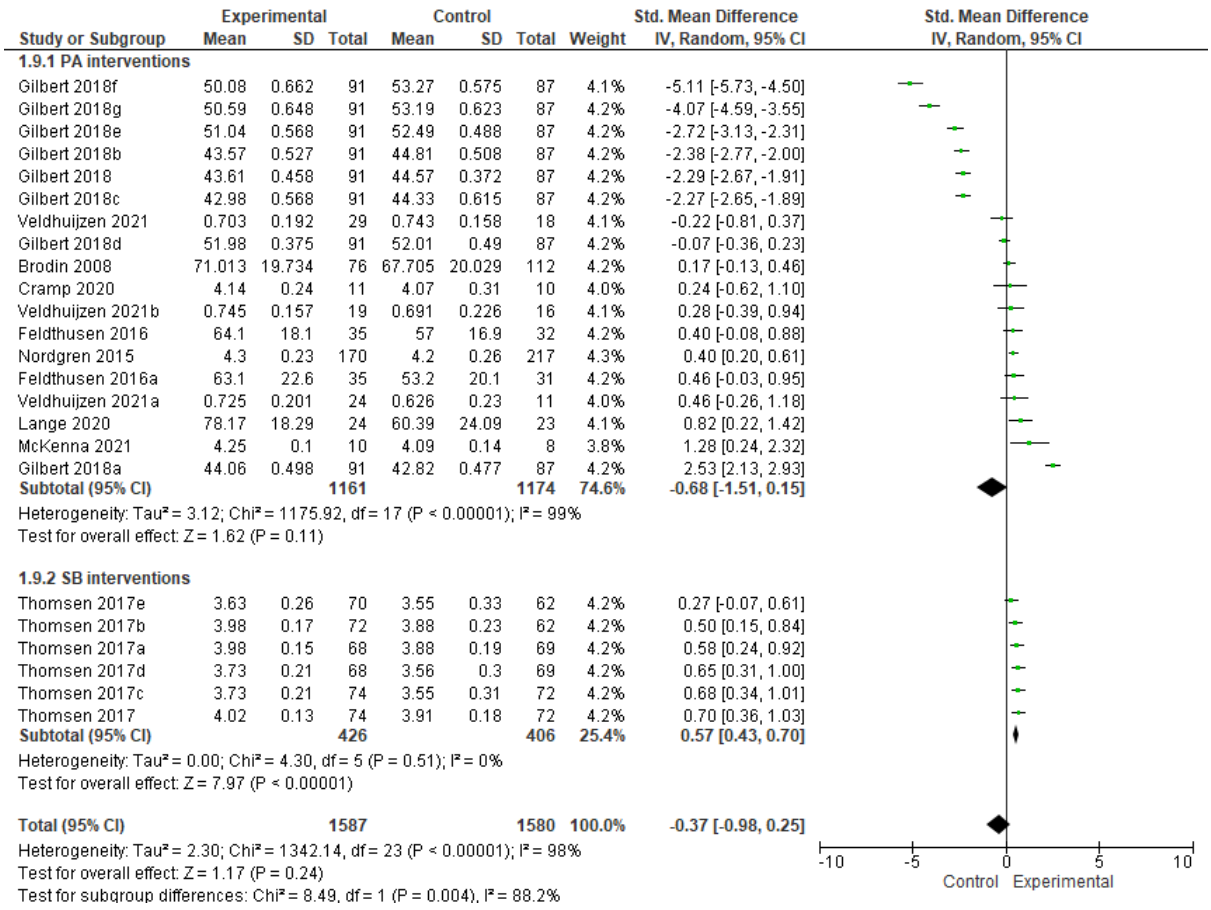
Supplementary Figure 2.5: The effects of interventions on anxiety. SD= standard deviation, 95% CI= 95% confidence interval.



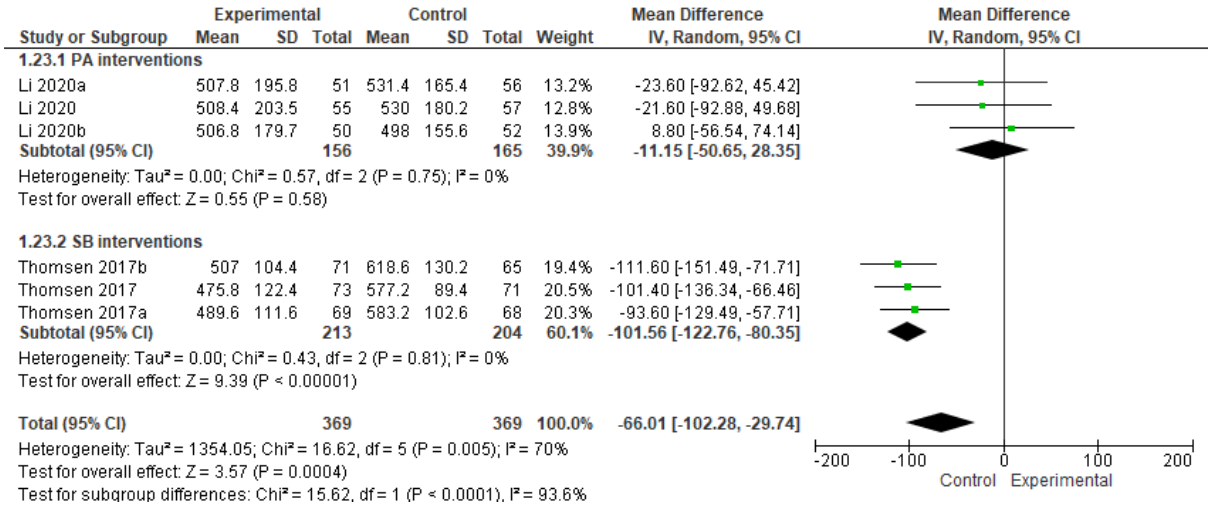
Supplementary Figure 2.6: The effects of interventions on depression (non-normally distributed). SD= standard deviation, 95% CI= 95% confidence interval.



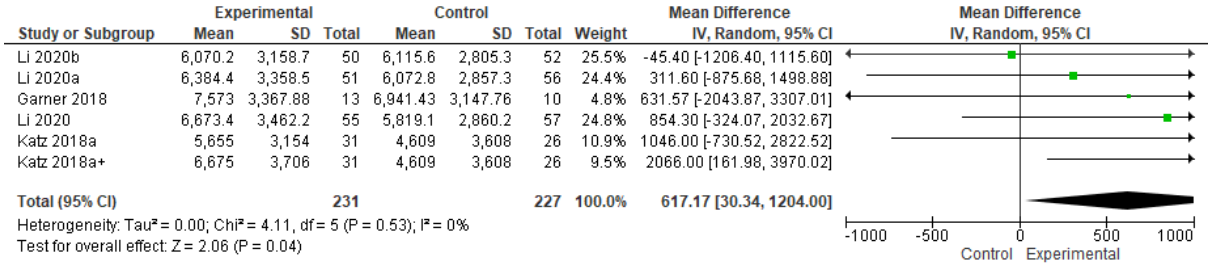
Supplementary Figure 2.7: The effects of interventions on depression (normally distributed). SD= standard deviation, 95% CI= 95% confidence interval.



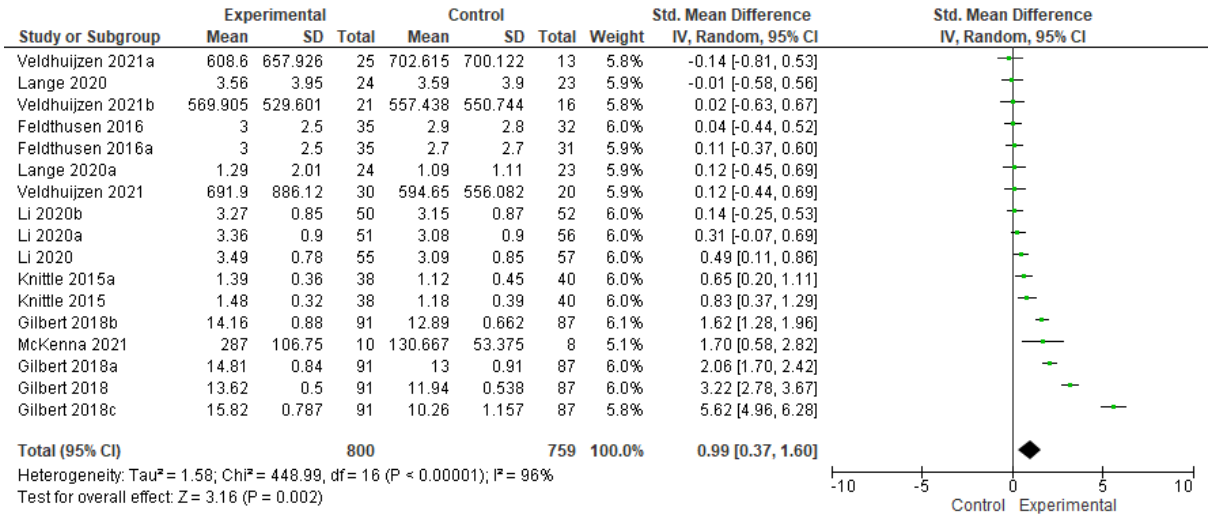
Supplementary Figure 2.8: The effects of interventions on quality of life. SD= standard deviation, 95% CI= 95% confidence interval.



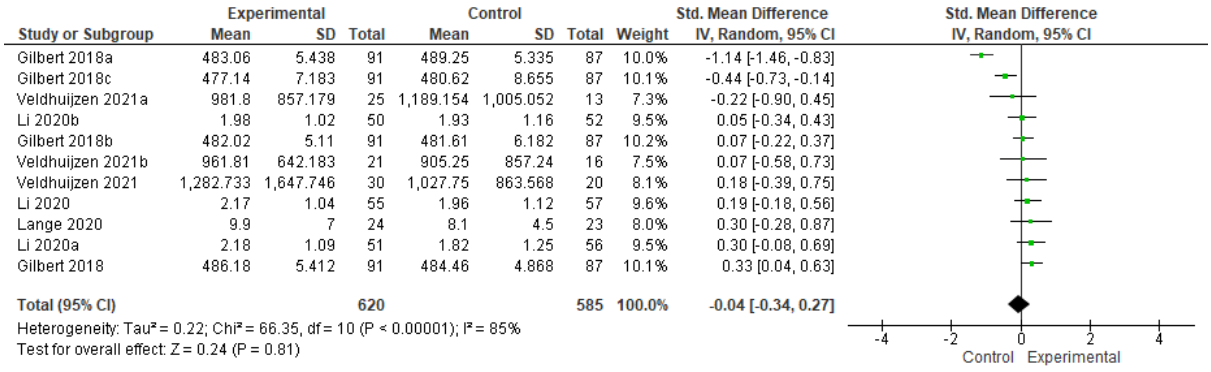
Supplementary Figure 2.9: The effects of interventions on sedentary time. SD= standard deviation, 95% CI= 95% confidence interval.



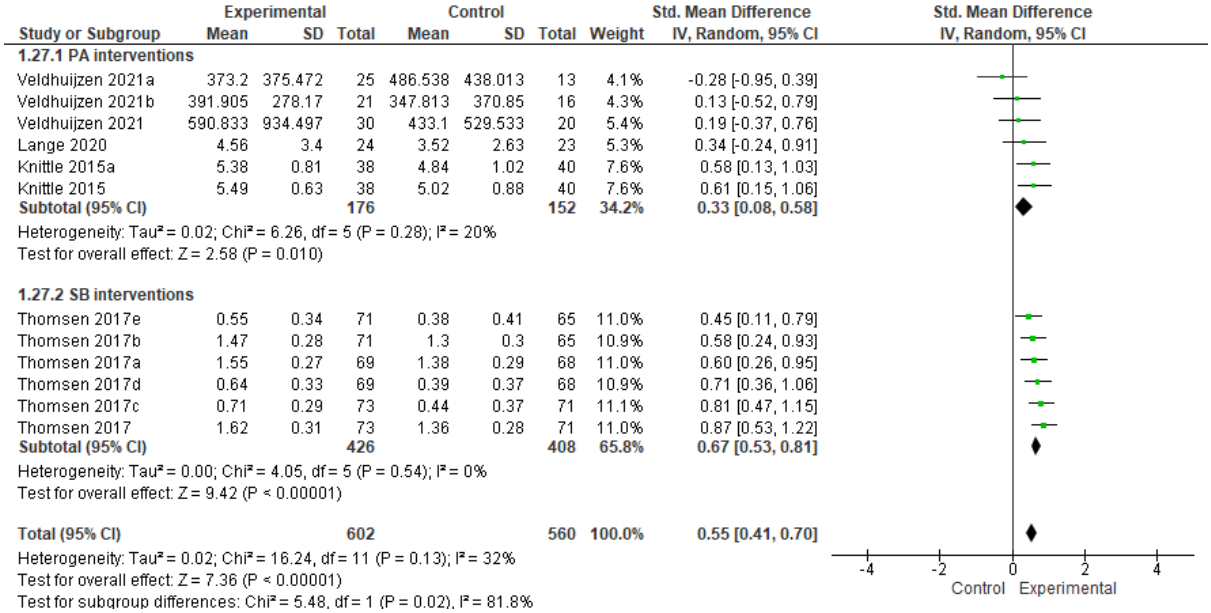
Supplementary Figure 2.10: The effects of interventions on daily steps. SD= standard deviation, 95% CI= 95% confidence interval.



Supplementary Figure 2.11: The effects of interventions on moderate to vigorous physical activity. SD= standard deviation, 95% CI= 95% confidence interval.

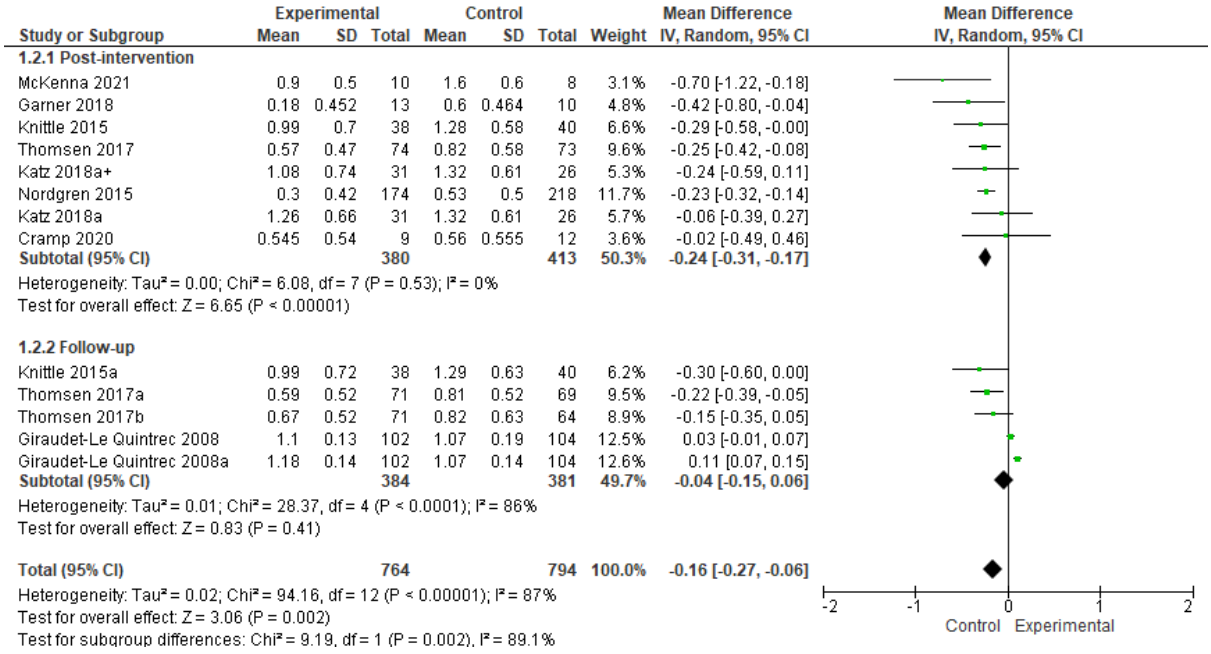


Supplementary Figure 2.12: The effects of interventions on total physical activity. SD= standard deviation, 95% CI= 95% confidence interval.

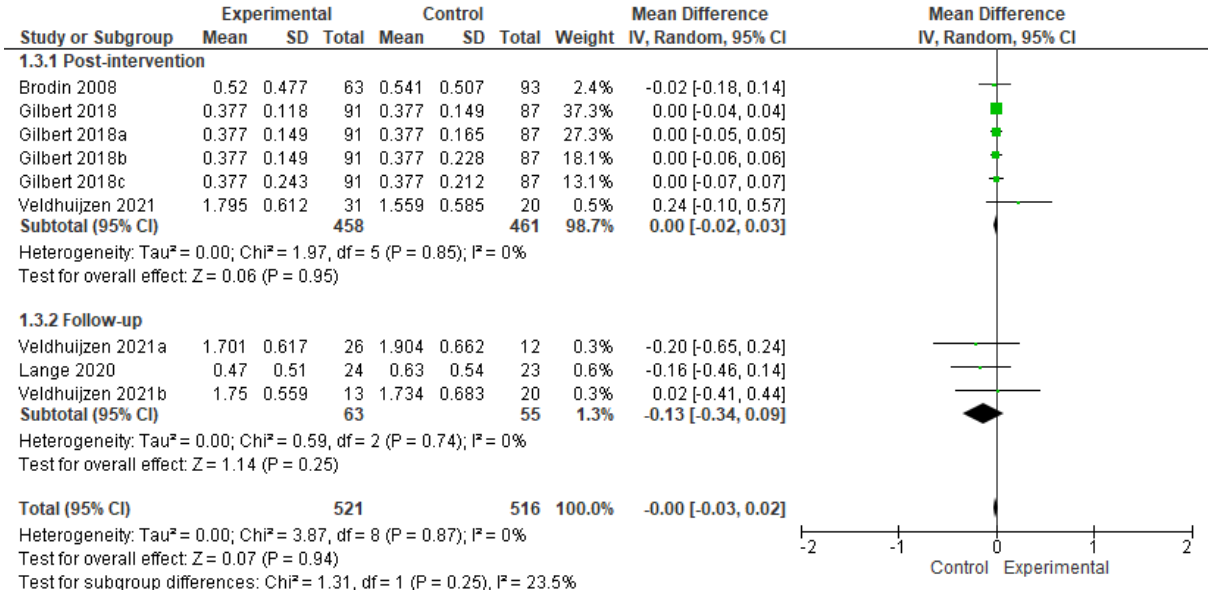


Supplementary Figure 2.13: The effects of interventions on leisure/light intensity physical activity. SD= standard deviation, 95% CI= 95% confidence interval.

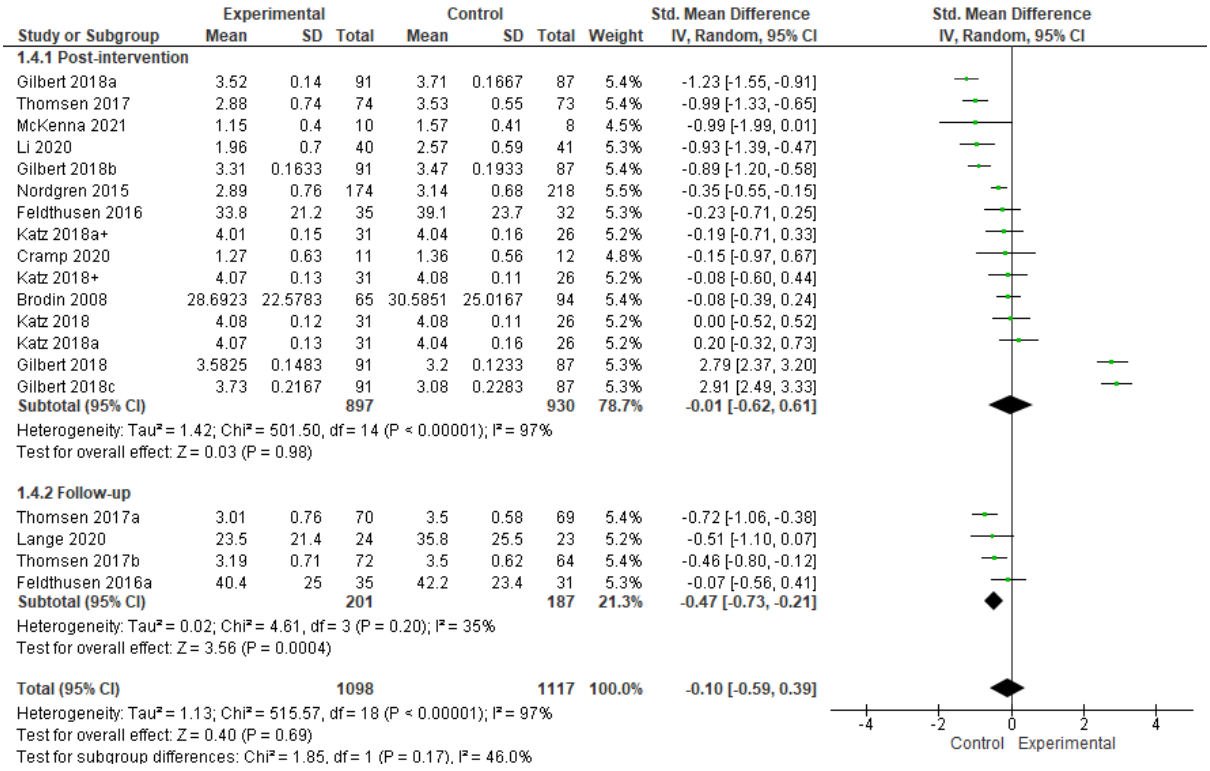
Supplementary Figures 2.14-2.26: Forest plots for secondary outcomes- Post-intervention vs follow-up



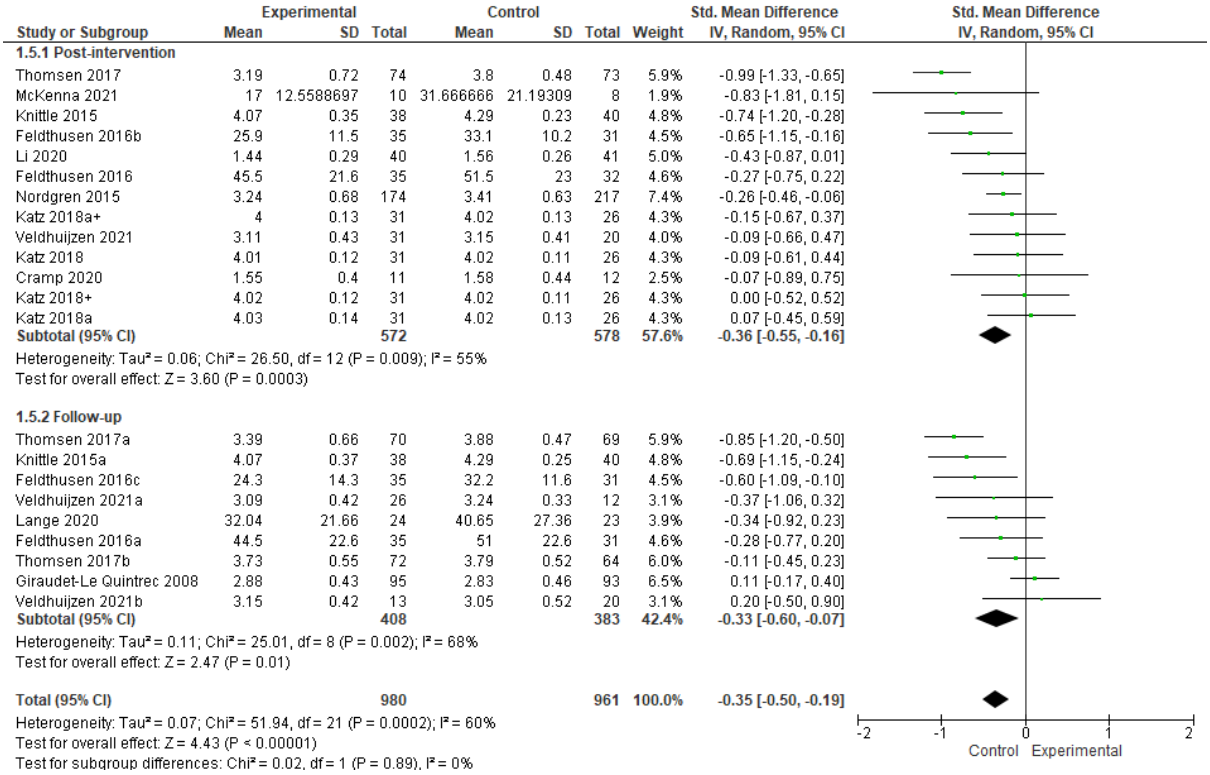
Supplementary Figure 2.14: The effects of interventions at post-intervention and follow-up for functional ability (normally distributed). SD= standard deviation, 95% CI= 95% confidence interval.



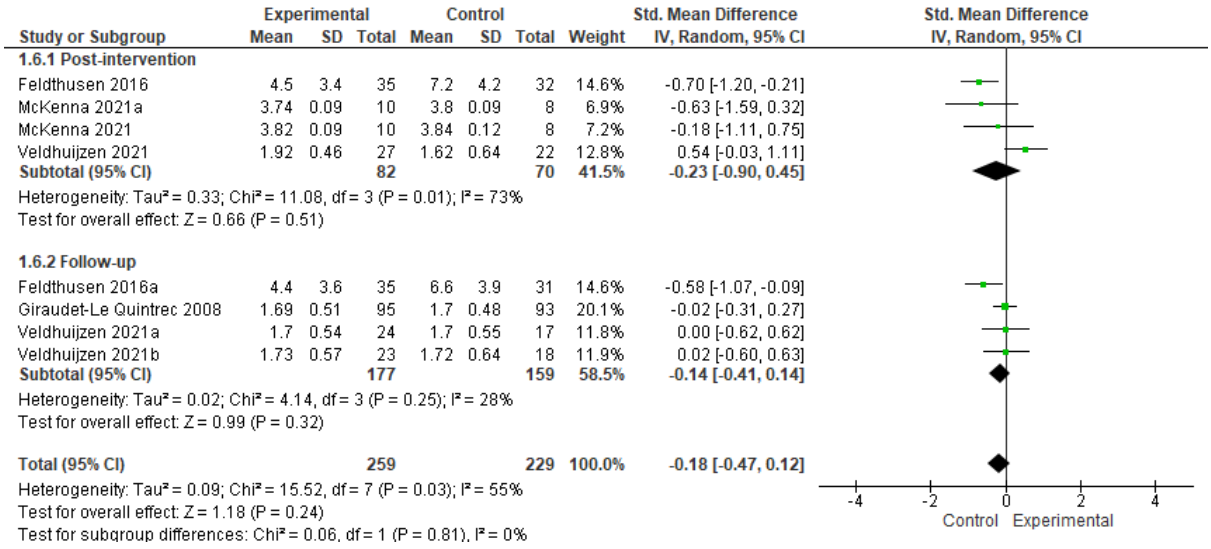
Supplementary Figure 2.15: The effects of interventions at post-intervention and follow-up for functional ability (non-normal). SD= standard deviation, 95% CI= 95% confidence interval.



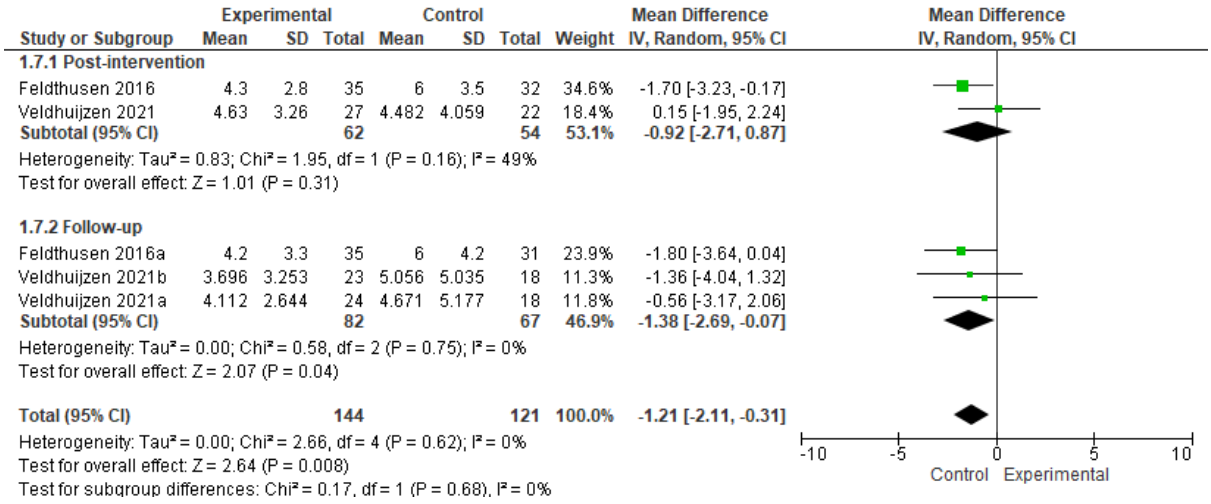
Supplementary Figure 2.16: The effects of interventions at post-intervention and follow-up for pain. SD= standard deviation, 95% CI= 95% confidence interval.



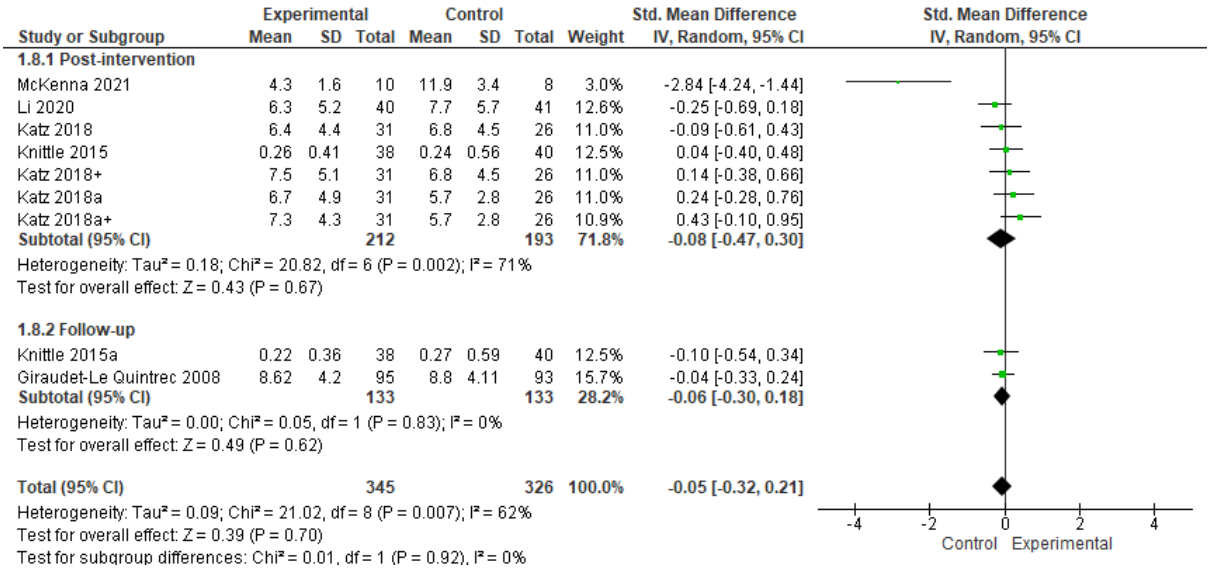
Supplementary Figure 2.17: The effects of interventions at post-intervention and follow-up for fatigue. SD= standard deviation, 95% CI= 95% confidence interval.



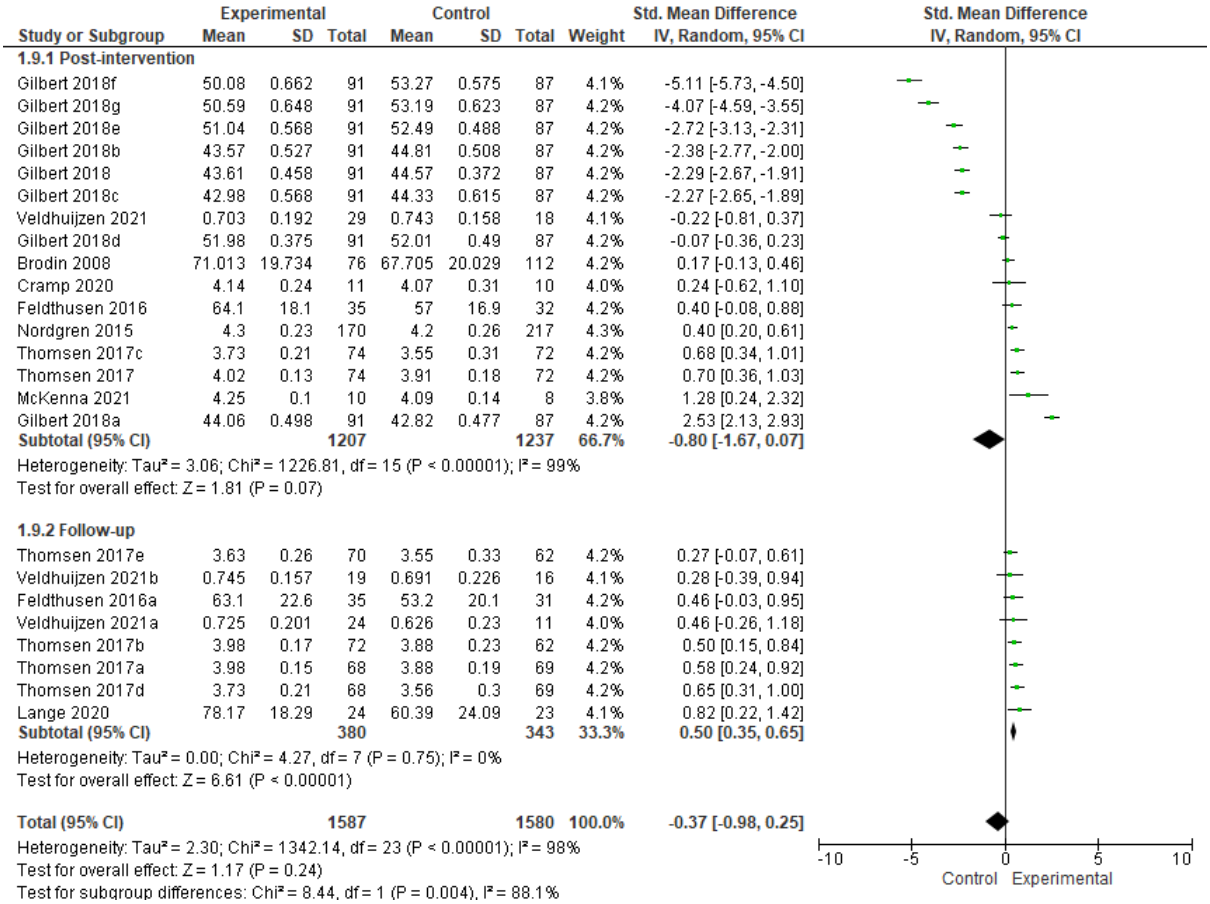
Supplementary Figure 2.18: The effects of interventions at post-intervention and follow-up for anxiety. SD= standard deviation, 95% CI= 95% confidence interval.



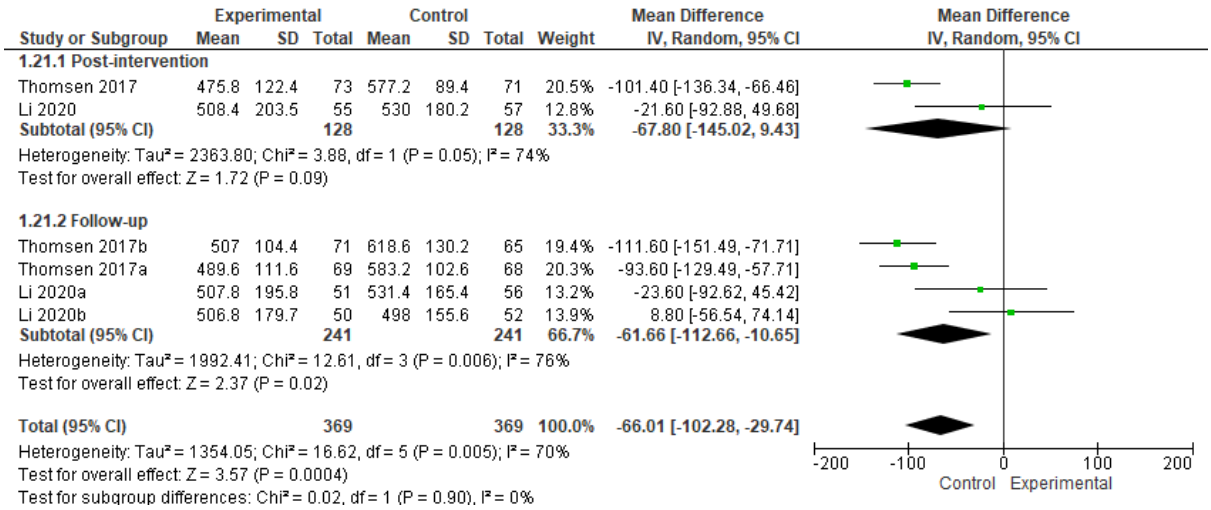
Supplementary Figure 2.19: The effects of interventions at post-intervention and follow-up for depression (non-normal). SD= standard deviation, 95% CI= 95% confidence interval.



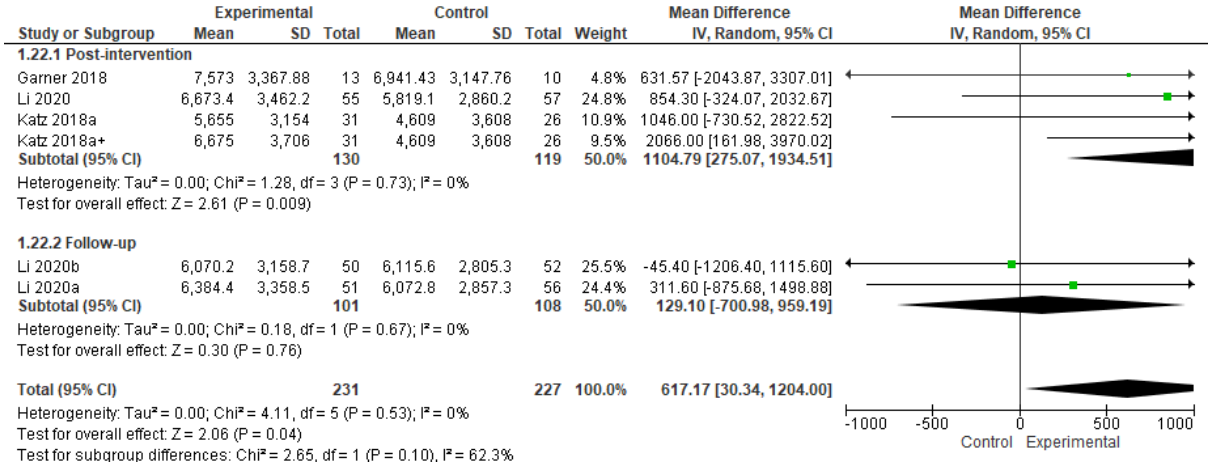
Supplementary Figure 2.20: The effects of interventions at post-intervention and follow-up for depression (normally distributed). SD= standard deviation, 95% CI= 95% confidence interval.



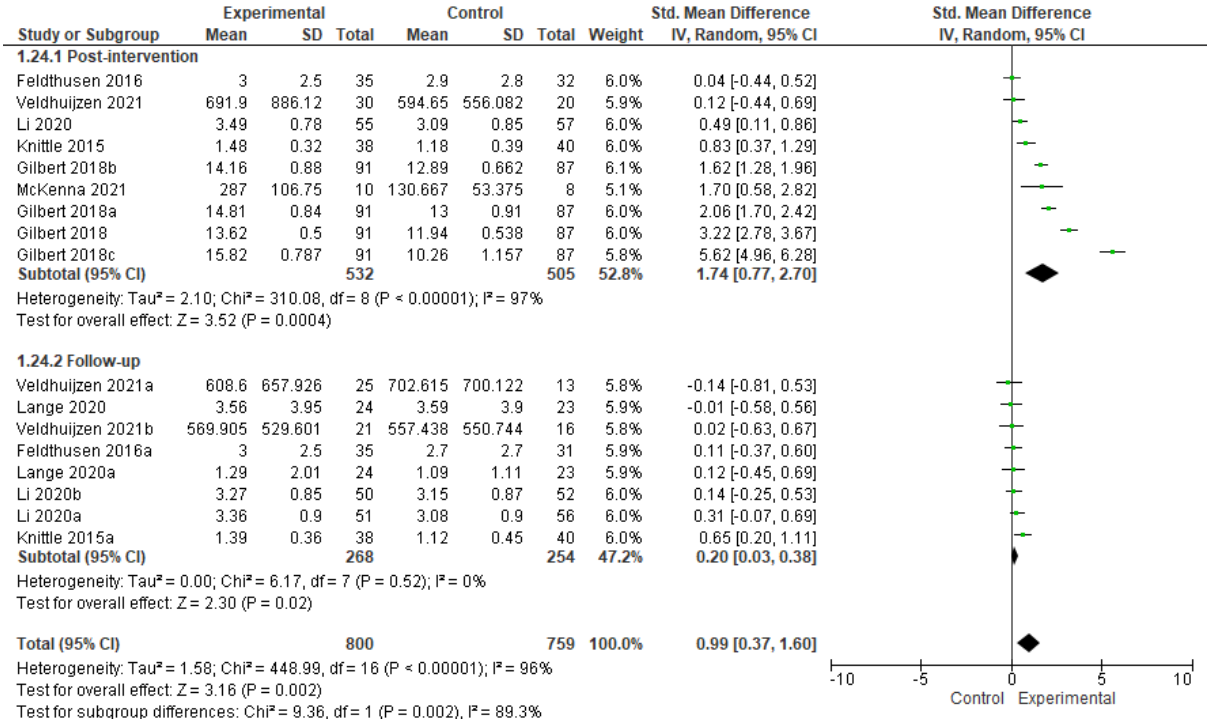
Supplementary Figure 2.21: The effects of interventions at post-intervention and follow-up for quality of life. SD= standard deviation, 95% CI= 95% confidence interval.



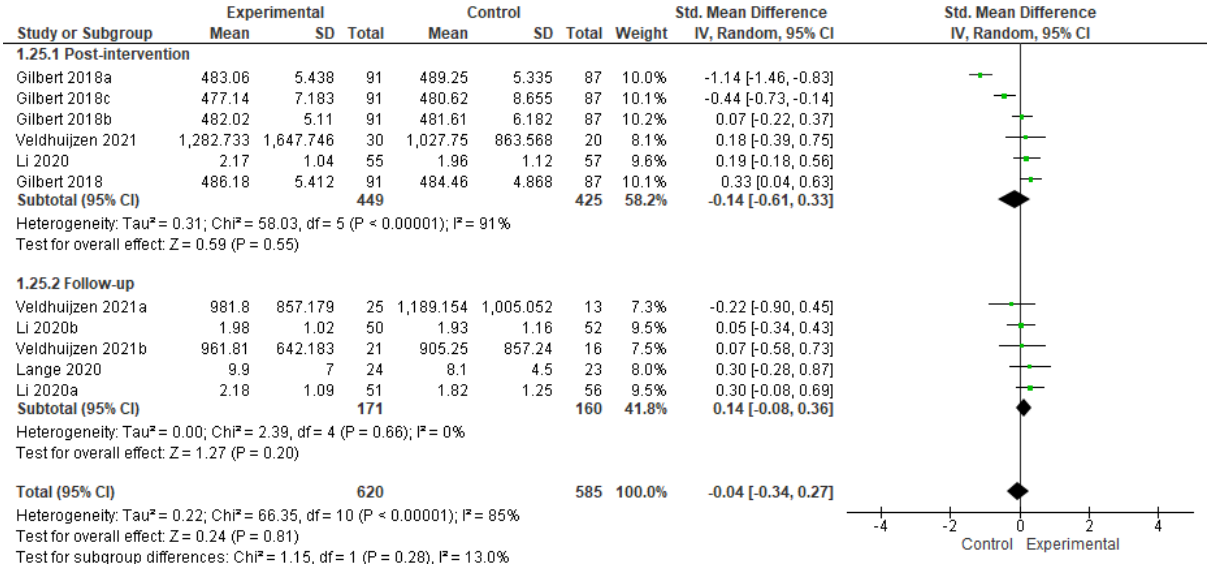
Supplementary Figure 2.22: The effects of interventions at post-intervention and follow-up for sedentary time. SD= standard deviation, 95% CI= 95% confidence interval.



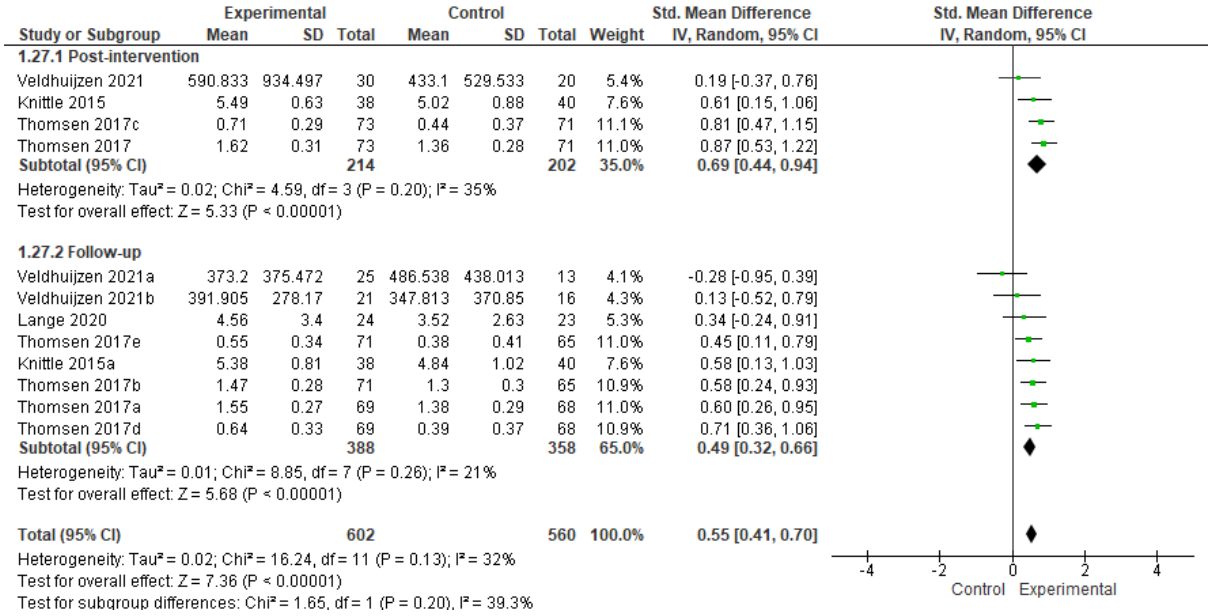
Supplementary Figure 2.23: The effects of interventions at post-intervention and follow-up for daily steps. SD= standard deviation, 95% CI= 95% confidence interval.



Supplementary Figure 2.24: The effects of interventions at post-intervention and follow-up for moderate to vigorous physical activity. SD= standard deviation, 95% CI= 95% confidence interval.



Supplementary Figure 2.25: The effects of interventions at post-intervention and follow-up for total physical activity. SD= standard deviation, 95% CI= 95% confidence interval.



Supplementary Figure 2.26: The effects of interventions at post-intervention and follow-up for leisure/light intensity physical activity. SD= standard deviation, 95% CI= 95% confidence interval.

CHAPTER 3: THE INTER-RATER AND TEST-RETEST RELIABILITY OF THREE MODALITIES OF QUANTITATIVE SENSORY TESTING IN HEALTHY ADULTS AND PEOPLE WITH CHRONIC LOWER BACK PAIN OR RHEUMATOID ARTHRITIS

Abstract

Quantitative Sensory Testing (QST) modalities used to assess central pain mechanisms require different protocols in people with different musculoskeletal conditions. We explored possible effects of musculoskeletal diagnosis and test site (rheumatoid arthritis (RA) n=18- lower leg, lower back pain (LBP) n=25- forearm, plus healthy controls n=20- lower leg, and n=25- forearm) on QST inter-rater and test-retest reliability. QST modalities used were Pressure Pain detection Threshold (PPT), Temporal Summation (TS) and Conditioned Pain Modulation (CPM). TS was calculated as difference or ratio of single and repeated punctate stimuli, and CPM used single or mean of multiple unconditioned PPTs. Intraclass correlation coefficients (ICC) were compared between different subgroups. High to very high reliability was found for all assessments of PPT at or between anatomical sites and participants (ICC≥ .77). Moderate to high reliabilities were found for TS (ICC= .64 to .88), both at the lower leg and the forearm. Reliability was higher when TS was calculated as a difference rather than a ratio. The ICCs of CPM indicated no to moderate reliability (ICC= .01 to .64) at leg or forearm, and in people with RA or LBP. In conclusion, PPT and TS are transferable tools to quantify pain sensitivity at different testing sites in different musculoskeletal diagnoses. Low apparent reliability of CPM protocols might indicate minute to minute dynamic pain modulation, rather than a stable trait of individuals with musculoskeletal pain.

Introduction

Pain is a multidimensional sensory experience. Quantification is essential to understand how pain signals are processed (Arendt-Nielsen & Yarnitsky, 2009). Quantitative Sensory Testing (QST) is an umbrella term for a battery of different psychophysical tissue-stimulation tests that provide important information about different types of pain processing (Arendt-Nielsen & Yarnitsky, 2009; Courtney et al., 2010) and peripheral or central sensitisation at/remote from sites of injured tissue (Arendt-Nielsen & Yarnitsky, 2009; Middlebrook et al., 2020). QST is used to explore central mechanisms underpinning local and global pain modulation in people with musculoskeletal conditions with high levels of chronic pain (Fingleton et al., 2015; Pavlakovic & Petzke, 2010), such as Rheumatoid Arthritis (RA) or Lower Back Pain (LBP) (da Rocha Castelar Pinheiro et al., 2013; Heiberg & Kvien, 2002; Joharatnam et al., 2015; Sokka et al., 2001).

There are several QST modalities, but there is incomplete consensus on what are the most appropriate QST research protocols, and there is limited standardisation of reporting in clinical populations (Rolke et al., 2006a). QST protocols may need to be adapted for specific diagnoses. For example, assessment of the lower or upper limb may be confounded by neuropathic features from lumber or cervical nerve root compression (Osborne et al., 2018). Furthermore, other aspects of diagnosis might influence QST outcomes, for example ongoing pain, joint distribution or inflammatory disease (Joharatnam et al., 2015; Osborne et al., 2018; Suokas et al., 2012).

QST can be categorised into: “static” (e.g. pressure pain detection threshold (PPT)) and “dynamic” (e.g. temporal summation (TS) and conditioned pain modulation (CPM)) modalities. Protocols can include tests using a pressure algometer to assess pain thresholds (PPT); a sensitivity assessment using repeated pressure stimulation on the skin (TS); and the use of conditioning stimuli alongside a test stimulus to modulate pain (CPM) (McWilliams & Walsh, 2017). In combination, these assessments provide insight into the presence of central integration (TS, CPM) and pain sensitivity

(PPT) through ascending and descending nociceptive pathways (Uddin & MacDermid, 2016). QST, therefore, is a valuable diagnostic assessment tool (Arendt-Nielsen & Yarnitsky, 2009). In LBP, QST has been used to investigate temporal changes in pain mechanisms and sensitivity (Marcuzzi et al., 2018), and the contribution of central components to pain (Corrêa et al., 2015; Imamura et al., 2013). In RA, QST has been used to characterise pain mechanisms affected during disease progression (Lee et al., 2018; Pavlakovic & Petzke, 2010).

Use of QST in research or clinical practice presumes measurement of a relatively stable characteristic with tools that will give the same result, irrespective of the assessor (inter-rater reliability) and when the test is repeated (test-retest reliability) (Hogan et al., 2000). QST reliability previously has been reported in healthy people (Chesterton et al., 2007; Graven-Nielson et al., 2015; Marcuzzi et al., 2017), and people with neuropathic or osteoarthritis pain (Middlebrook et al., 2020; Suokas et al., 2012). In RA, Lee et al. (2018) report low to high inter-rater reliability for PPT, TS and CPM, and strong associations with disease activity. In LBP, Paungmali et al. (2012) examined test-retest reliability of PPT conducted on the primary area of pain, thought to be largely influenced by peripheral sensitisation. However, central mechanisms might predominantly determine PPT at sites distant from the site of pathology (i.e., at forearm in people with LBP, and in lower leg in people with RA) (Suokas et al., 2012). Different diagnoses and research questions might therefore require QST to be undertaken at different sites. PPTs vary between body sites, possibly due to differences in innervation of subcutaneous tissues, or depth of overlying soft tissue (Hogeweg et al., 1996; Kosek et al., 1993). However, data are sparse comparing reliability of QST protocols that have been modified for different clinical populations (Gerecz-Simon et al., 1989; Lee et al., 2018; Paungmali et al., 2012). The few QST reliability studies in RA and LBP have been conducted on varied body sites, with little between- and within-study consistency, and report little information about the methodologies used (Geber et al., 2011; Georgopoulos et al., 2019; Lee et al., 2018; Paungmali et al., 2012). A standardised, reliable, QST protocol, which could be employed across multiple musculoskeletal

conditions, would enable the collection of more harmonious data (Middlebrook et al., 2020). Although the main population of interest of this thesis is in people with RA, the inclusion of participants with chronic LBP in this study enables the exploration of whether QST modality reliability and validity is transferable across different musculoskeletal diagnoses.

The primary aim (Aim 1) of this study was to evaluate the test-retest and inter-rater reliability of similar protocols of a) PPT, b) TS and c) CPM that had been adapted for use at different testing sites in different clinical populations. The secondary aim (Aim 2) was to define optimally reliable calculation methods for a) TS and b) CPM. Finally, Aim 3 of this study was to assess the validity of QST measures.

Methods

Participants

People living with RA were recruited from Rheumatology outpatient clinics at Russells Hall Hospital, Dudley Group NHS Foundation Trust. In addition, people living with LBP or RA were recruited from a list of participants who had already participated in an observational study at the University of Birmingham (RA^{leg}) or the University of Nottingham ($LBP^{forearm}$) and had consented to be contacted about future studies. Healthy individuals ($Healthy^{leg}$ for comparison with RA^{leg} and $Healthy^{forearm}$ for comparison with $LBP^{forearm}$) were recruited to assess and compare the reliability of QST modalities when conducted at different testing sites. Healthy individuals affiliated with the School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham or the School of Medicine, Department of Academic Rheumatology at the University of Nottingham (students, lecturers, senior academics) were invited to participate. Written informed consent was obtained from all individuals before their participation in the study.

Inclusion criteria for *healthy individuals* were: adults (≥18 years old), being currently healthy (defined as having no acute or chronic pain at any part of their body), and understanding the English language. Exclusion criteria were: diagnosed with acute or chronic pain, participation in an NHS rehabilitation program and pregnancy. Inclusion criteria for *patients* were: adults (≥18 years old), physician diagnosis of RA (for RA^{leg} group) or chronic LBP (for $LBP^{forearm}$ group), and be able to speak and understand English. Patients were excluded if unable to give informed consent due to cognitive impairment or otherwise, had history of additional co-morbidities (e.g., cancer, diabetes neuropathies, fractures and/or other conditions) which cause greater disability than their RA or back pain, or were pregnant.

Favourable ethical opinions were granted from the University of Birmingham Ethics Committee, Black Country Regional Ethics Committee (ERN: 16/WM/0371), Faculty of Medicine & Health Sciences Research Ethics Committee of the University of Nottingham (ERN: 264-1803) and the East Midlands - Nottingham 1 Research Ethics Committee of the Health Research Authority (REC: 18/EM/0049).

Study Procedures

To participate, individuals with RA (RA^{leg}) visited Russells Hall Hospital, Dudley, and individuals with LBP ($LBP^{forearm}$) visited the Back Pain Unit of the King's Mill Hospital, Sutton-in-Ashfield. Healthy participants visited the School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham ($Healthy^{leg}$) or Academic Rheumatology, University of Nottingham ($Healthy^{forearm}$) to take part. For all groups, each participant undertook two QST testing sessions (baseline/follow-up) separated by at least a week, and up to 3 weeks apart. These timeframes were considered appropriate periods of time between sessions to minimise learning effects and reduce the risk of potential recall bias for self-reported data (Middlebrook et al., 2020). Baseline and follow-up examinations in both the healthy and patient cohort were performed by the same researcher (Rater 1 for $Healthy^{leg}$ and RA^{leg} participants, Rater 3 for $LBP^{forearm}$ and $Healthy^{forearm}$ participants). As part of

the baseline session, a second rater (Rater 2) also performed identical examinations to Rater 1 and Rater 3 in both healthy groups only ($Healthy^{leg}$ and $Healthy^{forearm}$) in order to assess inter-rater reliability. All participants completed the protocol in full, with a mean baseline to follow-up period of 8 days for $Healthy^{forearm}$ and $LBP^{forearm}$ participants, 13 days for $Healthy^{leg}$ participants and 12 days for RA^{leg} participants.

Raters 1, 2 and 3 were fully trained on how to conduct the QST modalities, and procedures were standardised and rehearsed to ensure they used the same language when interacting with participants, and employed the same measurement techniques when administering protocols. All sessions took place in the same temperature-controlled room (18-20 °C). All participants were advised and encouraged to maintain their routines and daily activities throughout the testing period. The full QST protocol used by all raters can be found in **Appendix 3.1**.

Quantitative Sensory Testing

The QST protocol comprised both "static" (PPT) and "dynamic" (TS and CPM) modalities (Arendt-Nielsen et al., 2018; Rolke et al., 2006a; Yarnitsky et al., 2015). For $Healthy^{leg}$ and RA^{leg} participants, testing was on the tibialis anterior muscle (5cm distal to the tibial tuberosity and knee joint) (for PPT and CPM modalities), and 5cm above the patella on the skin above the rectus femoris (for TS) of both legs. For $Healthy^{forearm}$ and $LBP^{forearm}$ participants, the testing site was 5cm distal from the lateral epicondyle of individual's elbow joint on both arms, corresponding with the body of the brachioradialis muscle.

Pressure Pain Threshold: For measuring PPT, an electronic handheld algometer (Medoc-AlgoMed Advanced Medical Systems – Computerised Pressure Algometer, Israel) was used connected to a laptop. The laptop displayed the amount of applied pressure from the algometer on the screen. Increasing pressure with an $1cm^2$ rubber probe of the handheld algometer was applied over the tibialis anterior of the participant's dominant leg ($Healthy^{leg}$ and RA^{leg} participants) or the

radiobrachialis of the participants' non-dominant forearm (Healthy^{forearm} and LBP^{forearm} participants) at a rate of 50kPa/sec (Rolke et al., 2006a). Each participant was asked to press a button on a device held in their dominant hand as soon as the sensation of pressure started to become painful. The procedure was initially applied for familiarisation purposes on the opposite leg or arm. Data were collected from 3 repeats of the PPT (Rolke et al., 2006a).

Temporal Summation: The TS was assessed by repeated application of a stimulus using the retractable blunt needle of a specially manufactured pen (256mN Pinprick, MRC-Systems, Germany). The participants maintained a relaxed position and a single stimulus with the blunt needle was applied to the skin approximately 5cm above the patella over the rectus femoris (Healthy^{leg} and RA^{leg} participants) or 5cm distal to the lateral epicondyle over the radiobrachialis muscle of their dominant forearm (Healthy^{forearm} and LBP^{forearm} participants). This was followed by ten repetitive stimuli at a rate of 1/sec (Arendt-Nielsen & Yarnitsky, 2009). Immediately after the single stimulus, each participant was asked to rate the experienced intensity of pain or sharpness on a 0 to 10 Numerical Rating Scale (NRS) (Healthy^{leg} and RA^{leg} participants) or a 10cm Visual Analogue Scale (VAS) (Healthy^{forearm} and LBP^{forearm} participants) where the lowest and highest extremes signified no pain/sharpness and worst imaginable pain/sharpness respectively. After the 10 stimuli, they were asked to rate the average intensity of pain or sharpness on the same scales. For familiarisation purposes, TS testing was initially applied on the opposite leg or arm, and the testing procedure was collected from 2 repeats of the single and 10 stimuli.

Conditioned Pain Modulation: For the purposes of CPM testing, participants' unconditioned PPT was assessed for all participants in an identical way as described for PPT testing. The participants' conditioned PPT was assessed again while pain was induced in their non-dominant (Healthy^{leg} and RA^{leg} participants) or dominant (Healthy^{forearm} and LBP^{forearm} participants) arm by application of a 15cm wide blood pressure cuff and induction of ischaemic pain. The cuff was inflated above systolic

pressure to occlude arterial blood flow to the arm, and participants repeatedly squeezed a small foam ball (conditioning stimulus). Once pain had reached 4/10 rating, the conditioned PPT was performed in the dominant leg (Healthy^{leg} and RA^{leg} participants) or non-dominant arm (Healthy^{forearm} and LBP^{forearm} participants). Immediately after the conditioned PPT, the pressure cuff was released.

Data Reduction

Participants' PPT was taken as the arithmetic mean of 3 replicate measurements (PPT^{mean}) with low PPT indicating greater pain sensitivity. TS pain was calculated as the difference (subtraction between the score of the single stimulus and the average pain experienced during the ten subsequent stimuli, TS^{WUD}). The mean of the two TS^{WUD} values was used for analysis. CPM (CPM$^{PPT\text{-}mean}$) was taken to be the single conditioned PPT measurement (PPT^{Con}) minus the arithmetic mean of all the replicated unconditioned PPT measurements from the study visit (PPT^{mean}) (Yarnitsky, 2010; Yarnitsky et al., 2015).

The windup ratio, TS^{WUR} , was also calculated to define optimally reliable calculation methods for TS (Aim 2a), as average pain during the 10 stimuli divided by pain rating of single stimulus. In both TS^{WUD} and TS^{WUR} a larger positive value of TS indicated greater sensitivity. CPM (CPM^{Unc}) was also calculated (to define optimally reliable calculation methods for CPM (Aim 2b)) using the single conditioned PPT measurement (PPT^{Con}) minus the interim unconditioned PPT measurement (the single measure taken immediately before the conditioning stimulus, PPT^{Unc}). In both calculation methods (CPM$^{PPT\text{-}mean}$, CPM^{Unc}) a lower value indicated higher pain sensitivity (less efficient descending inhibition) (Marcuzzi et al., 2018).

Data Analysis

Sample size calculations for this study were performed with type I and type II errors as .05 to .20 respectively (Walter et al., 1998). Considering that each substudy comprised two different sessions, featuring one measurement for each modality within each session, and with minimally accepted

reliability of ρ=.5 and expected reliability of ρ=.8 (Micalos et al., 2009; Rhudy & France, 2007), the minimum sample size was calculated to be 22 participants (Manresa et al., 2011).

Data normality was assessed by Shapiro-Wilk normality testing. Where data distributions differed significantly from normal, non-parametric statistical tests were used. TS^{WUD} and TS^{WUR} distributions in $Healthy^{leg}$ participants, TS^{WUR} in RA^{leg} participants, and all PPT, TS and CPM variables in $Healthy^{forearm}$ $LBP^{forearm}$ participants significantly differed from normality and were positively skewed. Non-parametric statistical analyses were conducted for these modalities, whereas parametric tests were conducted for all other modalities.

The following analyses were used to assess the reliability of all QST modalities as part of Aims 1 and 2 of this study. To assess differences between variables, paired samples t-tests (paired normal data) and Wilcoxon signed-rank tests (paired non-normal data) were performed. Unpaired t-tests (normal data) and Mann-Whitney U tests (non-normal data) were conducted to examine if there were differences between the participant groups.

The test-retest and inter-rater reliability of the PPT, TS and CPM modalities were established using methods that focused on the measurement of reliability (Bisset et al., 2015; Manresa et al., 2014; Middlebrook et al., 2020; Suokas et al., 2012; Walton et al., 2011). For each separate modality, a two-way random effects absolute agreement model for single measures was used to measure the inter-rater reliability (rater 1 and 2 or rater 3 and 2) for the healthy groups as well as the test-retest reliability for the single rater (rater 1 for RA^{leg} participants and rater 3 for $LBP^{forearm}$ participants) for both patient groups. A single measures intraclass correlation coefficient (ICC) with 95% confidence intervals (95% CI) were reported to express each reliability. For interpretation purposes, reliability values (ICC) of <.5 low correlation, .50 - .74 moderate correlation, .75 - .9 high correlation and >.90 very high correlation (Portney & Watkins, 2009). Further analysis involved determining the statistical significance between ICCs from separate disease populations (RA^{leg} compared to $LBP^{forearm}$

populations). These were derived from testing differences in variances using F-distributions (Feldt et al., 1987).

Bland-Altman analysis was conducted to give a visual representation of the data, identify outliers, and allow identification of systematic differences between measurements for each outcome. Plots show the mean difference (mean bias) between the two measurements, and 95% limits of agreement (LoA):
(mean difference of raters ± 2 × standard deviations of the difference between raters) (Manresa et al., 2014). An even distribution of points across the Bland Altman plots indicates no systematic bias (Bland & Altman, 1999).

To test validity of the QST measures (Aim 3) and explore the accuracy of QST conducted in different populations and at different body sites, correlations between modalities were conducted (Pearson correlation coefficient for normal data and Spearman correlation coefficient for non-normal data). In addition, correlation tests were assessed between modalities with participant age, and exploration of differences in QST modalities between sexes (using t-tests (normal data) and Mann-Whitney U tests (non-normal data)). Finally, differences between diagnoses for scoring of each modality were investigated using t-tests and Mann-Whitney U tests.

Data were analysed using IBM SPSS Version 26 and R version 3.4.2, and $p \leq .05$ was used to indicate of statistical significance.

Results

Participant characteristics are displayed in **Table 3.1**. Overall, the study groups for QST at the forearm comprised n=25 $Healthy^{forearm}$ and n=25 $LBP^{forearm}$ participants. The study groups for QST at the lower leg were n=18 RA^{leg} and n=20 $Healthy^{leg}$ participants. Healthy participants were younger than disease

groups, with no significant differences observed between healthy and diseased groups for sex (**Table 3.1**).

Table 3.1: Characteristics of the participants

	Healthy^{leg}	***RA^{leg}***	***Healthy^{forearm}***	***LBP^{forearm}***
N	20	18	25	25
Age median (IQR) years	26 (23 to 32) [a]	58 (55 to 65) [a]	31 (28 to 46) [b]	57 (48 to 65) [b]
Sex (n= female (%))	10 (50.0)	13 (72.2)	15 (60)	17 (68)

Note: IQR = interquartile range, QST = Quantitative Sensory Testing

[a]=significant difference between Healthy^{leg} and RA^{leg} participants in demographic data, determined by independent samples t-tests (age) and Chi-square tests (sex), [b]= significant difference between Healthy^{forearm} and LBP^{forearm} participants in demographic data, determined by Mann Whitney U tests (age) and Chi-square tests (sex). ($p<.05$)

Aim 1a

At baseline, PPT measurements were similar between raters and between test-retest replicates (**Table 3.2**). The inter-rater and test-retest ICCs for PPT were between .77 and .95, being classified as high to very high at the forearm and very high reliability at the lower leg (**Table 3.3**). Bland-Altman plots supported the reliability of PPT between measurements (**Figure 3.1a, 3.1b, 3.2a, 3.2b, 3.3a, 3.3b** and **Supplementary Table 3.1**). The ICCs for inter-rater reliability were similar between lower leg and forearm, except that the test-retest ICC for PPT was significantly higher in Healthy^{leg} population (ICC= .95) compared to the Healthy^{forearm} population (ICC= .77, $F(19,24)= 4.6$, $p<.001$), although ICCs both demonstrated high to very high reliability.

Aim 1b

Baseline TS^{WUD} measurements were similar between raters at both forearm and lower leg, and were similar in test-retest replicates. Although, RA^{leg} showed a significant change over time ($z = -2.32$, $p=.02$, **Table 3.2**). ICCs for inter-rater and test-retest ranged from .64 and .88, displaying moderate to high reliability at the lower leg, and a high reliability at the forearm (**Table 3.3**). Bland-Altman plots (**Figure 3.1c, 3.1d, 3.2c, 3.2d, 3.3c, 3.3d** and **Supplementary Table 3.1**) confirmed the reliability of

TS^{WUD}. The ICCs for inter-rater reliability of TS^{WUD} were statistically similar between lower leg and forearm.

Aim 1c

Baseline $CPM^{PPT\text{-}mean}$ showed no differences in measurements between raters and in test-retest reliability (**Table 3.2**). The ICCs for $CPM^{PPT\text{-}mean}$ were heterogeneous with values between .01 and .64, classified within the range of no to moderate reliability. For $CPM^{PPT\text{-}mean}$ Bland Altman plots, LoA between measurements from raters were generally wide (**Figure 3.1e, 3.1f, 3.2e, 3.2f, 3.3e, 3.3f** and **Supplementary Table 3.1**). No statistically significant differences were found for ICCs for $CPM^{PPT\text{-}Mean}$ between the lower leg and forearm.

Aim 2a

In order to define the optimal measurement method for TS, reliability of windup ratio using TS^{WUR} was assessed. Measurements of windup ratio showed rater 2 reported higher TS^{WUR} at baseline than rater 3 in $Healthy^{forearm}$ participants (median rater 3= 2.5, rater 2= 3.6, z= -2.46, p=.01); and TS^{WUR} was higher at follow-up than at baseline in $Healthy^{leg}$ (median baseline= 1.7, follow-up= 2.0, z= -2.27, p=.02, **Table 3.2**). The ICC reliability of TS^{WUR} appeared heterogeneous between study populations, with $Healthy^{leg}$ inter-rater and $Healthy^{forearm}$ test-retest showing low reliability (**Table 3.3**). Other measurements of TS^{WUR} showed moderate to high reliability at both leg and forearm sites (**Table 3.3**). The test-retest and inter-rater Bland-Altman plots appeared to show greater variability at larger values of TS^{WUR}, particularly in disease populations (**Supplementary Figure 3.1a-3.1f**). The ICCs for inter-rater reliability were statistically similar for TS^{WUR} between lower leg and forearm, with the exception of inter-rater reliability for $Healthy^{leg}$ (ICC= .21) and $Healthy^{forearm}$, (ICC= .71, $F(24,18)= 2.7$, $p<.001$).

Aim 2b

Baseline CPM^{Unc} showed statistically similar measurements between raters and in test-retest reliability (**Table 3.2**) but also displayed heterogeneous ICC values in healthy adults at both lower leg and forearm (ICC= .19 to .71) (**Table 3.3**). The CPM^{Unc} measures also showed no test-retest reliability in either RA or LBP (ICC= −.02 and −.10, respectively) (**Table 3.3**). Bland-Altman plots for CPM^{Unc} are shown in **Supplementary Figure 3.2a–3.2f**. Comparisons of the 2 different measurements of PPT that were used to calculate CPM scores (PPT^{Mean} and PPT^{Unc}) showed high to very high reliability (ICC= .71 to .98) and no statistically significant differences across all different populations, body sites and timepoints (data not shown). No statistically significant differences were found for ICCs of CPM^{Unc} between the lower leg and forearm.

Aim 3

Correlations between modalities demonstrated that a higher PPT was associated with a lower TS^{WUD} in people with RA and LBP, a higher CPM^{Unc} in $Healthy^{leg}$ participants, and higher $CPM^{PPT\text{-}mean}$ in $Healthy^{forearm}$ participants (**Supplementary Table 3.2**). Participant age was not significantly correlated with QST outcomes for most modalities (**Supplementary Table 3.3**). Significantly lower PPT was reported by female participants for all rater 1 comparisons at the tibialis anterior (lower leg) and at baseline for rater 3 at the brachioradialis (forearm) (**Supplementary Table 3.4**). $LBP^{forearm}$ participants had a higher rater 3 baseline TS^{WUR} than $Healthy^{forearm}$ participants (Mann-Whitney U= 200.00, p=.03) (**Table 3.2**). In addition, when compared to $Healthy^{leg}$ participants, RA^{leg} participants had lower rater 1 baseline CPM^{Unc} (t = 2.35, p=.02) and higher follow-up TS^{WUD} (Mann-Whitney U = 110.50, p=.04) (**Table 3.2**).

Table 3.2: QST measurements of all participants at baseline and follow-up

	Quantitative Sensory Testing	Baseline			Follow up	
		$Healthy^{leg}$		RA^{leg}	$Healthy^{leg}$	RA^{leg}
		Rater 1	*Rater 2*	*Rater 1*	*Rater 1*	*Rater 1*
Lower leg	***PPT (kPa)***	483.0 (259.9 to 689.3)	441.8 (281.8 to 567.5)	333.0 (232.7 to 488.6)	498.5 (269.2 to 688.0)	310.2 (173.1 to 650.2)
	TS^{WUD} ***(-10 to 10)***	1.0 (0.5 to 1.9)	1.3 (0.6 to 1.5)	1.5 (0.5 to 2.1) [b]	1.1 (1.0 to 2.0) [c]	2.6 (0.9 to 3.6) [b c]
	TS^{WUR} ***(Ratio)***	1.7 (1.2 to 2.2) [b]	1.7 (1.3 to 2.0)	2.0 (1.3 to 5.1)	2.0 (1.5 to 4.5) [b]	2.9 (2.1 to 4.5)
	$CPM^{PPT\text{-}mean}$ ***(kPa)***	76.2 (7.9 to 204.9)	117.6 (53.6 to 167.4)	67.3 (22.3 to 159.3)	133.9 (54.5 to 202.7)	93.1 (36.8 to 193.7)
	CPM^{Unc} ***(kPa)***	122.0 (26.3 to 219.5)[c]	107.3 (56.1 to 178.7)	74.0 (-27.6 to 106.1) [c]	103.9 (53.7 to 208.8)	95.6 (-2.0 to 211.2)
		Baseline			**Follow up**	
		$Healthy^{forearm}$		$LBP^{forearm}$	$Healthy^{forearm}$	$LBP^{forearm}$
		Rater 3	*Rater 2*	*Rater 3*	*Rater 3*	*Rater 3*
Forearm	***PPT (kPa)***	222.0 (176.9 to 249.5)	206.3 (147.0 to 275.4)	271.5 (195.5 to 305.3)	224.0 (178.4 to 251.9)	216.5 (164.6 to 281.6)
	TS^{WUD} ***(-10 to -10)***	1.2 (0.5 to 2.2)	1.4 (0.5 to 2.2)	1.5 (0.5 to 2.5)	0.9 (0.3 to 2.0)	1.3 (0.4 to 2.3)
	TS^{WUR} ***(Ratio)***	2.5 (1.9 to 3.8) [a c]	3.6 (2.0 to 5.4) [a]	5.0 (2.3 to 9.5) [c]	2.6 (1.7 to 4.6)	3.5 (2.1 to 7.5)
	$CPM^{PPT\text{-}mean}$ ***(kPa)***	87.2 (50.4 to 119.9)	109.3 (42.1 to 173.0)	55.2 (24.2 to 91.8)	66.6 (36.9 to 131.0)	62.7 (31.0 to 99.3)
	CPM^{Unc} ***(kPa)***	92.1 (37.2 to 163.6)	120.5 (30.3 to 213.6)	47.0 (-6.9 to 98.0)	55.9 (5.9 to 95.0)	38.2 (11.8 to 81.4)

Note: Data are presented as median (IQR). [a] = Paired samples t-test (normal) or Wilcoxon signed-rank test (non-normal) demonstrating significant difference between baseline measurements from rater 1 or 3 with rater 2 in healthy participants ($p<.05$). [b] = Paired samples t-test (normal) or Wilcoxon signed-rank test (non-normal) demonstrating significant difference between baseline and follow-up measurements ($p<.05$). [c] = Independent samples t-test (normal) or Mann-whitney U test (non-normal) demonstrating significant differences in QST modalities between healthy and diseased participants ($p<.05$).
PPT = Mean Pressure-Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, TS^{WUR}: Temporal Summation calculated as a ratio, $CPM^{PPT\text{-}mean}$ = Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus, CPM^{Unc} = Conditioned Pain Modulation where a unique PPT measurement was used as an unconditioned stimulus, kPa= kilopascals.

Table 3.3: Inter-rater and test-retest reliability in all participants

	Lower leg		
	Healthy^{leg}		***RA^{leg} patients***
	Inter-rater (Rater 1** - **Rater 2**) (n=20)	***Test-Retest (Rater 1**) (n=20)***	***Test-Retest (Rater 1**) (n=18)***
	ICC (95% CI)	***ICC (95% CI)***	***ICC (95% CI)***
PPT	.92 (.82, .97)	.95 (.88, .98)	.94 (.84, .98)
TS^{WUD}	.82 (.60, .92)	.80 (.57, .92)	.64 (.20, .85)
TS^{WUR}	.21 (-.22, .59)	.74 (.45, .89)	.77 (.49, .91)
CPM$^{PPT\text{-}mean}$	.01 (-.45, .46)	.64 (.30, .84)	.11 (-.34, .53)
CPM^{Unc}	.19 (-.29, .58)	.71 (.39, .87)	-.02 (-.40, .41)
	Forearm		
	Healthy^{forearm}		***LBP^{forearm} patients***
	Inter-rater (Rater 3** - **Rater 2**) (n=25)	***Test-Retest (Rater 3**) (n=25)***	***Test-Retest (Rater 3**) (n=25)***
	ICC (95% CI)	***ICC (95% CI)***	***ICC (95% CI)***
PPT	.86 (.72, .94)	.77 (.54, .89)	.92 (.83, .96)
TS^{WUD}	.88 (.75, .94)	.76 (.52, .89)	.78 (.56, .86)
TS^{WUR}	.71 (.45, .86)	.48 (.11, .73)	.71 (.44, .86)
CPM$^{PPT\text{-}mean}$	.46 (.09, .72)	.43 (.06, .70)	.44 (.07, .71)
CPM^{Unc}	.55 (.21, .77)	.50 (.15, .74)	-.10 (-.44, .27)

Note: Intraclass Correlation Coefficient (ICC) with 95% confidence intervals (CI) are presented. PPT= Pressure Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, TS^{WUR}= Temporal Summation calculated as a ratio $CPM^{PPT\text{-}mean}$= Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus, CPM^{Unc}= Conditioned Pain Modulation where a unique PPT measurement was used as an unconditioned stimulus.

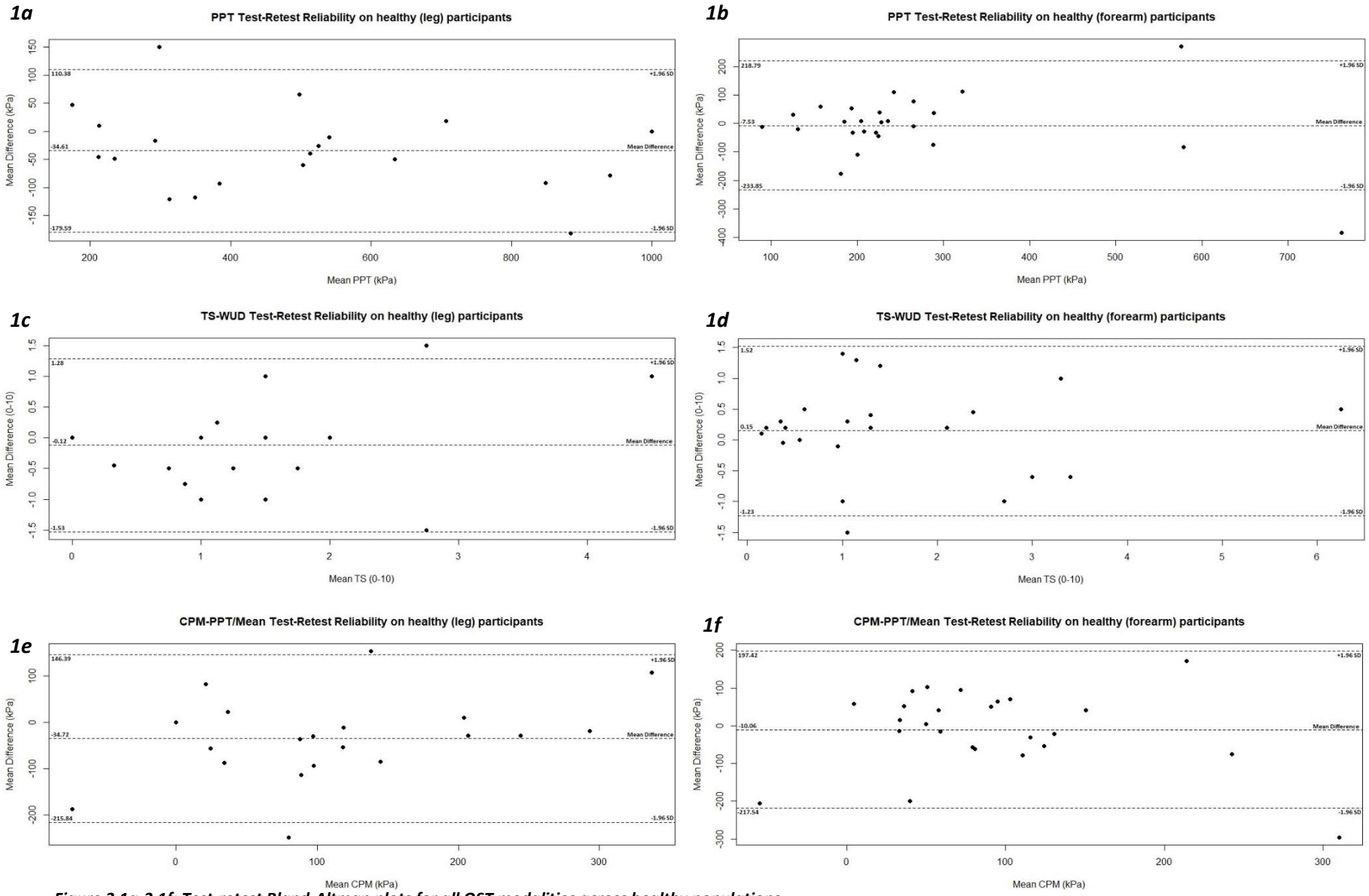


Figure 3.1a-3.1f: Test-retest Bland-Altman plots for all QST modalities across healthy populations

Note: PPT= Pressure Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, $CPM^{PPT\text{-}mean}$= Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus, kPa= kilopascals.

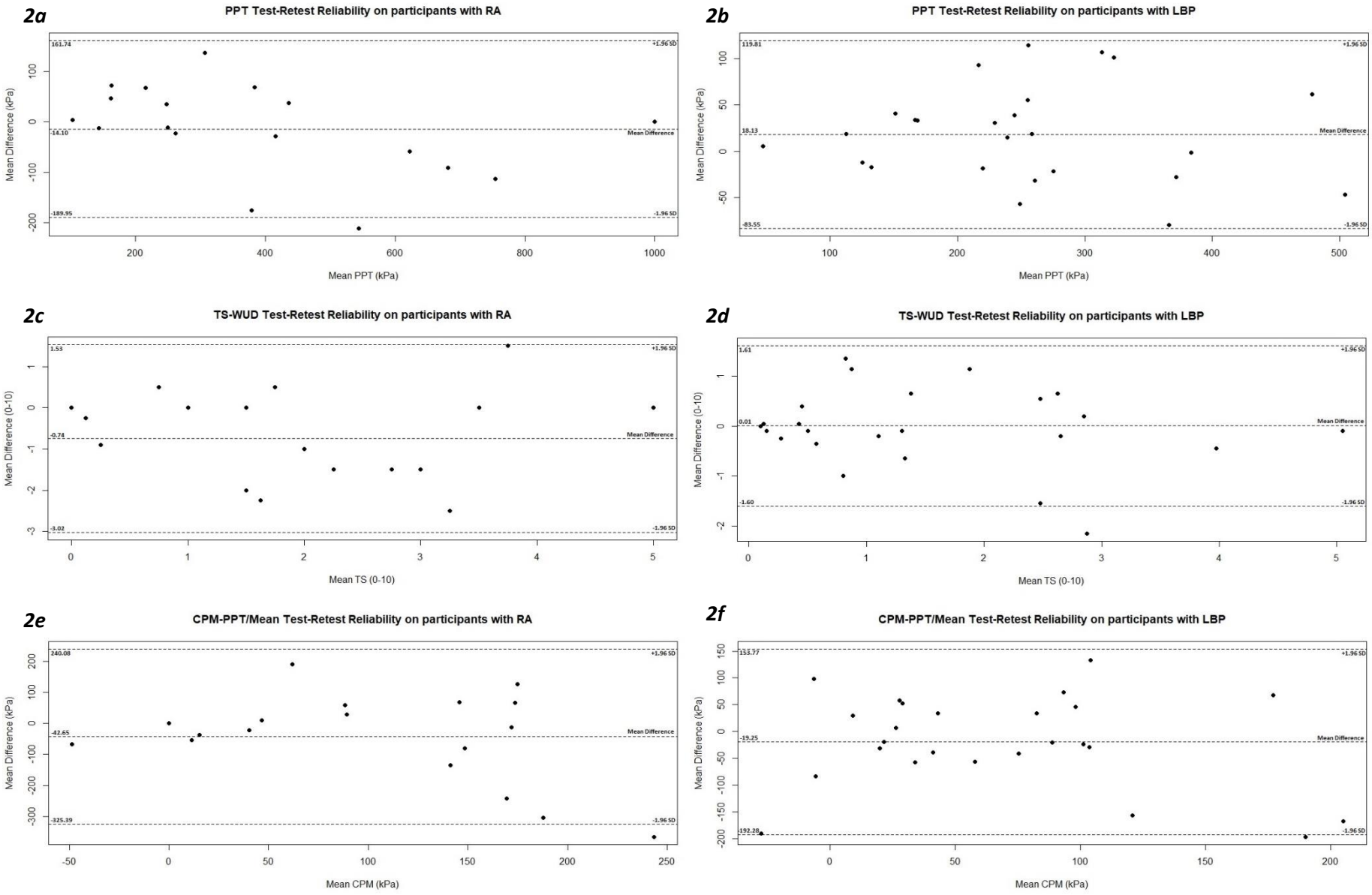


Figure 3.2a-3.2f: Test-retest Bland-Altman plots for all QST modalities across RA^{leg} and LBP^{forearm} populations

Note: PPT= Pressure Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, CPM$^{PPT\text{-}mean}$= Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus, kPa= kilopascals.

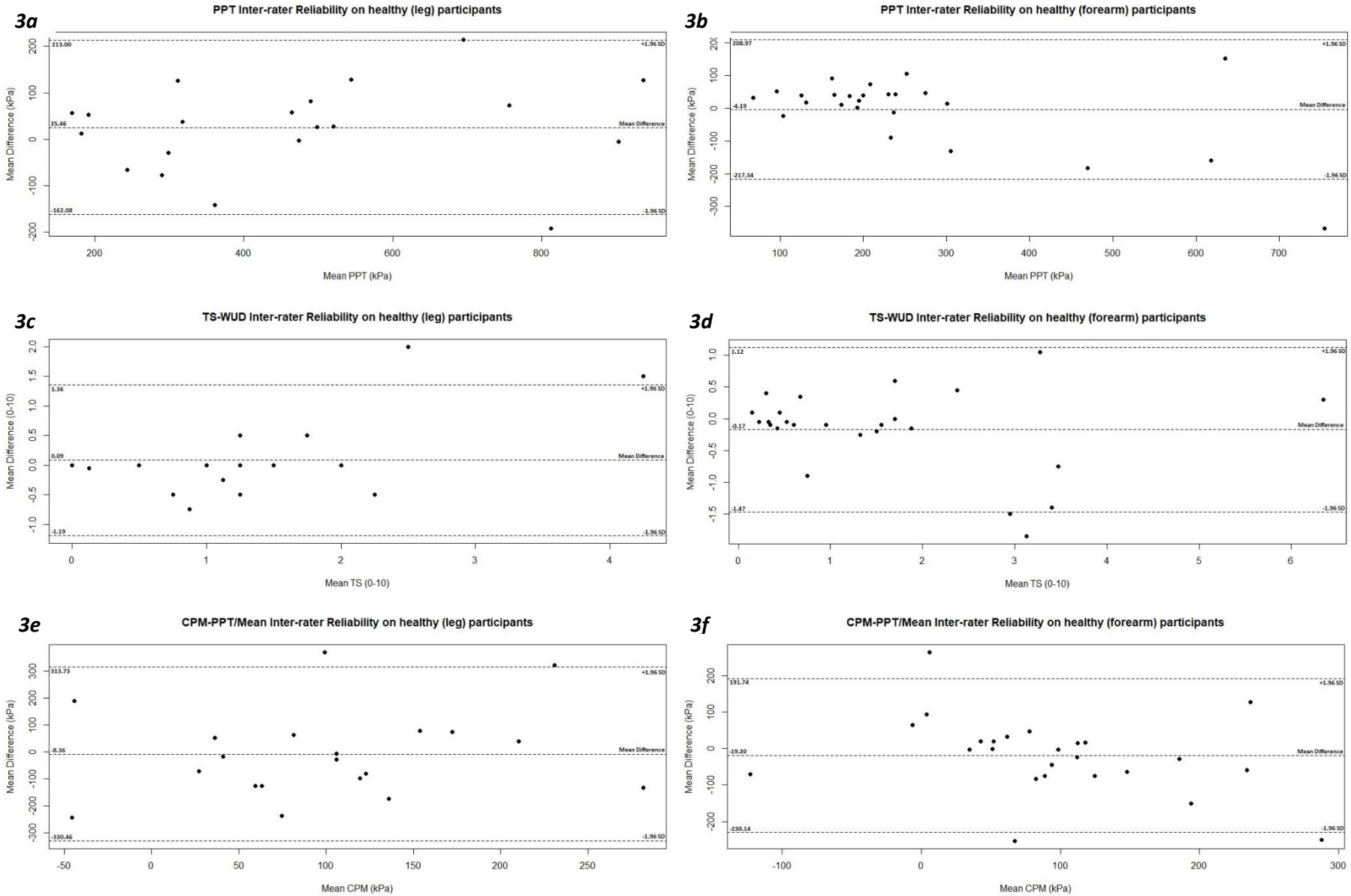


Figure 3.3a-3.3f: Inter-rater Bland-Altman plots for all QST modalities across healthy populations

Note: PPT= Pressure Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, $CPM^{PPT\text{-}mean}$= Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus. kPa= kilopascals.

Discussion

Aim 1

This study found that PPT and TS appeared to be reliable modalities to measure aspects of central pain processing in healthy participants and patients alike. CPM demonstrated heterogeneous reliability in different participants and anatomical sites. For example, reliability of CPM was not present in RA^{leg} or $LBP^{forearm}$ patients, which both displayed no to low reliability, with no to moderate reliability displayed for test-retest and inter-rater ICCs in healthy participants. The results of this study suggest that application of both static and dynamic QST modalities (PPT and TS) within a single assessment session can yield reliable data to quantify some aspects of pain sensitivity and central sensitisation in healthy participants and in individuals with a chronic musculoskeletal disorder.

We found that PPT appeared to be the most reliable QST modality across populations, timepoints, and raters. This study extends previous studies that have demonstrated high reliability (ICC= .75 to .94) of PPT in healthy participants (Chesterton et al., 2007; Chung et al., 1992; Fabio Antonaci, 1998; Nussbaum & Downes, 1998; Park et al., 2011), at different timepoints in healthy individuals (10 minutes to 6 hours) (Chesterton et al., 2007; Pelfort et al., 2015), in patients with knee osteoarthritis and neuropathic pain (Geber et al., 2011; Wylde et al., 2011), as well as in individuals with RA or LBP (Lee et al., 2018; Paungmali et al., 2012). Bland-Altman plots in this study illustrated narrow LoA, indicating little variability between PPT measurements in all sets of participants. PPT had previously been found as the most reliable QST modality in healthy and patient populations (Wylde et al., 2011). We now extend those findings to show similar results in healthy, RA, and LBP participants, with PPT conducted at different body sites, suggesting that this may be a transferable and generalisable finding. My study used a longer gap between test-retest sessions than previous studies (Suokas et al., 2012; Wylde et al., 2011), and the very high level of test-retest reliability over 2-3 weeks suggests

that pain pressure sensitivity is a highly stable trait. ICCs for PPT at the different testing sites were all within the high to very high reliability categories, despite significantly differing between sites.

Findings from TS in this study are consistent with existing evidence that finds this dynamic modality to be sensitive and reliable to assess centrally driven hypersensitivity in patient populations in general, and RA or chronic LBP in particular (Arendt-Nielsen et al., 2018). TS^{WUD} test-retest and inter-rater reliability were also moderate to high in this study, demonstrated across raters, populations, and timepoints. When the TS^{WUD} was calculated, reliability at the forearm was consistently rated as high. Findings of this study are consistent with past research in healthy participants that also demonstrated moderate to high TS test-retest reliability (ICC= .67 to .87) (Cathcart et al., 2009; Graven-Nielson et al., 2015; Kong et al., 2013). However, some studies in healthy participants showed low test-retest (ICC= .43) and inter-rater reliability (ICC= .41) (Pigg et al., 2010). No differences were observed between ICCs for TS^{WUD}, at different sites, or between people with RA, LBP or healthy individuals, and all ICCs were within the moderate to very high reliability category. Both PPT and TS at both anatomical sites therefore appear to be reliable metrics.

Test-retest and inter-rater ICCs for $CPM^{PPT\text{-}mean}$ showed no to moderate reliability, when calculated with the mean PPT value as an unconditioned stimulus. CPM is a dynamic modality, attempting to measure descending inhibition, which is distinct from TS. Abandoning the assessment of CPM may lead to the loss of important information about pain mechanisms that are not captured elsewhere. However, my findings suggest that care should be taken to optimise CPM protocols and ensure validity throughout studies. My findings reinforce the heterogeneity of previous reports of CPM in healthy participants (ICC= .60- .82) (Lewis et al., 2012a), people with chronic LBP (ICC= .59) (Martel et al., 2013), shoulder pain (ICC= .54) (Valencia et al., 2014), and chronic pancreatitis (ICC= .10) (Olesen et al., 2012). However, it should be noted that CPM has been assessed using substantially different methodologies across different studies. The synthesis of multiple measurements and participant self-

assessments into a single value will add to the variability of CPM, so it is not surprising that its reliability can be much lower. It is also possible that the underlying mechanisms involved in CPM measures might be less stable ("more dynamic") than the other QST modalities. Therefore, differences between observations reflect real changes in descending pain modulation.

Altered CPM is a characteristic of populations with chronic pain (Kennedy et al., 2016; Lewis et al., 2012b; O'Brien et al., 2018a; Yarnitsky, 2010), although it is unclear whether that is due to a fully activated endogenous inhibition or a reduced ability to modulate pain. Nevertheless, obtaining CPM reliability may be elusive (Kennedy et al., 2016), and establishing the association between CPM responses and clinical manifestations of pain merits further investigation (Fernandes et al., 2019). There may be a narrow window between tolerable and intolerable pain that is breached by application of a conditioning stimulus, particularly for clinical populations experiencing chronic pain. Previous studies have shown that when a test stimulus has become intolerable, poor CPM test-retest reliability is observed (Olesen et al., 2012). Nevertheless, the magnitude of CPM measured in this study (PPT^{Con}–PPT^{Unc}) in patient and healthy participants is similar to those previously reported with patient populations and healthy controls (Corrêa et al., 2015; Owens et al., 2015). The variability and fluctuating nature of musculoskeletal pain (Gooberman-Hill et al., 2007), and the subjective nature of pain perception (Wylde et al., 2011), may each contribute to low CPM ICCs. In addition, due to its requirement to measure pain processing on two different occasions through identical pain indices, CPM is subject to more unknown confounders than other modalities.

Aim 2

This study also provided evidence that TS might be more reliable and valid if calculated as a difference (TS^{WUD}) rather than a ratio (TS^{WUR}). This should be considered when adopting the recommendations of the German Research Network on Neuropathic Pain (Rolke et al., 2006a; Rolke et al., 2006b). Conceptually, TS may describe the excitability of spinal cord neurons as it plateaus

after frequent stimulation (Rolke et al., 2006a), and can be easily utilised in routine clinical assessment (Rolke et al., 2006b). Distortion from low denominator values may adversely affect statistical properties of TS^{WUR}, although we also found that TS^{WUD} distributions differed significantly from normality. However, TS^{WUD} appeared more consistently reliable than TS^{WUR} with respect to ICC (particularly in healthy populations) and Bland Altman plots. There was also a lack of significant correlation between TS^{WUR} with other QST indices of central pain processing (**Supplementary Table 3.2**). This suggests a statistical and methodological advantages of TS^{WUD} over TS^{WUR}.

My findings indicate that where the unconditioned stimulus for CPM is the same as that used elsewhere in the QST protocol (PPT^{mean}), its repeat at the beginning of CPM testing in order to obtain a CPM^{Unc} value (PPT^{Unc}) may not be necessary. Indeed, multiple testing with painful stimuli may itself modulate central pain processing (as observed during TS), and increased sensitivity developing during the test protocol may lead stimuli to approach the pain tolerance threshold. Therefore, by forfeiting the interim PPT stimulus, this could minimise the chances of moderated patient response to stimuli by increased or intolerable pain. In addition, When CPM was calculated with an unconditioned stimulus in patient populations (CPM^{Unc}), the test-retest reliability was negative, indicating no correlation between timepoints using this method, similar to past test-retest findings (ICC= −.40) with ischaemic pain as the conditioning stimulus (Lewis et al., 2012a). As a result, $CPM^{PPT\text{-}mean}$ demonstrates statistical, methodological, and application advantages over CPM^{Unc}.

Aim 3

Significant associations were demonstrated between PPT and TS^{WUD} in patient populations, confirming the inter-association of these modalities as measures of central sensitisation. The weakness of associations between CPM with other modalities in patient populations suggests that CPM measures different aspects of pain processing, but might also reflect the low reliability of CPM measurement. Central sensitisation results from multiple processes, and different QST modalities

might reflect different aspects of central sensitisation, rather than each being estimates of a shared 'central sensitisation' phenomenon. Patients with RA and LBP have reduced PPTs, increased TS and deficient CPM (Marcuzzi et al., 2018; Müller et al., 2019), suggesting changes at multiple levels in pain processing pathways.

Limitations

Although this study had strengths from use of shared protocols across sites with multiple researchers, it is subject to a number of limitations. The small sample size may have increased uncertainty in the ICC estimates, reducing power to elucidate important differences between groups. Inter-rater reliability was evaluated in healthy participants, but not in the patient populations, due to the chronic pain prevalent in these groups, to avoid subjecting participants to two consecutive QST protocols in one session. QST involves complex procedures influenced by interacting variables, and future research might explore mechanisms related to anatomical site, diagnosis and assessor which underlie observed differences in reliability. Although we studied diverse pain populations, extension of my findings to other chronic pain diagnoses would require further research.

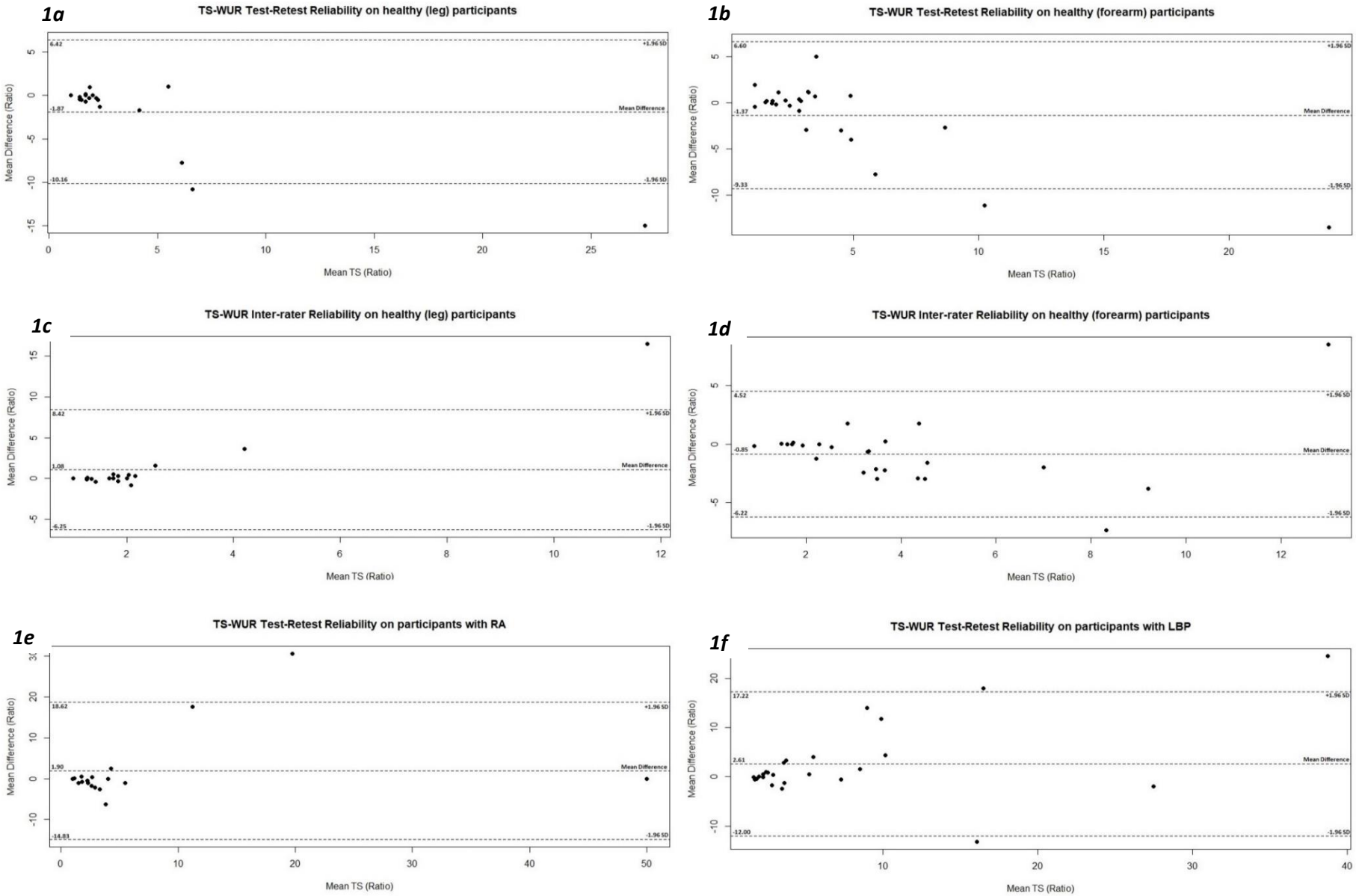
Conclusion

To conclude, a QST protocol consisting of PPT and TS is a reliable form of assessing and quantifying pain mechanisms in healthy participants and patients with RA and LBP, when assessed on either the forearm or leg. Further research is needed to improve reliability of CPM, in larger samples and different musculoskeletal populations, and to confirm the findings of this study.

Declarations

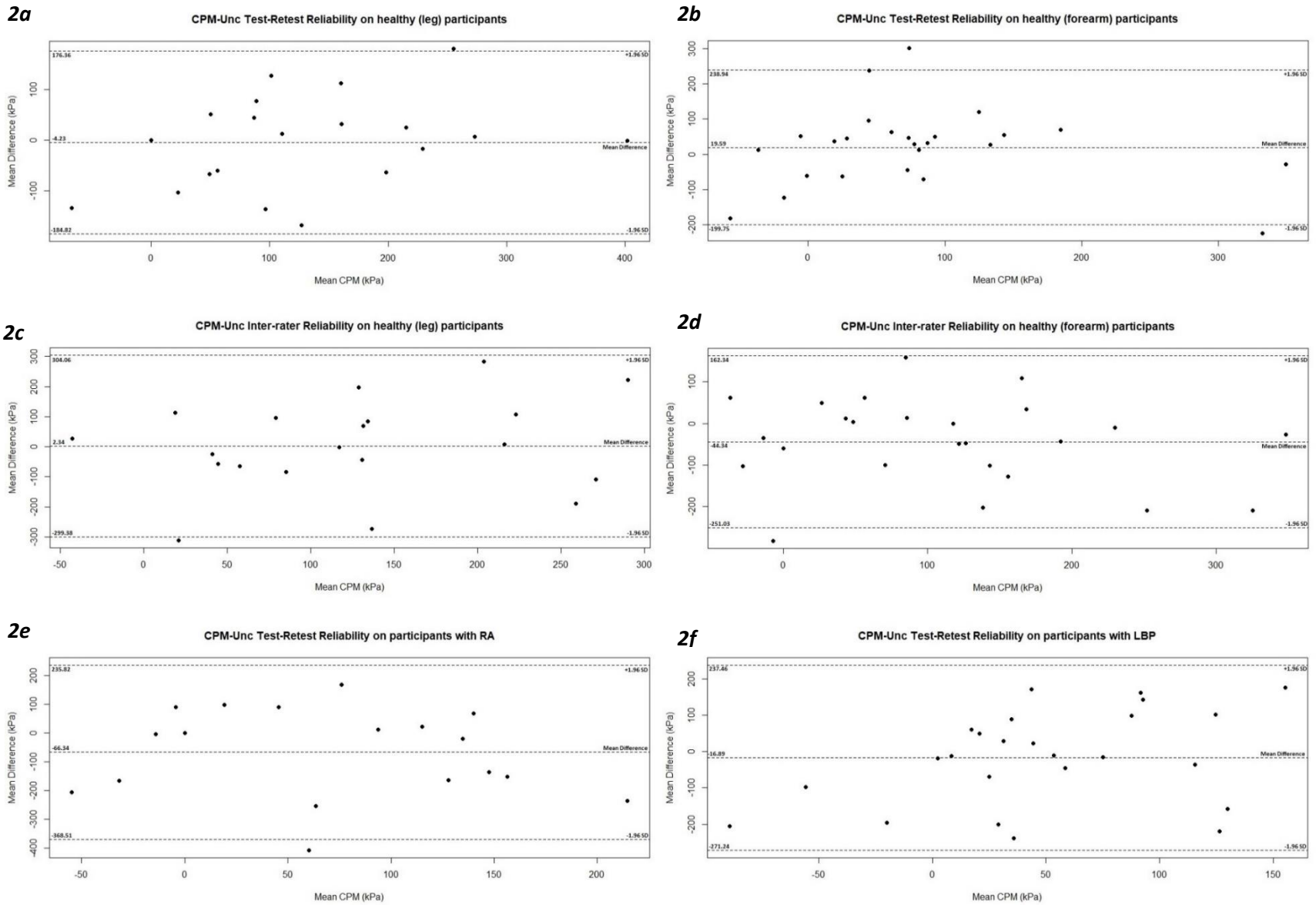
AUTHOR CONTRIBUTORSHIP: Conception and design of the study: Sophia Brady, Sally Fenton, David Walsh, Daniel McWilliams, Vasileios Georgopoulos, Jet Veldhuijzen van Zanten. Data acquisition: Sophia Brady, Daniel McWilliams, Vasileios Georgopoulos. Data analysis: Sophia Brady, Daniel

McWilliams, Vasileios Georgopoulos, David Walsh. Data interpretations and drafting of manuscript: all authors. Final approval of manuscript: all authors.



Supplementary Figure 3.1a-3.1f: Alternate TS calculation method (TS^{WUR}) across populations and raters

Note: TS^{WUR}= Temporal summation calculated as a wind-up ratio, kPa= kilopascals.



Supplementary Figure 3.2a-3.2f: Alternate CPM calculation method (CPM^{Unc}) across populations and raters

Note: CPM^{Unc} = Conditioned Pain Modulation where a unique PPT measurement was used as an unconditioned stimulus, kPa= kilopascals.

Supplementary Table 3.1: Mean differences and limits of agreement for test-retest and inter-rater reliability based on Bland Altman analysis results

		Test-retest reliability			***Inter-rater reliability***		
		Mean difference	*-LoA (95% CI)*	*+LoA (95% CI)*	*Mean difference*	*-LoA (95% CI)*	*+LoA (95% CI)*
Healthy^{leg}	***PPT***	-34.60	-179.59 (-239.56, -119.63)	110.38 (50.41, 170.34)	25.46	-162.08 (-239.65, -84.52)	213.00 (135.44, 290.57)
	TS^{WUD}	-0.12	-1.53 (-2.11, -0.95)	1.28 (0.70, 1.87)	0.09	-1.19 (-1.72, -0.66)	1.36 (0.83, 1.89)
	TS^{WUR}	1.08	-6.25 (-9.28, -3.22)	8.42 (5.38, 11.45)	-1.87	-10.16 (-13.58, -6.72)	6.42 (2.99, 9.84)
	CPM$^{PPT\text{-}mean}$	-34.72	-215.84 (-290.74, -140.93)	146.39 (71.49, 221.30)	-8.37	-330.46 (-463.67, -197.25)	313.73 (180.52, 446.94)
	CPM^{Unc}	-4.23	-184.82 (-259.51, -110.13)	176.36 (101.67, 251.05)	2.34	-299.38 (-424.16, -174.59)	304.06 (179.27, 428.84)
Healthy^{forearm}	***PPT***	-7.53	-233.85 (-316.41, -151.30)	218.79 (136.23, 301.34)	-4.19	-217.34 (-295.10, -139.59)	208.97 (131.22, 286.72)
	TS^{WUD}	0.15	-1.23 (-1.73, -0.73)	1.52 (1.02, 2.02)	-0.17	-1.47 (-1.94, -1.00)	1.12 (0.65, 1.60)
	TS^{WUR}	-1.36	-9.33 (-12.24, -6.43)	6.60 (3.69, 9.50)	-0.85	-6.22 (-8.18, -4.26)	4.52 (2.56, 6.48)
	CPM$^{PPT\text{-}mean}$	-10.06	-217.55 (-293.22, -141.86)	197.42 (121.73, 273.10)	-19.20	-230.14 (-307.09, -153.20)	191.74 (114.80, 268.69)
	CPM^{Unc}	19.59	-199.75 (-279.76, -119.74)	238.94 (158.93, 318.95)	-44.34	-251.03 (-326.42, -175.64)	162.34 (86.95, 237.74)
RA^{leg}	***PPT***	-14.10	-189.15 (-267.23, -112.68)	161.74 (84.47, 239.02)			
	TS^{WUD}	-0.74	-3.02 (-4.02, -2.02)	1.53 (0.53, 2.53)			
	TS^{WUR}	1.90	14.83 (-22.17, -7.48)	18.62 (11.27, 25.97)			
	CPM$^{PPT\text{-}mean}$	-42.65	-325.39 (-449.64, -201.14)	240.09 (115.84, 364.33)			
	CPM^{Unc}	-66.34	-368.51 (-501.29, -235.72)	235.82 (103.03, 368.61)			
LBP^{forearm}	***PPT***	18.13	-83.55 (-120.64, -46.46)	119.81 (82.72, 156.90)			
	TS^{WUD}	0.01	-1.60 (-2.18, -1.01)	1.61 (1.03, 2.20)			
	TS^{WUR}	2.61	-12.00 (-17.33, -6.67)	17.22 (11.89, 22.55)			
	CPM$^{PPT\text{-}mean}$	-19.26	-192.28 (-255.39, -129.16)	153.77 (90.65, 216.88)			
	CPM^{Unc}	-16.89	-271.24 (-364.02, -178.46)	237.46 (144.68, 330.24)			

Note: LoA= limit of agreement, CI= confidence interval, PPT= Pressure Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, TS^{WUR}= Temporal Summation calculated as a ratio, $CPM^{PPT\text{-}mean}$= Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus, CPM^{Unc}= Conditioned Pain Modulation where a unique PPT measurement was used as an unconditioned stimulus.

Supplementary Table 3.2: Correlations between QST modalities in different participant groups

	Healthy^{leg}: *Rater 1*		***Healthy^{leg}:*** *Rater 2*	***RA^{leg}:*** *Rater 1*		***Healthy^{forearm}:*** *Rater 3*		***Healthy^{forearm}:*** *Rater 2*	***LBP^{forearm}:*** *Rater 3*	
	Baseline	*Follow-up*	*Baseline*	*Baseline*	*Follow-up*	*Baseline*	*Follow-up*	*Baseline*	*Baseline*	*Follow-up*
PPT^{mean} – TS^{WUD}	-.38 (.10)	-.29 (.22)	-.33 (.16)	-.59 (.01) [a]	-.42 (.09)	-.05 (0.79)	-.31 (.14)	-.04 (.86)	-.42 (.03) [a]	-0.53 (.01) [a]
PPT^{mean} – TS^{WUR}	-.36 (.12)	.03 (.91)	-.29 (.22)	-.26 (.30)	.16 (.54)	-.14 (0.49)	-.18 (.38)	.08 (.75)	.11 (.62)	-.03 (.90)
TS^{WUD} – $CPM^{PPT\text{-}mean}$	-.28 (.24)	.02 (.94)	.03 (.91)	-.22 (.39)	.01 (.98)	-.05 (.79)	-.03 (.90)	-.28 (.18)	-.03 (.90)	-.19 (.38)
PPT^{mean} – $CPM^{PPT\text{-}mean}$	-.10 (.68)	-.13 (.60)	.01 (.97)	-.14 (.59)	.25 (.31)	-.002 (.99)	.19 (.36)	.39 (.047) [a]	.26 (.21)	.26 (.22)
PPT^{mean} – CPM^{Unc}	.03 (.91)	-.13 (.60)	.45 (.045) [a]	-.19 (.45)	.42 (.08)	.85 (<.01) [a]	.91 (<.01) [a]	.73 (<.01) [a]	.85 (<.01) [a]	.93 (<.01) [a]

Note: data is presented as r (p-value), Pearsons (normally data) and Spearmans correlation coefficient (non-normal data) were conducted. [a] = p<.05, showing statistically significant correlation between the two variables.

PPT^{mean} = Mean Pressure-Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, TS^{WUR}= Temporal Summation calculated as a ratio, $CPM^{PPT\text{-}mean}$= Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus, CPM^{Unc}= Conditioned Pain Modulation where a unique PPT measurement was used as an unconditioned stimulus.

Supplementary Table 3.3: Correlations between QST modalities with age in pooled participant groups

	Healthy^{leg} and RA^{leg} participants			***Healthy^{forearm} and LBP^{forearm} participants***		
	Rater 1 (n=38)		*Rater 2 (n=20)*	*Rater 3 (n=50)*		*Rater 2 (n=25)*
	Baseline	*Follow-up*	*Baseline*	*Baseline*	*Follow-up*	*Baseline*
PPT	-.18 (.28)	-.18 (.28)	-.24 (.14)	.01 (.97)	-.02 (.89)	-.10 (.65)
TS^{WUD}	.24 (.14)	.40 [a] (.01)	.39 (.09)	.08 (.60)	.10 (.51)	-.18 (.38)
CPM$^{PPT\text{-}mean}$	-.12 (.49)	-.12 (.49)	-.13 (.43)	-.20 (.17)	.02 (.88)	-.10 (.63)

Note: data is presented as r (p-value), Pearsons (conducted for normal data) and Spearmans correlation coefficient (for non-normally distributed data). [a] = p<.05, showing statistically significant correlation between the two variables.

PPT = Mean Pressure-Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, CPM$^{PPT\text{-}mean}$ = Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus

Supplementary Table 3.4: Sex differences between QST modalities in pooled participant groups

	Healthy^{leg} and RA^{leg} participants						***Healthy^{forearm} and LBP^{forearm} participants***					
	Rater 1 (n=38)				*Rater 2 (n=20)*		*Rater 3 (n=50)*				*Rater 2 (n=25)*	
	Baseline		*Follow-up*		*Baseline*		*Baseline*		*Follow-up*		*Baseline*	
	Male	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
PPT	605.87 (204.93) [a]	328.79 (200.92) [a]	655.99 (244.32) [a]	337.24 (212.41) [a]	558.38 (216.42)	362.14 (225.44)	339.85 (159.85) [a]	217.37 (84.76) [a]	320.81 (209.20)	219.80 (84.90)	373.92 (292.56)	191.10 (84.95)
TS^{WUD}	0.95 (0.63)	1.81 (1.57)	1.30 (0.76)	2.27 (1.58)	1.13 (0.56)	1.24 (1.12)	1.73 (1.42)	1.44 (1.34)	1.61 (1.44)	1.38 (1.44)	2.00 (1.47)	1.50 (1.68)
CPM$^{PPT\text{-}mean}$	82.47 (127.13)	95.20 (96.45)	137.30 (115.53)	123.02 (107.81)	149.71 (94.44)	62.00 (102.02)	63.90 (116.36)	76.94 (45.26)	92.10 (80.70)	83.99 (98.46)	64.46 (126.22)	131.94 (108.37)

Note: data are presented as means (standard deviations). [a] = p<.05, showing statistically significant sex difference in measurements calculated using independent samples t-tests (normal data) and Mann-Whitney U tests (non-normal data).

PPT = Mean Pressure-Pain Threshold, TS^{WUD}= Temporal Summation calculated as a difference, CPM$^{PPT\text{-}mean}$ = Conditioned Pain Modulation where the mean of the three PPT measurements was used as an unconditioned stimulus.

CHAPTER 4: COMPARABILITY OF TRIAXIAL AND UNIAXIAL CUT-POINTS AND WRIST AND HIP PLACEMENT OF THE ACTIGRAPH GT9X FOR ASSESSING FREE-LIVING PHYSICAL ACTIVITY AND SEDENTARY TIME

Chapter 4 Prelude

One of the main aims of this thesis was to recruit further participants and obtain data to add to an already existing database of accelerometer-assessed physical activity (PA) in people with Rheumatoid Arthritis (RA) (***Chapter 5***). This was to answer thesis aims 4 and 5, to investigate associations between PA and sedentary behaviour (SB) with core RA-related patient- and clinician-important health outcomes in a large-scale observational study (O'Brien et al., 2018b). In the initial study as part of the existing database, data was captured using the validated and reliable ActiGraph GT3X accelerometer (O'Brien et al., 2020). However, since this was conducted, a newer ActiGraph model with improved features has been released, the GT9X. Less is known about the validity and reliability of this model.

In previous studies of healthy adults, the GT9X has demonstrated high comparability and agreement in measurements of raw acceleration, vector magnitude (VM) and different PA intensities compared to the GT3X, when worn on the hip (Clevenger et al., 2020a; Montoye et al., 2018b). Part of this thesis involved assessing inter-monitor reliability between the GT3X and GT9X, in order to ensure data collected using the new GT9X model could be combined with the older GT3X data. This was done by employing an almost identical measurement protocol, data processing and reduction methods to those used by O'Brien et al. (2018b). Minor differences in data processing were that, in the current study, we employed triaxial cut-points (i.e., cut-points calculated using acceleration data from 3 axes) to accelerometer data, whilst, O'Brien et al. (2018b) employed the uniaxial cut-points (i.e., cut-points calculated using acceleration data from 1 axis only) developed by Troiano et al. (2008) to GT3X data. We chose to use triaxial cut-points as recent literature has suggested that triaxial cut-points should be applied to triaxial accelerometers for accurate assessment of free-living PA and sedentary time (ST) (Kozey-Keadle et al., 2014; Leeger-Aschmann et al., 2019). In addition, recent studies have demonstrated that Troiano cut-points are not accurate at measuring free-living ST in people with RA (O'Brien et al., 2020), perhaps due to their uniaxial nature. Therefore, triaxial

cut-points were chosen for use in this study, as they may provide more accurate measurement of PA and ST when applied to individuals with RA in a study as part of ***Chapter 5***.

Growing evidence has shown that triaxial and uniaxial cut-points significantly differ in their estimates of PA, and therefore, estimates of PA are highly dependent on data processing methods (Sagelv et al., 2019). Although data processing methods in this study were slightly different to those used by O'Brien et al. (2018b), Aim 1 of ***Chapter 4*** will investigate if triaxial and uniaxial cut-points, when applied to the same accelerometer data, give comparable estimates of different intensities of PA and ST. This section, as a prelude to ***Chapter 4***, reports the inter-monitor reliability of the GT9X compared to the GT3X for measuring different intensities of free-living PA and ST, when worn on the hip, in a group of healthy adults.

Methods

Participants, inclusion criteria, protocol and methods were identical to those employed in the main part of ***Chapter 4***. Inclusion criteria were being able to speak English, aged 18-65 years and able to replicate normal activities of daily living during the study week. In total, 37 healthy adults visited the School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham to participate.

Participants were asked to wear both ActiGraph GT3X and GT9X accelerometers on their right hip for 7 days whilst continuing normal daily activities. They were instructed to remove both accelerometers only for water-based activities and for sleep. The devices were attached directly next to one another, with the GT9X in-front of the GT3X (i.e., GT9X towards the centre of the participant's body), and worn on an elasticated belt on the participants right hip (Montoye et al., 2020; Rhudy et al., 2020). To track monitor removal and replacement, participants completed a log book. After 7 days of wear, participants returned accelerometers to the School of Sport, Exercise and Rehabilitation Sciences.

ActiGraph GT3X: The ActiGraph GT3X is a small, lightweight (19g; 4.6 x 3.3 x 1.5 cm) triaxial accelerometer which records acceleration on the vertical (Y), antero-posterior (Z), and medio-lateral

(X) axes. *ActiGraph GT9X:* A newer ActiGraph triaxial accelerometer model (14g; 4.18 x 3.98 x 1.13 cm) records acceleration identically to the GT3X. The GT3X and GT9X accelerometers were initialised at 60Hz sampling rate and sampled movement in 1 second epochs. Data processing and reduction were identical to those described in the main part of ***Chapter 4***.

ActiLife software (ActiGraph, LLC., Pensacola, Florida, USA – version 6) was used to analyse the data. This software uses raw accelerometer data from the 3 axes to calculate VM (VM= $\sqrt{(axisY^2 + axisZ^2 + axisX^2)}$), and data are then compressed and converted into activity counts. Activity counts are then used to identify periods of non-wear, and to quantify time spent in different intensities of PA and ST. To determine non-wear, data was scanned using the Troiano et al. (2008) algorithm, which defines non-wear time as ≥60 consecutive minutes of zero VM counts, with a spike tolerance of 2 minutes of counts between 0-100 counts (Semanik et al., 2010). Furthermore, a day was deemed valid when a participant had at least 10 hours wear. Participants were retained for inclusion in further analysis when they had recorded valid data (i.e., ≥10 hours wear) on ≥4 days (Semanik et al., 2010; Troiano et al., 2008). Log books were used to ensure wear and non-wear time scanned using the Troiano algorithm was correct. Wear and non-wear time was manually amended on wear-time validation outputs if required. A time filter of between 10am – 8pm was applied to the data to ensure data extracted was identical between participants in regards to the periods of time being compared. For both devices, identical wear days and times of assessment were used in the data analyses to allow for comparison between devices.

In order to estimate different PA intensities, count-based "cut-points" were applied to the counts per minute (cpm) data, in order to categorise PA into ST, LPA and MVPA. In this study, triaxial cut-points created by Sasaki et al. (2011) were employed, defined as: ST= ≤150 cpm, LPA= 151–2690 cpm, moderate intensity PA (MPA)= 2691–6166 cpm, vigorous intensity PA (VPA)= 6167–9642 cpm, and very vigorous intensity PA (VVPA)= >9642 cpm. As the existing dataset used MVPA as the outcome of

interest, we combined MPA, VPA and VVPA to create an MVPA outcome, to enable comparison across different studies, monitors and cut-points.

Data analysis methods and statistical tests were identical to those described in the main part of ***Chapter 4***. Statistical tests were: paired samples t-tests to assess inter-monitor differences, Pearsons correlation coefficient and intraclass correlation coefficient (ICC) (using two way mixed effects analysis of variance (ANOVA), testing for absolute agreement, with single measures) to examine agreement between the GT3X and GT9X for ST, LPA, and MVPA. Bland Altman plots with 95% limits of agreement (LoA) were constructed to give a graphical representation of data and to illustrate any bias.

Results

Table 4.1: Means (SD) for Sedentary Time, Light intensity Physical Activity and Moderate to Vigorous Physical Activity for the hip-worn ActiGraph GT3X and GT9X

	HipGT3X	***HipGT9X***
ST (min/day)	466.71 (25.05)	464.38 (25.96) [a]
LPA (min/day)	65.17 (15.06)	64.99 (14.57)
MVPA (min/day)	68.12 (15.64)	70.63 (16.15) [a]

Note: n=30. [a] = paired-samples t-test indicates significant difference between hip-worn GT3X and GT9X, p<.05.
HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, LPA= Light intensity Physical Activity, MVPA= Moderate to Vigorous Physical Activity, min/day= minutes per day.

Valid data were available for n=30 healthy adult participants (mean age= 30.0 ± 11.2 years, 63% female). **Table 4.1** displays the means and standard deviations (SD) for all outcome variables. Paired samples t-tests revealed significant differences between the 2 monitors in computed levels of ST and MVPA, but with not for LPA. As reported in **Table 4.1**, ST estimated by GT3X was higher and MVPA estimated by GT3X was lower compared to GT9X. Although ST and MVPA estimates were significantly different between accelerometers, 95% confidence intervals (CI) for the true mean difference were -3.12 to -1.53 min/day for ST and 1.28 to 3.73 min/day for MVPA, indicating that the difference between measurements from each accelerometer was relatively small.

Pearsons correlation coefficient demonstrated positive correlations between ST, LPA and MVPA between accelerometers, with r-values of >.99, .98 and .98, respectively (data not shown). The results of ICC are displayed in **Table 4.2**. Very high agreement was demonstrated for ST, LPA, and MVPA between the hip-worn GT3X and GT9X accelerometers. Finally, Bland Altman analysis illustrated small between-monitor mean differences and narrow 95% LoA between GT3X and GT9X estimates of ST, LPA, and MVPA (**Figure 4.1a, 4.1b and 4.1c**).

Table 4.2: Intraclass correlation coefficient results for Sedentary Time, Light intensity Physical Activity and Moderate to Vigorous Physical Activity between hip-worn ActiGraph GT3X and GT9X

HipGT3X and HipGT9X

	ICC	*95% CI*	
		Lower limit	*Upper limit*
ST	.99	.91	<1.00
LPA	.98	.96	.99
MVPA	.97	.86	.99

Note: n=30. ICC= Intra-class correlation coefficient, CI= Confidence Interval, HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, LPA= Light intensity Physical Activity, MVPA= Moderate to Vigorous Physical Activity, min/day= minutes per day.

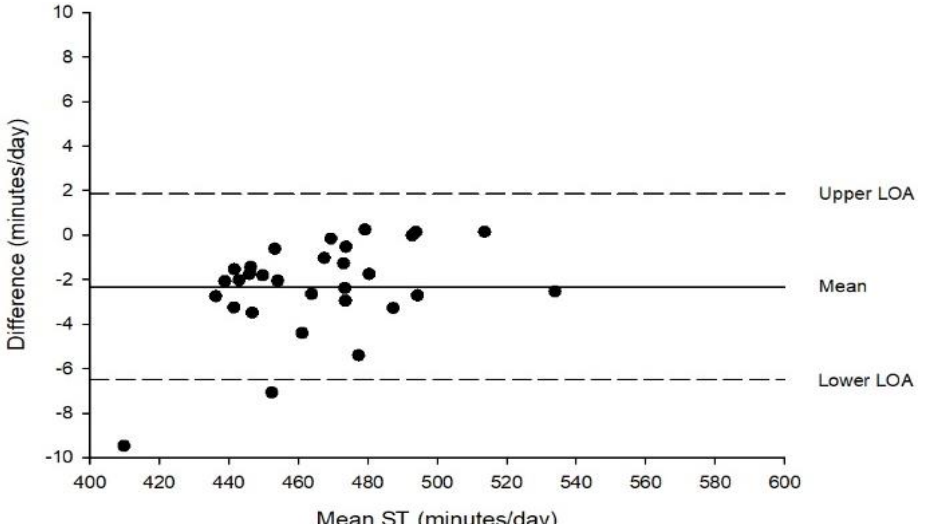
Conclusion

The results of this analysis show that the ActiGraph GT3X and GT9X have very high inter-monitor reliability for assessment of ST, LPA and MVPA when both worn on the right hip, with the same triaxial cut-points applied. Although t-tests did display some differences between ST and MVPA measurements, exploration of t-test 95% CIs and Bland Altman analysis showed that there were consistent differences (mean bias) between monitors within-participants. The GT9X consistently gave a higher measure of MVPA and lower ST than the GT3X. However these differences were too small to be meaningful (e.g., 3.12 to -1.53 min/day CIs for ST and 1.28 to 3.73 min/day CIs for MVPA), and ICCs still demonstrated very high inter-monitor agreement. Therefore, this suggests that these devices can be worn interchangeably, giving highly comparable estimations of different PA intensities, when the same cut-points are applied. As the GT3X accelerometer was used previously by

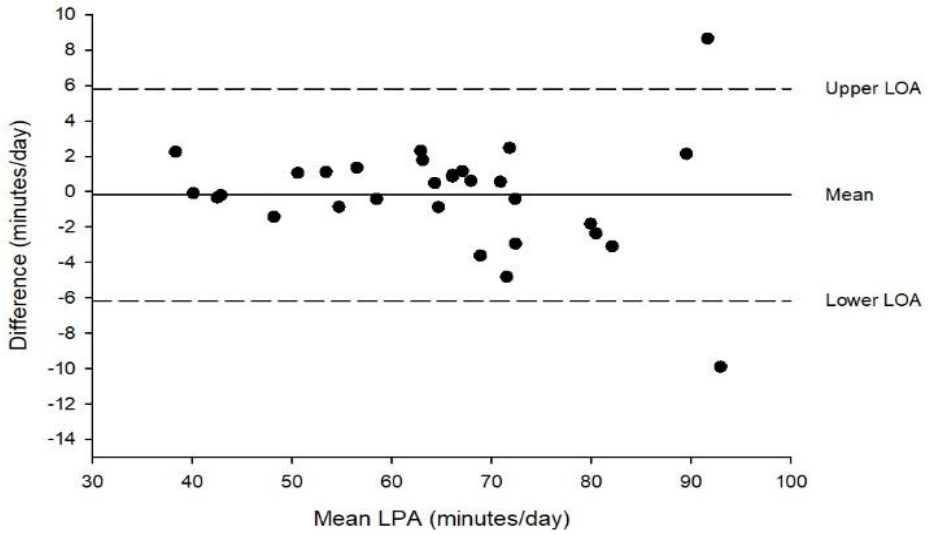
my research team in an existing dataset, this analysis has proven that the hip-worn GT9X, with identical cut-points applied, is a reliable and comparable replacement for the GT3X.

Relevance to Chapter 4

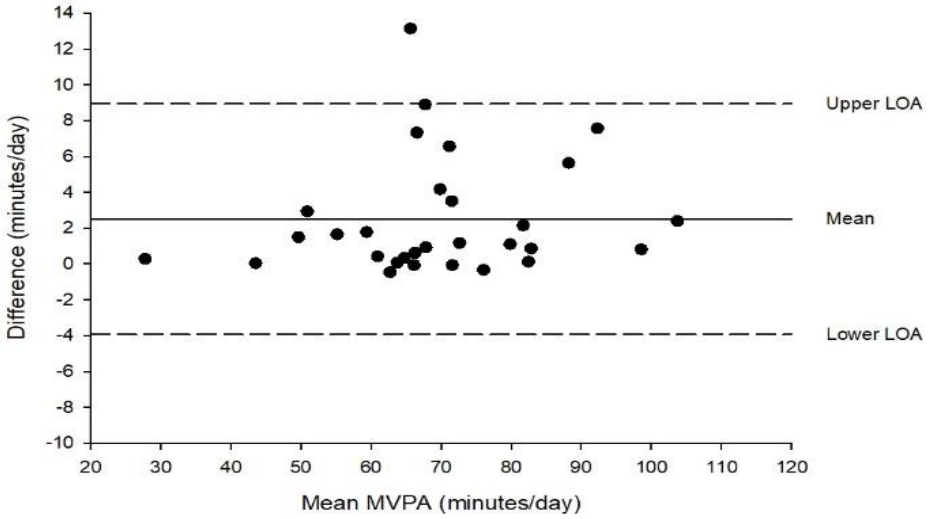
These data demonstrate there is little difference between the GT3X and GT9X in regards to data capture, where the same device placement and analytical processes are applied. Consequently, the decision was made to utilise the GT9X for the studies focussed on RA in this thesis, prior to COVID-19 impacting this work. The remainder of this chapter (***Chapter 4***), therefore, employs the ActiGraph GT9X as the model of interest, when answering questions around how device placement (hip vs wrist) and analytical processing decisions around cut-points, influence accelerometer outcomes.



4.1a.) *ST agreement between HipGT3X and HipGT9X*



4.1b.) *LPA agreement between HipGT3X and HipGT9X*



4.1c.) *MVPA agreement between HipGT3X and HipGT9X*

Figure 4.1: Inter-monitor reliability Bland Altman plots for a.) Sedentary Time; b.) Light Physical Activity, and c.) Moderate to Vigorous Physical Activity

Note: HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, LPA= Light intensity Physical Activity, MVPA= Moderate to Vigorous Physical Activity, min/day= minutes per day, Solid line= mean difference/bias between the two accelerometers, dashed lines= 95% limits of agreement (LoA)

Abstract

In order to assess the associations between PA and SB with health, it is essential to ensure accurate assessment of these behaviours. Data capture (i.e., cut-points applied) and device placement (i.e., wrist vs hip placement) have previously shown to significantly impact estimations of free-living LPA and MVPA, and ST for the ActiGraph GT9X. Aims of this study were: 1.) to compare PA and ST determined by uniaxial hip (*Troiano*), triaxial hip (*Sasaki*), and triaxial wrist (*Montoye*) cut-points; 2.) to assess the validity of these cut-points for measurement of ST, compared to the activPAL™. Participants (n=37) wore a hip- and wrist-worn GT9X accelerometer and an activPAL posture sensor for 7 days. ActiGraph data processing involved application of uniaxial (*Troiano*) and triaxial (*Sasaki*) hip cut-points to hip-worn GT9X data, and triaxial wrist (*Montoye*) cut-points to wrist-worn GT9X data. Differences and agreement between cut-points and devices were assessed with paired samples t-tests, ICCs and Bland-Altman analysis. *Troiano*, *Sasaki* and *Montoye* cut-points all demonstrated significantly different estimations of free-living PA and ST, with no to moderate agreement between cut-points (ICC= .02 − .74). For assessment of ST, cut-points demonstrated significant differences and low agreement with the activPAL, with ICCs= .07 − .21. Estimates of PA and ST quantified by the ActiGraph GT9X are not comparable where different protocols for data capture (uniaxial vs triaxial) and device placement (hip vs wrist) are employed. These cut-points displayed poor validity at assessing free-living ST, compared to the activPAL.

Introduction

The World Health Organisation (WHO) guidelines state that adults should undertake at least 150 minutes of MPA, or 75 minutes VPA per week (Bull et al., 2020). PA has numerous health benefits, including improving cardiovascular risk factors and reducing mortality (Nocon et al., 2008; Stamatakis et al., 2007; Warburton et al., 2006). SB is defined as "any waking behaviour characterized by an energy expenditure ≤1.5 MET (metabolic equivalent of task) while in a sitting or reclining posture", whereby 1 MET equals the amount of oxygen consumed at rest (i.e., 3.5 $ml \cdot kg^{-1} \cdot min^{-1}$) (Ainsworth et al., 2011; Sedentary Behaviour Research Network, 2012). Evidence conducted in healthy young and older adults has demonstrated that SB is an independent risk factor for health, including poor cardiometabolic and cardiovascular health, cancer and metabolic syndrome (Biswas et al., 2015; Chastin et al., 2015; de Rezende et al., 2014; Green et al., 2014).

In order to examine the role of PA and SB for health, it is important to be able to accurately quantify these behaviours (Chastin et al., 2018; Edwardson et al., 2017; Healy et al., 2011). Accurate assessment is also important to evaluate the efficacy of interventions targeting changes in PA and ST. Research suggests self-report methods for PA and ST measurement are subject to self-report bias (Healy et al., 2011), and self-report questionnaires have previously demonstrated poor reliability and inconsistent results when compared against the doubly labelled water method criterion for assessment of PA (Sylvia et al., 2014; Westerterp, 2009). Alternatively, device-based assessment tools can provide a more objective assessment of PA and ST (Healy et al., 2011; Matthews et al., 2008), relative to self-report, and are increasingly used in PA and SB research.

Device-based measures include accelerometers, which capture movement via compression sensed by a piezoelectric instrument, which records accelerations of the area of the body they are attached to. Uniaxial accelerometers can record body movements from 1 axis (vertical axis (Y) only). However, newer accelerometer models can record accelerations from 3 axes (vertical (Y), antero-posterior (Z), and medio-lateral (X)). These are termed triaxial accelerometers (Vanhelst et al., 2012). As triaxial

accelerometers are able to record accelerations from 3 axes, therefore, they are able to sense and record a greater reflection of body movements and activities as people move in 3 dimensions (Vanhelst et al., 2012). Such studies have indicated that triaxial cut-points minimise measurement error, and non-wear periods can be more easily distinguished due to the additional data available (Choi et al., 2012; Evenson et al., 2015).

ActiGraph accelerometers are widely used in PA and ST research. They have been validated against indirect calorimetry (Kelly et al., 2013), and display high reliability for measurement of PA and ST (Calabro et al., 2014). These devices are easy to administer to participants (Vanhelst et al., 2012), and offer a user-friendly platform for analysis – ActiLife (ActiGraph, LLC., Pensacola, Florida, USA). ActiLife software takes raw accelerometer data collected via the X, Y, and Z axis, and compresses it into a unit called "activity counts" for use in analysis. "Cut-points" or "thresholds" can then be applied to these activity counts, in order to classify activities as ST, LPA, MPA or VPA.

The most commonly used ActiGraph cut-points were developed for use with a previous uniaxial ActiGraph accelerometer version – the 7164 – by Troiano et al. (2008). These are used to calculate time spent in ST, MPA and VPA using acceleration data from the vertical axis only (uniaxial) (Matthews et al., 2008; Troiano et al., 2008). The uniaxial Troiano cut-points were developed by calculating a weighted average from results of multiple calibration studies assessing walking and running activities only (Brage et al., 2003; Freedson et al., 1998; Leenders et al., 2001; Yngve et al., 2003), and subsequently used in a large-scale epidemiological study in healthy adults (Matthews et al., 2008; Troiano et al., 2008). However, they have displayed poor validity at quantifying PA intensities when compared against indirect calorimetry, and poor validity at measuring ST when compared against the activPAL (Crouter et al., 2013; Koster et al., 2016; Watson et al., 2014). Despite this, Troiano cut-points remain widely employed in research. More recently, cut-points have been developed for triaxial accelerometer models, such as the ActiGraph GT9X, which utilise data measured from all 3 axes (Sasaki et al., 2011). A growing number of studies have suggested triaxial

cut-points provide more comprehensive measurement of different types of PA relative to uniaxial cut-points (Evenson et al., 2015; Kozey-Keadle et al., 2014; Leeger-Aschmann et al., 2019; O'Brien et al., 2020). Further research has indicated that levels of PA and ST quantified by triaxial vs uniaxial cut-points differ substantially (Kozey-Keadle et al., 2014; Luzak et al., 2017; Sagelv et al., 2019). However, uniaxial cut-points are still frequently used in research (Duncan et al., 2020; Fenton et al., 2017; Fenton et al., 2018b; Hibbing et al., 2020). Consequently, researchers using triaxial accelerometers and corresponding cut-points continue to draw inappropriate comparisons between their data and that quantified using uniaxial data (Rhudy et al., 2020).

More recently, wrist-worn device-based measures are becoming a popular option for measuring ST and PA among the general population (Montoye et al., 2020). They are increasingly used in large-scale epidemiological research. They demonstrate a higher wear compliance than hip-worn devices (Rhudy et al., 2020; Troiano et al., 2014; Van Hees et al., 2011), and a superior ability to measure PA and ST in people with atypical gait and to capture non-ambulatory activities (i.e., ironing, sweeping) (Diaz et al., 2018; Troiano et al., 2008). However, similar to the triaxial vs uniaxial debate, growing research suggests that wrist-worn accelerometers produce significantly different estimates of raw acceleration, accelerometer counts, MVPA, and ST when compared to hip-worn accelerometers (Ellis et al., 2016; Hildebrand et al., 2014; Loprinzi & Smith, 2017). This is especially true where placement specific (i.e., hip vs wrist) cut-points are not employed (Rhudy et al., 2020). This is the case even where the more comprehensive triaxial cut-points are employed to wrist-worn device data (Clevenger et al., 2020a; Clevenger et al., 2020b; Montoye et al., 2020).

Together, research points to the importance of utilising triaxial, placement-specific (hip vs wrist) cut-points to accurately quantify PA and ST, as both of these factors impact data capture, and therefore the outcomes reported. Recently, new triaxial cut-points have been developed specifically for the ActiGraph models, when worn at the hip (Sasaki et al., 2011), and the wrist (Montoye et al., 2020). Sasaki et al. (2011) triaxial cut-points were developed for the hip-worn ActiGraph GT3X, which has

demonstrated good agreement with the GT9X model for estimating free-living PA and ST (Montoye et al., 2018b). Subsequently, Sasaki cut-points have been used as a criterion in development of newer triaxial cut-points (Clevenger et al., 2020a; Montoye et al., 2020), and have shown good accuracy and precision compared to other cut-points for measurement of energy expenditure in free-living older adults (Aguilar-Farias et al., 2019). Triaxial wrist cut-points have also been developed, by Montoye et al. (2020), in a controlled laboratory and using 3-8 hours of directly observed free-living data. Receiver operating characteristic analysis was performed to determine optimal triaxial count-based cut-points for ST, LPA and MVPA for the wrist-worn GT9X. These cut-points subsequently demonstrated high accuracy for assessing free-living PA and good wear compliance (Montoye et al., 2020). The Sasaki and Montoye triaxial cut-points have not been compared against each other and against uniaxial Troiano cut-points in free-living healthy adults.

The primary aim (Aim 1) of this study was to employ these 3 sets of cut-points, to determine the extent to which PA and ST outcomes differ according to a) data captured (uniaxial vs triaxial), and b) device placement (hip vs wrist). Specifically, Aim 1 of this study compared PA and ST determined by the extensively employed Troiano uniaxial hip cut-points (*HipGT9X-Troiano*), the Sasaki triaxial hip cut-points (*HipGT9X-Sasaki*), and the Montoye triaxial wrist cut-points (*WrGT9X-Montoye*). Results illustrated the extent to which between-study comparisons of PA and ST measured by the GT9X are appropriate, given variability in analytical methods and device placement sites. Based on the current literature, I hypothesise that triaxial cut-points will estimate less ST and more time spent in LPA and MVPA than uniaxial cut-points, due to the 3 axes used to record body accelerations (Kozey-Keadle et al., 2014; Luzak et al., 2017; Sagelv et al., 2019). When comparing hip vs wrist specific cut-points, previous findings suggest that there may be greater agreement with other triaxial hip cut-points compared to uniaxial hip cut-points (Rhudy et al., 2020). However, as Montoye wrist cut-points have not been extensively used or validated (Clevenger et al., 2020a; Clevenger et al., 2020b; Montoye et al., 2020), I hypothesise that there may be poor agreement and significant differences with both

Troiano uniaxial and Sasaki triaxial hip cut-points, although, for this reason, I cannot predict the direction of these differences.

Accelerometers, overall, are limited in their ability to accurately measure ST (Koster et al., 2016; Kuster et al., 2021). Little is known about the relative accuracy of accelerometer data processing for estimation of ST. Indeed, accelerometers classify ST based on lack of movement (i.e., low accelerations or "activity counts"), rather than posture (Chan et al., 2017). For assessment of free-living ST, the activPAL is considered to be the "gold standard" device-based measure and has demonstrated high validity in different populations when compared to direct observation (Edwardson et al., 2017; Kozey-Keadle et al., 2011; Sellers et al., 2016). Currently, the validity of the GT9X, with different cut-points applied, and worn at different placement sites, for quantification of ST, has not yet been established. A secondary aim of this study, Aim 2, was to assess the criterion validity of these uniaxial and triaxial hip and wrist GT9X cut-points for measurement of ST, when compared to the activPAL. For Aim 2, I hypothesise that the GT9X cut-points and placement sites will estimate significantly more ST than the activPAL, due to previous studies indicating the poor ability of accelerometers and their cut-points for ST assessment (Koster et al., 2016; Kuster et al., 2021).

Methods

Participants and Recruitment

Healthy adults were recruited from the School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham, using invitational emails and through word of mouth. To be eligible to participate, individuals were required to be aged between 18-65 years, speak English, and be able to replicate normal activities of daily living during the study week. Written informed consent was taken, and this study was approved by the University of Birmingham Research Governance (ERN_18-1811).

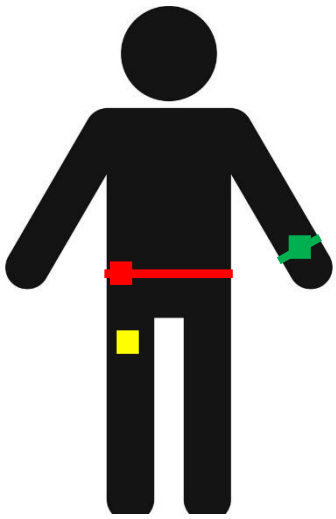


Figure 4.2: Wear sites for the 3 accelerometers.
Note: Red: ActiGraph GT9X on right hip, Yellow: activPAL™ on right thigh, Green: ActiGraph GT9X on non-dominant wrist

Protocol

A total of n=37 participants (100% white ethnicity) visited the School of Sport, Exercise and Rehabilitation Sciences at the University of Birmingham to take part. Upon arrival, participants were fitted with 3 devices, as illustrated in **Figure 4.2**. Specifically, an ActiGraph GT9X device was attached to an elasticated belt on the right hip (HipGT9X), and another on the non-dominant wrist (WrGT9X) using a purposely designed watch strap. The activPAL was also attached to mid-anterior right thigh (activPAL™), using a waterproof adhesive waterproof dressing (Tegaderm, Farmaline) (Edwardson et al., 2017).

Participants were asked to wear the devices for 7 days, whilst continuing their normal daily activities. They were instructed to remove both GT9X monitors for water-based activities (due to not being waterproof), and HipGT9X was also removed for sleep. To keep track of monitor removal and replacement, participants were asked to complete a log book and activity diary. After 7 days of accelerometer wear, participants returned to the School of Sport, Exercise and Rehabilitation Sciences to remove and return accelerometers.

Measures

ActiGraph GT9X: The GT9X is a new ActiGraph triaxial accelerometer model, and is small and lightweight (14g; 4.18 x 3.98 x 1.13 cm). The device records accelerations on the X, Y and Z axes and these raw acceleration data are used to determine VM (VM= $\sqrt{(axisY^2 + axisZ^2 + axisX^2)}$) of these accelerations, which is subsequently compressed into activity counts using the ActiLife software (ActiGraph, LLC., Pensacola, Florida, USA). The GT9X accelerometers were initialised at 60Hz sampling rate and sampled movement in 1 second epochs.

activPAL3^{TM}: The activPAL is a postural classification device (9g; 2.35 x 4.3 x 0.5 cm), worn on the thigh in the mid-anterior position (Edwardson et al., 2017). By incorporating gravity into its accelerometer readings, this allows for postural allocation to determine free-living sedentary (sitting/lying), upright (standing) and ambulatory (walking) activities (Chan et al., 2017). The activPAL is considered the criterion "gold standard" measure for the assessment of free-living ST (Edwardson et al., 2017; Kozey-Keadle et al., 2011; Sellers et al., 2016).

Data Processing

ActiGraph GT9X: Data from ActiGraph devices was downloaded in 1 second epochs and raw data was converted into triaxial (VM) activity counts using ActiLife version 6 software (ActiGraph, LLC., Pensacola, Florida, USA). Data were scanned for wear-time using the Troiano 2007 algorithm, previously used in epidemiological studies on healthy adults (including the National Health and Nutrition Examination Survey, NHANES) (Semanik et al., 2010; Troiano et al., 2008). Using this algorithm, non-wear time was defined as ≥60 consecutive minutes of zero VM counts, with a spike tolerance of 2 minutes of counts between 0-100 counts. A valid day was defined as at least 10 hours (600 minutes) wear during waking hours. Where accelerometers defined a period as non-wear, log books were consulted to confirm this, and if wear/non-wear periods were incorrectly defined by accelerometers, researchers manually edited accelerometer data to match the log books.

Data Reduction

For inclusion in further analysis, all participants were required to have 10 hours of complete data recorded between the hours of 10am − 8pm, on each of the ≥4 valid days. These hours were chosen from manual exploration of participant data, which demonstrated that these hours were the most common "wear hours" within and between participants. This ensured that estimates of PA and ST would not be impacted by within-participant variability in wear-time (i.e., hour to hour). Where the complete 10 hours of data were recorded, these days were considered "valid wear days" (Semanik et al., 2010; Troiano et al., 2008). Once the number of "valid wear days" was determined, data from

each participant was checked to ensure that the "valid wear days" were the same across all 3 devices (HipGT9X, WrGT9X, and activPAL). This ensured that estimates of PA and ST would not be impacted by within-participant variability in daily PA and ST (i.e., day to day). See **Figure 4.3** for an illustrative example.

Cut-points applied to the data, to quantify time spent in PA and ST, were as follows; *HipGT9X-Troiano:* Troiano Adult 2008 uniaxial cut-points defined as: ST = <100 cpm, LPA = 100−2019cpm and MVPA = ≥2020cpm (Matthews et al., 2008; Troiano et al., 2008); *HipGT9X-Sasaki:* ST= ≤150cpm, LPA= 151–2690cpm, MPA= 2691–6166cpm, VPA= 6167–9642cpm, and VVPA= >9642cpm (Sasaki et al., 2011); and *WristGT9X-Montoye*: ST= <2860cpm, LPA= 2860−3940cpm, MVPA= ≥3941cpm (Montoye et al., 2020). For the purposes of this study, MPA, VPA and VVPA were combined to create an MVPA outcome, to enable comparison across different monitors and cut-points.

activPAL: PAL Connect software (PAL Technologies Ltd., Glasgow, UK) was used to download and export activPAL data (in 15 second epochs) to Microsoft Excel. Data was manually checked and cleaned for wear and non-wear time, and filtered in order to ensure data recorded between specific hours (10am – 8pm), and on "valid wear days" matched those of the filtered ActiGraph data. PAL Connect CREA enhanced analysis algorithm was used to calculate ST. This algorithm involves using raw acceleration data to calculate sitting, standing and walking events. Lying time is also classified as primary (in bed and non-wear) and secondary lying time (daytime naps). Together, secondary lying time is grouped with sitting to calculate daily waking ST (Carlson et al., 2021). For this study, ST was calculated by adding data for sitting time and lying time. Both primary and secondary lying time were used, as my time filter ensured any primary lying time was captured during waking hours (10am – 8pm), and would constitute ST during the day, rather than night-time sleep. This was confirmed by consultation of participants' activity diaries. Specifically, time spent in primary or secondary lying time were examined to ensure study estimates of ST represented only waking ST, and not ST in which participants were asleep during waking hours. Where diaries showed primary lying time did capture

sleep during waking hours, the specific primary lying time period was excluded from analysis for that participant.

Data Analysis

Data analysis methods were identical for Aims 1 and 2. Of the n=37 participants recruited to the study, valid data comparing device placement and corresponding cut-points (Aim 1), were available for analysis from n=30 (81.08%) participants. For comparisons between the GT9X uniaxial (Troiano) and triaxial cut-points (Sasaski and Montoye), and hip (Troiano and Sasaki) and wrist (Montoye) cut-points, with the activPAL for measurement of ST (Aim 2), data were available from n=28 (75.68%) participants. Data was excluded and deemed non-valid (from Aim 1 and 2) due to monitor malfunction (n=6), invalid ActiGraph data (≤4 days with 10 hours/day wear time) (n=1). For Aim 2 analysis only, further exclusions were due to activPAL malfunction (n=1), and reaction to adhesive dressing (n=1). No differences were observed in participant characteristics between valid and non-valid participants.

Descriptive statistics involved computing means and standard deviations (SD) for all continuous outcome variables for each of the device placements (HipGT9X, WrGT9X, activPAL) and corresponding cut-points (Sasaki, Troiano, Montoye). Shapiro-Wilk normality test results showed all variables except MVPA estimated from triaxial Montoye wrist cut-points were normally distributed. Paired samples t-tests (normally distributed data) or Wilcoxon signed rank tests (non-normally distributed data) were carried out to determine any significant differences in ST, LPA and MVPA quantified by each cut-point. Pearsons (normally distributed data) and Spearmans (non-normally distributed data) correlation coefficient tests were conducted to examine associations, and ICC tests were employed to examine agreement, between estimates of ST, LPA, and MVPA captured by each of the device placements and corresponding cut-points. ICCs employed a two way mixed effects ANOVA, testing for absolute agreement, with single measures. An ICC <.50 was considered low

agreement, .50−.74 moderate, .75−.90 high, and >.90 very high agreement (Portney & Watkins, 2009).

Bland Altman plots were produced to give a graphical representation of data, and to illustrate any bias in estimates of ST and PA using different device placements and cut-points. The plots display the difference between the cut-points/accelerometers (graph y axis) and mean of the 2 cut-points/accelerometers (x axis). A mean difference (mean bias) between cut-points/accelerometers was calculated, as well as 95% limits of agreement (95% LoA), using the formula:
mean difference ± 1.96 × SD of difference between 2 scores. If the mean difference was close to 0, with all data points within narrow 95% LoA, the results showed no significant difference in levels of ST and PA according to device (placement, cut-point, or accelerometer vs activPAL).

Count based data were all exported to Microsoft Excel and statistical analysis was conducted using Statistical Package for the Social Sciences (SPSS) Version 24, with Bland Altman analysis conducted using SigmaPlot Version 14.5. Statistical test were carried out using the $p<.05$ significance level.

Hour	*Participant*			
	1	*2*	*3*	*4*
00:00 – 00:59				
01:00 – 01:59				
02:00 – 02:59				
03:00 – 03:59				
04:00 – 04:59				
05:00 – 05:59				
06:00 – 06:59				
07:00 – 07:59				
08:00 – 08:59				
09:00 – 09:59				
10:00 – 10:59				
11:00 – 11:59				
12:00 – 12:59				
13:00 – 13:59				
14:00 – 14:59				
15:00 – 15:59				
16:00 – 16:59				
17:00 – 17:59				
18:00 – 18:59				
19:00 – 19:59				
20:00 – 20:59				
21:00 – 21:59				
22:00 – 22:59				
23:00 – 23:59				
Hours wear	10	9	15	17
Valid wear day (included vs excluded)	**Included**	**Excluded (<10 hours)**	**Excluded (missing data between 10am-8pm)**	**Included**

Figure 4.3: Extraction of daily data- illustration of valid and non-valid daily wear criteria for 4 participants

Indicates valid hour (i.e., 60 minutes of movement data recorded).

Indicates invalid hour (i.e., <60 minutes of movement data recorded).

Indicates hour to be filtered out to ensure valid wear day (i.e., before 10am or after 8pm).

Results

Participant characteristics are reported in **Table 4.3**. Data from the n=30 and n=28 valid participants provided a total of 162 and 149 valid days of recording for Aims 1 and 2, respectively.

Table 4.3: Participant Characteristics

	Aim 1	***Aim 2***
N	30	28
Valid days per participant	5.4 ± 1.2	5.3 ± 1.2
Age	30.0 ± 11.2	30.1 ± 11.6
Gender (% Female)	63.3	60.7

Note: age and valid days are presented as mean ± standard deviation. Gender is presented as a percentage of the total valid participants.

Aim 1: ActiGraph comparability

Table 4.4 reports the means and SD for daily time spent in ST, LPA, and MVPA, as calculated by the uniaxial hip-worn (HipGT9X-Troiano), triaxial hip-worn (HipGT9X-Sasaki) and triaxial wrist-worn (WrGT9X-Montoye) GT9X and corresponding placement-specific cut-points. Results of paired samples t-tests (normal data) and Wilcoxon signed rank tests (non-normal data) demonstrated differences between triaxial HipGT9X-Sasaki hip and uniaxial HipGT9X-Troiano hip cut-points for measurements of ST, LPA, and MVPA. The HipGT9X-Troiano cut-points demonstrated significantly greater amount of ST and less LPA and MVPA than HipGT9X-Sasaki cut-points. The 95% CI for the true mean difference were -50.01 to -40.24 min/day, 17.46 to 25.25 min/day and 21.32 to 26.21 min/day for ST, LPA and MVPA respectively.

Between uniaxial HipGT9X-Troiano hip and triaxial WrGT9X-Montoye wrist cut-points, there were significant differences between data for ST, LPA and MVPA. HipGT9X-Troiano cut-points gave higher estimates of ST and lower estimates of LPA and MVPA than WrGT9X-Montoye cut-points (**Table 4.4**). In addition, the 95% CI for the true mean difference were -66.45 to -52.59 min/day and 36.71 to 52.55 min/day for ST and LPA, respectively (no 95% CI available for MVPA due to non-normal data). These differences between cut-points were greater than differences between the two sets of hip cut-points, perhaps due to differences in both data capture (i.e., triaxial vs uniaxial) and placement site.

Finally, differences were also displayed between triaxial WrGT9X-Montoye wrist and triaxial HipGT9X-Sasaki hip cut-points for estimates of ST, LPA and MVPA (**Table 4.4**). The 95% CI for the true mean difference were -20.08 to -8.70 min/day for ST and 16.09 to 30.45 min/day for LPA (no 95% CI available for MVPA due to non-normal data), suggesting differences between the two triaxial cut-points were relatively small, compared to differences between the other cut-points.

Table 4.4: Means and standard deviations for Sedentary Time, Light Physical Activity and Moderate to Vigorous Physical Activity for HipGT9X-Troiano, HipGT9X-Sasaki and WrGT9X-Montoye

	HipGT9X- Troiano	*HipGT9X-Sasaki*	*WrGT9X- Montoye*
ST (min/day)	509.51 (19.61) [ab]	464.38 (25.96) [ac]	449.99 (28.32) [bc]
LPA (min/day)	43.63 (10.86) [ab]	64.99 (14.57) [ac]	88.26 (19.61) [bc]
MVPA (min/day)	46.86 (14.16) [ab]	70.63 (16.15) [ac]	61.75 (22.17) [bc]

Note: [a] = paired-samples t-test results showing significant difference between uniaxial HipGT9X-Troiano hip with triaxial HipGT9X-Sasaki hip cut-points. [b] = paired-samples t-test/Wilcoxon signed rank test results showing significant difference between uniaxial HipGT9X-Troiano hip with triaxial WrGT9X-Montoye wrist cut-points. [c] = paired-samples t-test/Wilcoxon signed rank test results showing significant difference between triaxial HipGT9X-Sasaki hip with triaxial WrGT9X-Montoye wrist cut-points (p<.05).
ST=Sedentary Time, LPA=Light Physical Activity, MVPA=Moderate to Vigorous Physical Activity, min/day= minutes per day.

Correlation coefficient and ICC results are displayed in **Tables 4.5 and 4.6**, respectively. Positive correlations were displayed between the uniaxial HipGT9X-Troiano hip with the triaxial HipGT9X-Sasaki hip cut-points for ST, LPA, and MVPA (**Table 4.5**), but ICC analysis revealed low agreement between HipGT9X-Troiano and HipGT9X-Sasaki for ST and PA intensities (ICC= .28 to .41) (**Table 4.6**), with lower bounds for 95% CIs less than 0. This indicated little agreement between activity intensity estimates when the two cut-points were compared. When assessing the correlation between uniaxial HipGT9X-Troiano hip and triaxial WrGT9X-Montoye wrist cut-points and placement site, analysis revealed significant positive associations for ST and MVPA. No correlations were visualised between cut-points for LPA (**Table 4.5**). Furthermore, ICCs revealed no to low agreement between cut-points for ST, LPA and MVPA (ICC= .02 to .21), with lower limits for 95% CI less than 0 (**Table 4.6**). In addition, positive correlations were demonstrated between the triaxial HipGT9X-Sasaki hip with triaxial WrGT9X-Montoye wrist cut-points and placement sites for all activity intensities (**Table 4.5**). ICC analysis demonstrated low to moderate agreement (ICC= .20 to .74) between WrGT9X-Montoye and HipGT9X-Sasaki cut-points for all activity intensities (**Table 4.6**).

Table 4.5: Correlation Coefficient results between accelerometers and cut-points for Aims 1 and 2

	Pearsons/Spearmans Correlation Coefficient (r)					
	Aim 1			***Aim 2***		
	HipGT9X-Troiano vs HipGT9X-Sasaki	*HipGT9X-Troiano vs WrGT9X-Montoye*	*HipGT9X-Sasaki vs WrGT9X-Montoye*	*activPAL vs HipGT9X-Troiano*	*activPAL vs HipGT9X-Sasaki*	*activPAL vs WrGT9X-Monotye*
ST	.87 [b]	.76 [b]	.85 [b]	. 69 [b]	.84 [b]	.67 [b]
LPA	.40 [a]	.12	.70 [b]			
MVPA	.92 [b]	.38 [a]	.43 [a]			

Note: for Aim 1 n=30, for Aim 2 n=28. Pearsons correlation coefficient was used for all correlations except for MVPA for WrGT9X-Montoye (whereby Spearmans correlation analysis was conducted).
[a]= p<.05, [b]= p<.01, ST=Sedentary Time, LPA=Light Physical Activity, MVPA=Moderate to Vigorous Physical Activity.

Bland-Altman analysis assessing differences between uniaxial and triaxial hip cut-points (HipGT9X-Troiano vs HipGT9X-Sasaki) illustrated a mean bias of −45 min/day, and LoA of ±25min/day for ST (**Figure 4.4a**). This bias demonstrated HipGT9X-Troiano uniaxial cut-points estimated more ST compared to the HipGT9X-Sasaki triaxial cut-points. For LPA, a mean bias of 21min/day was illustrated in plots, with ±20min/day LoA (**Figure 4.4b**). Finally, for MVPA, analysis demonstrated 24min/day mean bias and ±12min/day agreement between cut-points (**Figure 4.4c**). For all activity intensities, some systematic bias was visible. Greater agreement between HipGT9X-Troiano uniaxial and HipGT9X-Sasaki triaxial cut-points was visualised in participants who spent a greater amount of time in ST, LPA, and MVPA.

Bland-Altman analysis between uniaxial HipGT9X-Troiano hip and triaxial WrGT9X-Montoye wrist cut-points can be visualised in **Figures 4.5a, 4.5b and 4.5c**. For ST, the HipGT9X-Troiano cut-points measured an average of 59min/day more ST than WrGT9X-Montoye cut-points, and ±36min/day LoA. Some systematic bias was illustrated, with greater agreement between cut-points in participants with greater ST. For LPA, HipGT9X-Troiano measured an average of 45min/day less time spent in LPA than WrGT9X-Montoye cut-points, with ±41min/day LoA. Finally for MVPA, greater average agreement was observed between cut-points, with mean bias of 15min/day and ±44min/day LoA.

Finally, Bland-Altman analysis comparing triaxial HipGT9X-Sasaki hip and triaxial WrGT9X-Montoye wrist cut-points illustrated some differences between cut-points and placements sites for measurements of ST, LPA and MVPA (**Figure 4.6a, 4.6b and 4.6c**). For ST, the mean bias was 14min/day, with ±30min/day LoA. This indicated that WrGT9X-Montoye measured less ST in participants than the HipGT9X-Sasaki. For LPA, the mean bias was 23min/day, and LoA were ±38min/day, with the WrGT9X-Montoye estimating more free-living LPA than the HipGT9X-Sasaki. MVPA plots displayed a mean bias of -9 min/day and LoA of ±42min/day. No obvious systematic bias was observed between the cut-points measurements for each different activity intensity.

Table 4.6: ICC value and 95% CI for Sedentary Time, Light Physical Activity and Moderate to Vigorous Physical Activity for Aims 1 and 2

	Aim 1			***Aim 2***		
	HipGT9X- Troiano vs HipGT9X-Sasaki	***HipGT9X-Troiano vs WrGT9X-Montoye***	***HipGT9X-Sasaki vs WrGT9X-Montoye***	***activPAL vs HipGT9X-Troiano***	***activPAL vs HipGT9X-Sasaki***	***activPAL vs WrGT9X-Montoye***
	ICC (95% CI)	***ICC (95% CI)***	***ICC (95% CI)***	***ICC (95% CI)***	***ICC (95% CI)***	***ICC (95% CI)***
ST	.29 (-.04, .68)	.18 (-.04, .53)	.74 (.18, .90)	.07 (-.05, .28)	.20 (-.08, .55)	.21 (-.10, .54)
LPA	.28 (-.08, .65)	.02 (-.05, .14)	.20 (-.10, .50)			
MVPA	.41 (-.03, .78)	.21 (-.09, .50)	.34 (.01, .61)			

Note: ICC= Intraclass Correlation Coefficient, 95% CI= 95% confidence interval, ST=Sedentary Time, LPA=Light Physical Activity, MVPA=Moderate to Vigorous Physical Activity, min/day= minutes per day.

Aim 2: ActiGraph validity

Table 4.7 presents the means and SD for ST for HipGT9X-Troiano, HipGT9X-Sasaki, WrGT9X-Montoye and activPAL for the n=28 valid participants. Paired samples t-tests revealed the uniaxial HipGT9X-Troiano hip, triaxial HipGT9X-Sasaki hip, and triaxial WrGT9X-Montoye wrist cut-points all gave significantly higher measurements of daily ST compared to the activPAL. The 95% CI for the true mean difference were −157.22 to −118.62 min/day, −108.67 to −75.96 min/day, and −95.92 to −59.11 min/day for HipGT9X-Troiano, HipGT9X-Sasaki, WrGT9X-Montoye, respectively. The greatest

differences from activPAL ST estimates were observed for uniaxial Troiano hip cut-points, and smallest for triaxial Montoye wrist cut-points.

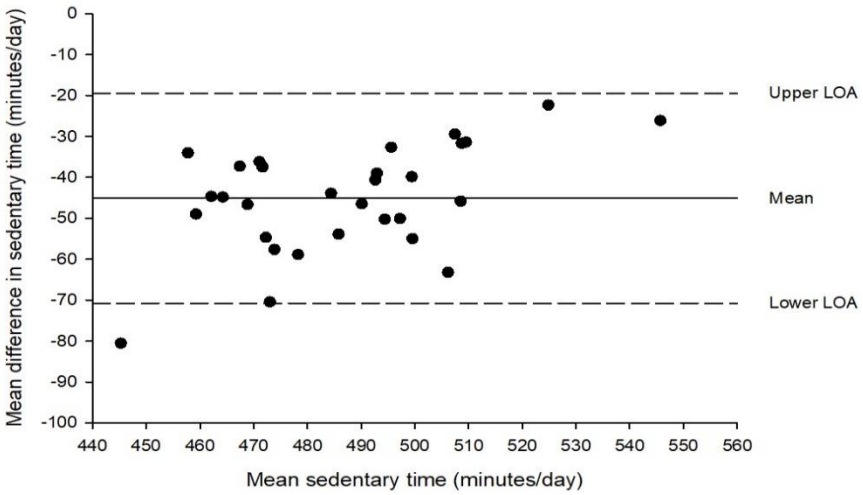
Table 4.7: Means and standard deviations for sedentary time for HipGT9X-Troiano, HipGT9X-Sasaki, WrGT9X-Montoye and activPAL

	HipGT9X-Troiano	*HipGT9X-Sasaki*	*WrGT9X- Montoye*	*activPAL*
ST (min/day)	507.90 (19.27) [a]	462.30 (25.05) [a]	447.50 (26.78) [a]	369.98 (61.02)

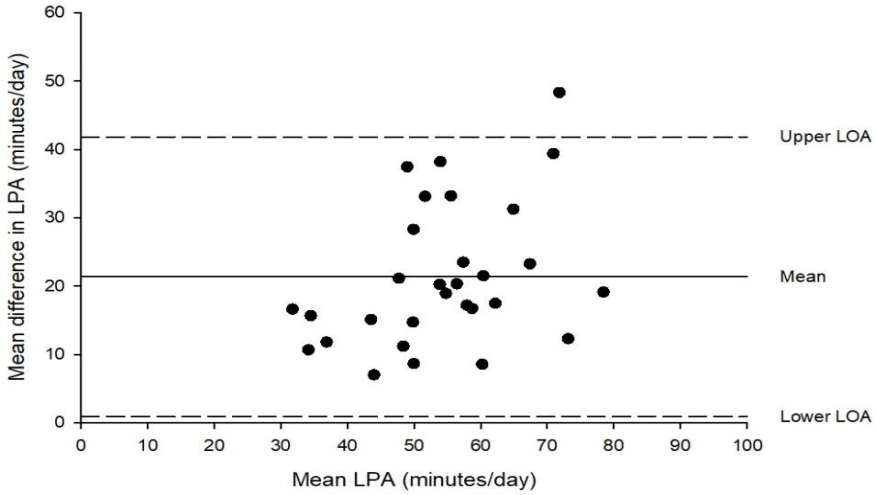
Note: n=28. [a] = paired-samples t-test results showing significant difference between uniaxial hip (HipGT9X-Troiano), triaxial hip (HipGT9X-Sasaki), or triaxial wrist (WrGT9X-Montoye) cut-points with activPAL for measurements of ST ($p<.05$). ST=Sedentary Time, min/day= minutes per day.

Pearsons correlation analysis (**Table 4.5**) revealed significant correlations between ST quantified with all GT9X device placement and corresponding cut-points with activPAL estimates of ST (r= .67 to .84). However, ICCs demonstrated no to low agreement, with ICCs of .07, .20, and .21 for the HipGT9X-Troiano, HipGT9X-Sasaki, and WrGT9X-Montoye, respectively (**Table 4.6**).

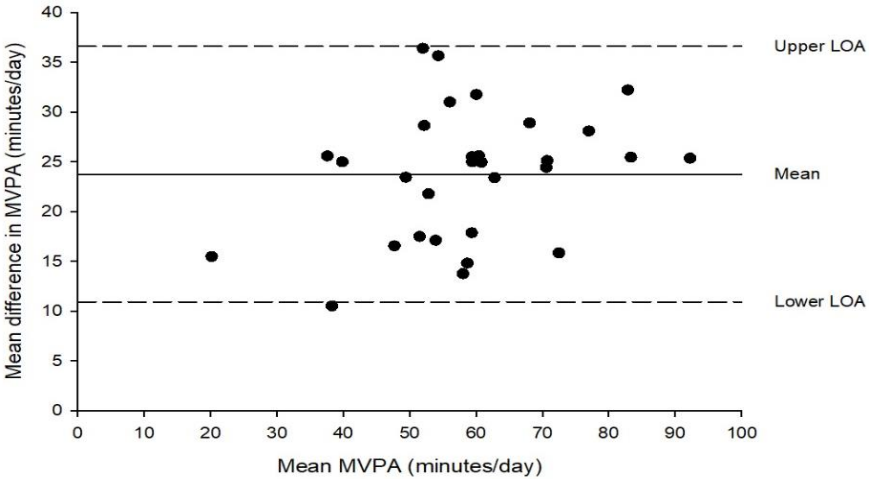
Bland-Altman plots for the different GT9X cut-points plotted against the activPAL for estimates of ST can be visualised in **Figures 4.7a, 4.7b and 4.7c**. In brief, plots illustrate that all GT9X cut-points estimated consistently greater amount of ST compared to the activPAL. Some systematic bias was also present, with greater agreement between accelerometers at higher measurements of ST. Specifically, for the uniaxial HipGT9X-Troiano hip cut-points, plots demonstrated a mean bias of 139min/day and ±96min/day LoA (**Figure 4.7a**). The triaxial HipGT9X-Sasaki hip cut-points had a mean bias of 93min/day and LoA within ±82min/day (**Figure 4.7b**). Finally, the triaxial WrGT9X-Montoye wrist cut-points displayed a mean bias of 78min/day and LoA within ±93min/day (**Figure 4.7c**).



4.4a.) *ST agreement between HipGT9X-Troiano and HipGT9X-Sasaki*



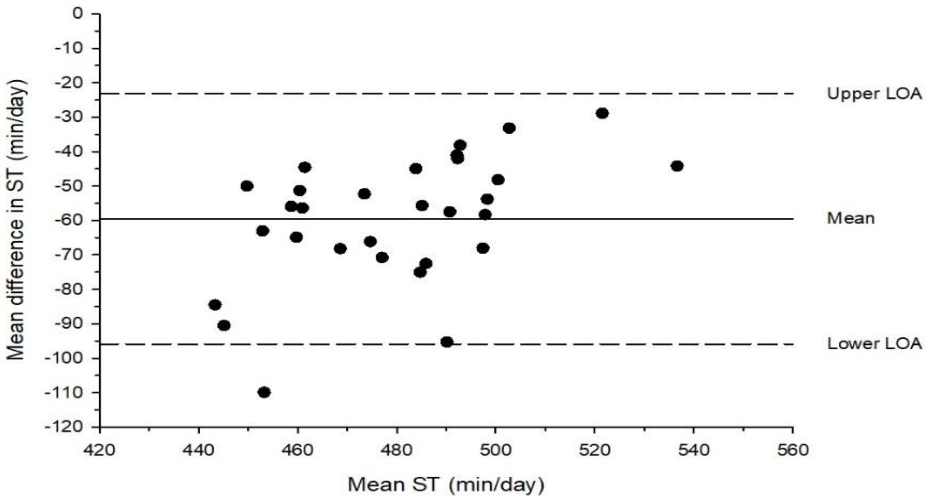
4.4b.) *LPA agreement between HipGT9X-Troiano and HipGT9X-Sasaki*



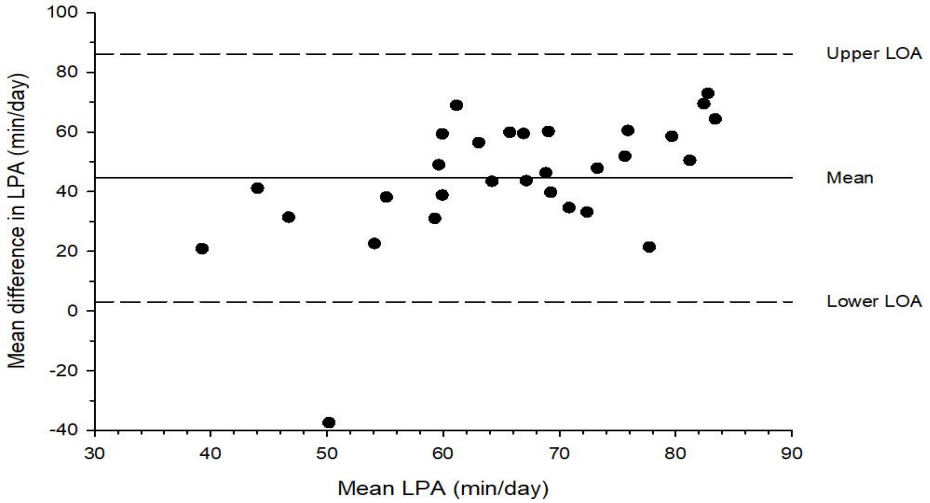
4.4c.) *MVPA agreement between HipGT9X-Troiano and HipGT9X-Sasaki*

Figure 4.4: Bland Altman plots showing differences between HipGT9X-Troiano and HipGT9X-Sasaki for a.) Sedentary Time; b.) Light Physical Activity, and c.) Moderate to Vigorous Physical Activity

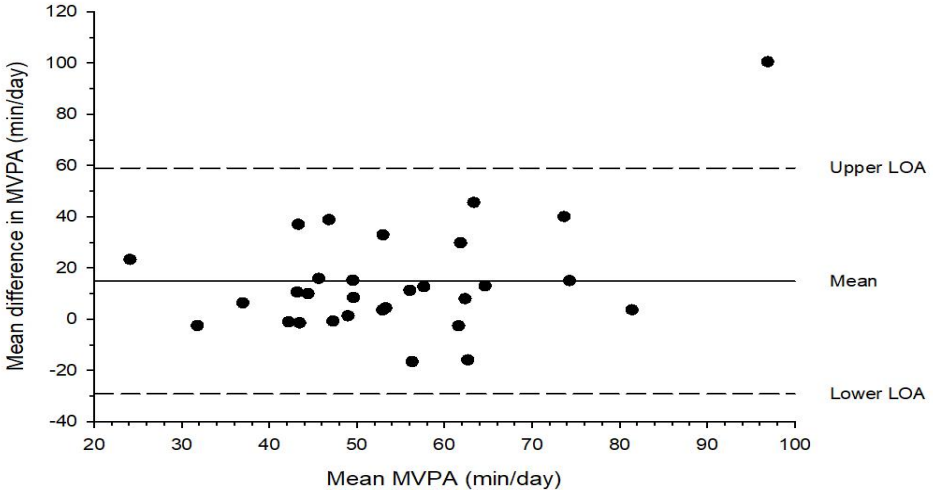
Note: HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, LPA= Light intensity Physical Activity, MVPA= Moderate to Vigorous Physical Activity, min/day= minutes per day, Solid line= mean difference/bias between the two accelerometers, dashed lines= 95% limits of agreement (LoA)



4.5a.) *ST agreement between HipGT9X-Troiano and WrGT9X-Montoye*



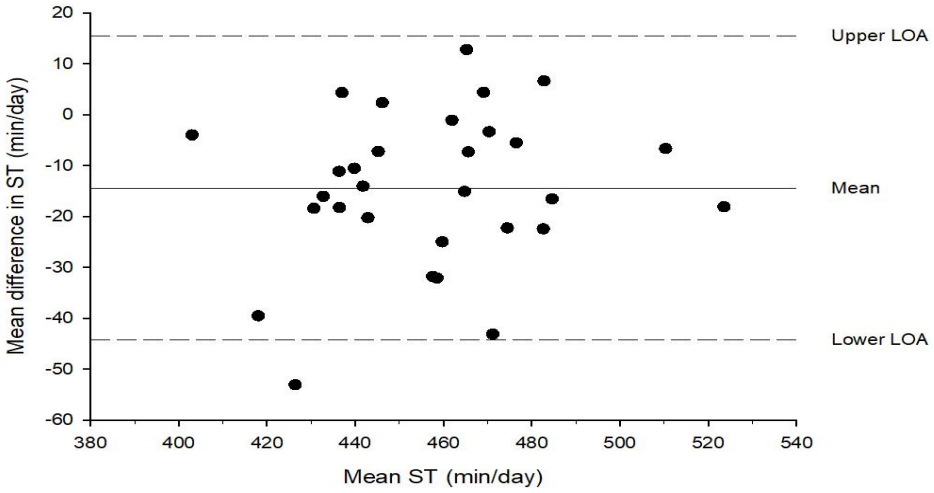
4.5b.) *LPA agreement between HipGT9X-Troiano and WrGT9X-Montoye*



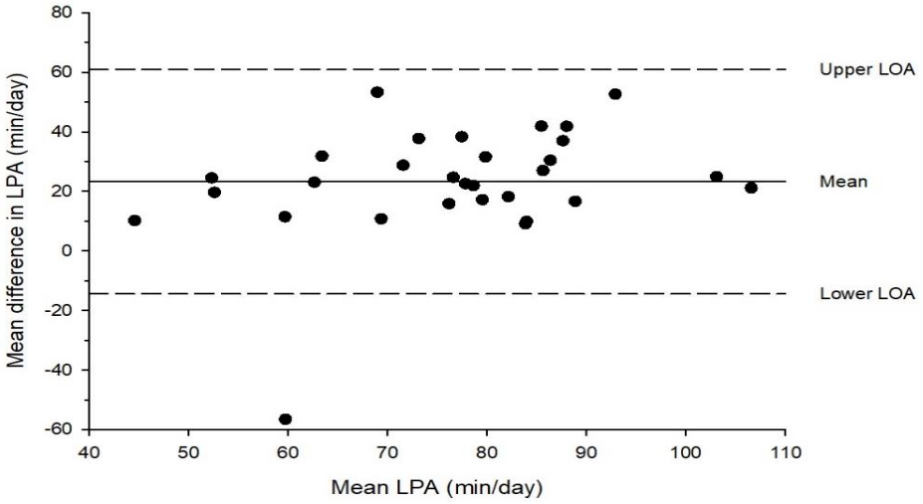
4.5c.) *MVPA agreement between HipGT9X-Troiano and WrGT9X-Montoye*

Figure 4.5: Bland Altman plots showing differences between HipGT9X-Troiano and WrGT9X-Montoye for: a.) Sedentary Time; b.) Light Physical Activity, and c.) Moderate to Vigorous Physical Activity

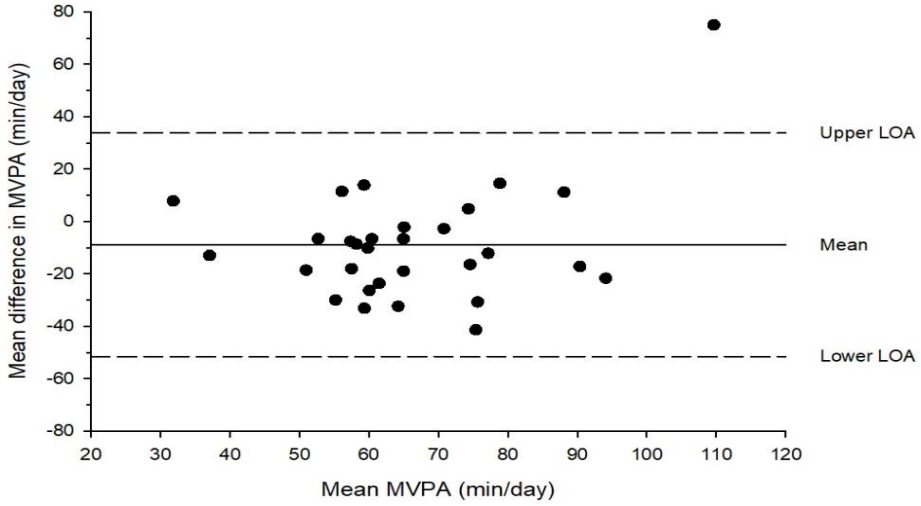
Note: HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, LPA= Light intensity Physical Activity, MVPA= Moderate to Vigorous Physical Activity, min/day= minutes per day, Solid line= mean difference/bias between the two accelerometers, dashed lines= 95% limits of agreement (LoA)



4.6a.) *ST agreement between HipGT9X-Sasaki and WrGT9X-Montoye*



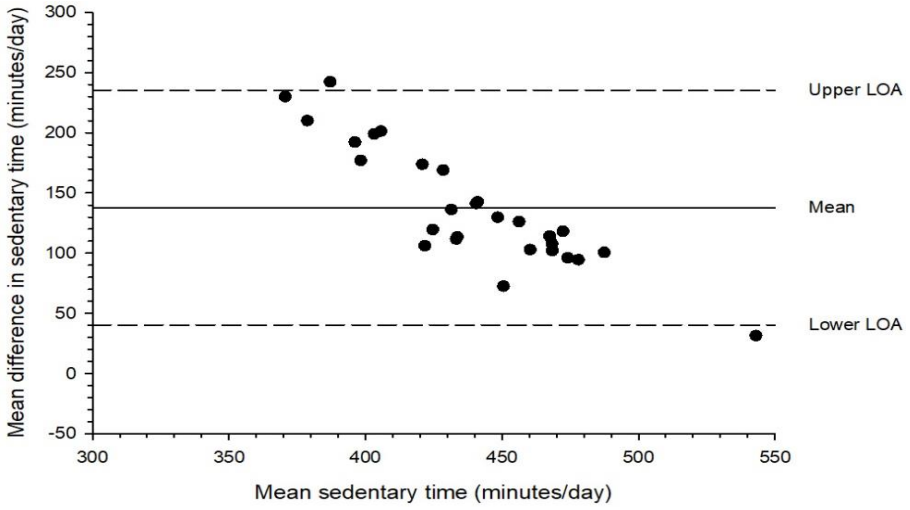
4.6b.) *LPA agreement between HipGT9X-Sasaki and WrGT9X-Montoye*



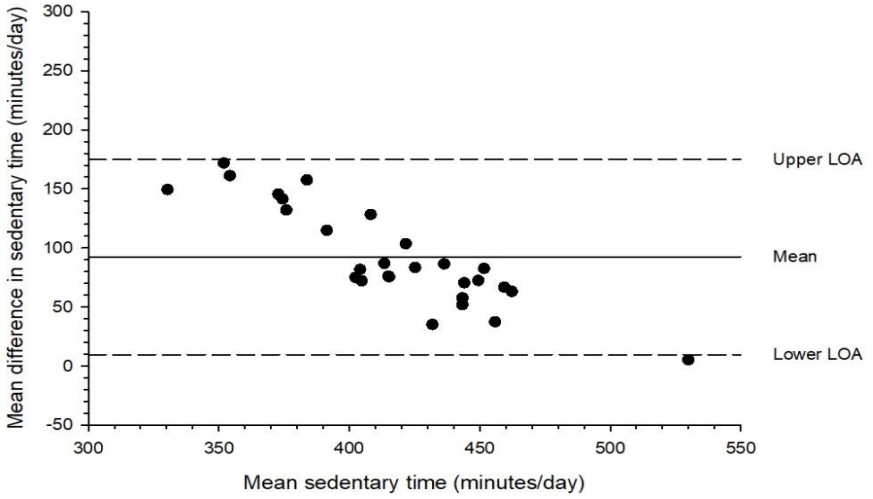
4.6c.) *MVPA agreement between HipGT9X-Sasaki and WrGT9X-Montoye*

Figure 4.6: Bland Altman plots showing differences between HipGT9X-Sasaki and WrGT9X-Montoye for: a.) Sedentary Time; b.) Light Physical Activity, and c.) Moderate to Vigorous Physical Activity

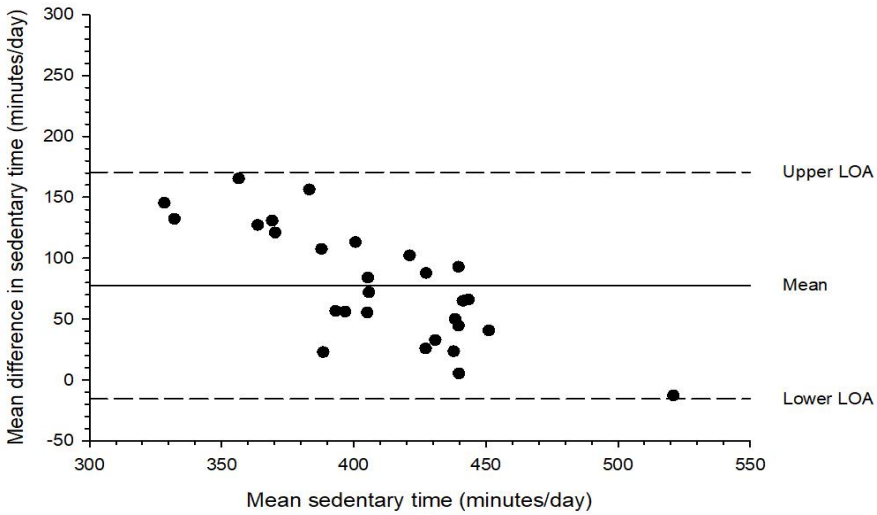
Note: HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, LPA= Light intensity Physical Activity, MVPA= Moderate to Vigorous Physical Activity, min/day= minutes per day, Solid line= mean difference/bias between the two accelerometers, dashed lines= 95% limits of agreement (LoA)



4.7a.) *ST agreement between HipGT9X-Troiano and activPAL*



4.7b.) *ST agreement between HipGT9X-Sasaki and activPAL*



4.7c.) *ST agreement between WrGT9X-Montoye and activPAL*

Figure 4.7: Bland Altman plots showing differences between a.) HipGT9X-Troiano; b.) HipGT9X-Sasaki and c.) WrGT9X-Montoye with activPAL for Sedentary Time (Aim 2)

Note: HipGT3X= hip-worn ActiGraph GT3X, HipGT9X= hip-worn ActiGraph GT9X, ST= Sedentary Time, Solid line= mean difference/bias between the two accelerometers, dashed lines= 95% limits of agreement (LoA)

Discussion

The purpose of this study was to compare the estimates of PA and ST according to data capture methods (i.e., cut-points applied) and device placement (i.e., hip or wrist placement). In addition this study assessed the validity of these data capture methods and device placements for assessment of ST, compared to the activPAL.

Aim 1: ActiGraph comparability

This study found significant differences between estimates of PA and ST when either the uniaxial Troiano or triaxial Sasaki hip cut-points were applied to the hip-worn GT9X data. In addition, significant differences were displayed between PA measurements from uniaxial Troiano hip and triaxial Montoye wrist cut-points. These findings are in agreement with study hypotheses, and were confirmed by ICCs and Bland-Altman plots displaying low agreement and substantial bias between activity estimates from the triaxial vs uniaxial, and wrist vs hip cut-points. The particular lack of agreement between uniaxial with triaxial cut-points in this study strongly suggest that triaxial and uniaxial cut-points generate significantly different estimates of PA, and should not be compared or used interchangeably within or between studies. This is particularly important when we consider the vast amount of studies which compare their data acquired with triaxial accelerometers and analysed with corresponding triaxial cut-points, with epidemiological datasets often employing uniaxial accelerometers and analytical procedures to quantify PA and ST (e.g., NHANES,(Matthews et al., 2008)) (Duncan et al., 2020; Koster et al., 2016; O'Brien et al., 2020).

In this study, the triaxial cut-points (Sasaki and Montoye) generated lower estimates of ST and higher estimates of LPA and MVPA than the uniaxial (Troiano) cut-points. This agrees with hypotheses based on previous studies of children, adults and older adults, reporting lower levels of ST and higher levels of LPA and MVPA quantified when using triaxial cut-points, compared to uniaxial cut-points (Kozey-

Keadle et al., 2014; Leeger-Aschmann et al., 2019; Luzak et al., 2017; Sagelv et al., 2019). Considering the relative accuracy of PA and ST measured by uniaxial vs triaxial accelerometers, research suggests the additional information available from 3 axes of data in triaxial cut-points may increase the accuracy of accelerometer outputs (Kozey-Keadle et al., 2014). Specifically, it is proposed that the higher levels of LPA and MVPA captured by triaxial accelerometers is due to higher acceleration estimates for Z (forward-backward) and X (left-right) axes (Kozey-Keadle et al., 2014; Leeger-Aschmann et al., 2019). As such, cut-points which only use vertical axis data (uniaxial) may not be valid tools to categorise certain behaviours and activities involving movement in 3 dimensions in all population groups (Leeger-Aschmann et al., 2019; Sasaki et al., 2011). However, additional research is needed to provide a consensus on the optimal methods and common approach to analyse PA from accelerometers. Given that uniaxial Troiano et al. (2008) cut-points are widely used in epidemiological research (Matthews et al., 2008), this does not emphatically imply that these cut-points are therefore valid and reliable for application to all models of accelerometer, worn at any placement site.

Regarding the comparability of the hip (Sasaki) vs wrist (Montoye) triaxial cut-points, my results displayed acceptable correlations. These were not all transferred to ICCs, which demonstrated low agreement for LPA and MVPA and moderate agreement for ST. In addition, Bland Altman plots demonstrated that triaxial Montoye wrist cut-points estimated less time spent in ST and MVPA and greater amount of LPA compared to the triaxial Sasaki hip cut-points. Study hypotheses were that significant differences between cut-points would be visualised, but the direction of these differences was unknown due to the fact that these cut-points have not been widely used or validated in multiple populations. For example, validation of the Montoye wrist cut-points involved a small sample of participants of various ages, wearing accelerometers for a short period of time (up to 8 hours) (Montoye et al., 2020). The results of this study agree with hypotheses and extend current knowledge on the comparability and differences between accelerometer wrist vs hip cut-points. This

limits generalisability of their findings, as validation may not have reflected all activities of daily life or captured habitual activity patterns (Montoye et al., 2020). Although triaxial Sasaki hip GT9X cut-points have been more widely used as a comparator or criterion in cut-point validation studies (Santos-Lozano et al., 2013), they were designed specifically for the ActiGraph GT3X model worn on the hip. Recent research has indicated that cut-points are device-dependent (Rhudy et al., 2020; Sasaki et al., 2011). So this suggest the Sasaki, Montoye, and Troiano cut-points may not be appropriate for different ActiGraph devices to the ones they were specifically designed for. Application of cut-points to different devices or placement sites could potentially lead to activity misclassification (Rhudy et al., 2020). As there are significant limitations in the design, validation and use of each of these cut-points, further research is required to elucidate and recognise the optimal cut-points for each ActiGraph model and placement site. Only then can comparisons between studies using these cut-points applied to accelerometers be deemed valid. Furthermore, caution should be applied when comparing PA intensity data between studies involving different accelerometer cut-points, even if placement site and the make and model of accelerometer used are the same (Loprinzi & Smith, 2017).

Aim 2: ActiGraph validity

With regards to the validity of the different GT9X cut-points for estimating ST, these results demonstrated no to low agreement with the activPAL for all cut-points. The uniaxial Troiano hip, triaxial Sasaki hip and triaxial Montoye wrist cut-points all significantly overestimated ST compared to the activPAL, in agreement with the study hypotheses. However, greater agreement for all cut-points was visualised where participants had recorded higher daily ST (**Figures 4.7a, 4.7b, and 4.7c**). Growing research has recognised that many accelerometer hip cut-points are too high, and thus misclassify standing time as ST (Aguilar-Farías et al., 2014; Koster et al., 2016). However, to my knowledge, no previous studies have assessed the validity of new triaxial Sasaki hip and Montoye

wrist cut-points, and compared their validity with uniaxial Troiano hip cut-points, for estimates of ST in a free-living setting, compared to the activPAL.

In this study, the widely used uniaxial Troiano hip cut-points resulted in the hip-worn monitors overestimating daily ST by over 2 hours compared with the activPAL. My finding that triaxial Sasaki hip cut-points had greater agreement with the activPAL than uniaxial Troiano cut-points agrees with previous research by Koster et al. (2016). Studies have indicated that the Troiano ST cut-points (<100cpm) are too high to reliably assess ST, with an alternative cut-point for ST for the vertical axis suggested as <22cpm or <25cpm (Aguilar-Farías et al., 2014; Koster et al., 2016). In comparison, triaxial hip cut-points have demonstrated greater specificity and sensitivity at measuring ST compared to uniaxial hip cut-points (Koster et al., 2016). Furthermore, as triaxial Sasaki hip cut-points for ST (<150cpm) displayed greater agreement with activPAL in this study, relative to the uniaxial Troiano cut-points, but still overestimated ST, this indicates that future research should concentrate on elucidating optimal triaxial cut-points to reliably assess ST.

Both the triaxial Sasaki hip and Montoye wrist cut-points gave similar estimates of daily ST in this study, despite both showing low agreement with the activPAL. The triaxial Montoye wrist ST cut-points (<2860cpm) were higher than other published triaxial cut-points developed on the non-dominant wrist, which have previously shown high agreement with the activPAL ((Koster et al., 2016) (<1853cpm)). However Koster et al. (2016) cut-points were developed for a previous ActiGraph model. As far as we are aware, the Montoye cut-points are the first triaxial ST cut-points developed specifically for the wrist-worn GT9X. Although they significantly overestimated ST compared to the activPAL, they have not been widely used in research. In addition, they were validated in a small sample of participants and aims of the validation study were to develop cut-points primarily for PA estimation, not for assessing ST (Montoye et al., 2020). Previous findings have concluded that cut-points need to be age group, axis and epoch length specific (Aguilar-Farías et al., 2014; Leeger-

Aschmann et al., 2019). Therefore, future research should refine current or develop new triaxial ST cut-points for wrist- and hip-worn GT9X accelerometers in order to give valid assessments of SB.

Limitations

This study was not without limitations. The initial sample size of n=37 was relatively small, with n=7 and n=9 participants data excluded for Aims 1 and 2, respectively. Future validation and comparability studies should aim to recruit larger numbers of participants in order to ensure sufficient power for statistical analysis employed herein. In addition, participants were young, healthy adults, with no disabilities or walking difficulties. This was due to this population matching those used to develop the Montoye, Troiano and Sasaki cut-points. Therefore, the results of this study can only be applied to healthy adult populations, and additional research is required to establish the validity and comparability of these cut-points and this protocol among people of different age groups, and patient or diseased populations.

As the focus of this thesis is on people with RA, it is important to consider the implications of these findings in this population. A large number of accelerometer studies in people with RA use the existing Troiano uniaxial hip cut-points (Fenton et al., 2020b; Pinto et al., 2020a; Summers et al., 2019), and only recently have RA-specific cut-points been developed, validated and shown to be superior to Troiano cut-points for estimation of PA and ST (O'Brien et al., 2020). Findings from this chapter emphasise the need for wear site and accelerometer model specific cut-points in healthy adults alone, and this may also be the case in other clinical populations, such as RA. Therefore, I recommend the need for caution when applying existing cut-points to accelerometers, and validity and reliability should be ensured before use in each different population. As my findings indicate cut-points should be designed and validated for every population, wear site and accelerometer model of interest, perhaps other methods of accelerometer data processing, such as using raw accelerometer data may be a more feasible and viable alternative (Rowlands et al., 2018).

Conclusion

To conclude, estimates of PA and ST quantified by ActiGraph GT9X accelerometers are not comparable where different protocols for data capture (uniaxial vs triaxial) and device placement (hip vs wrist) are employed. This is the case even where placement specific triaxial cut-points are used. Secondly, these cut-points displayed poor validity at assessing free-living ST, compared to the activPAL. Results of this study highlight the importance of acknowledging that PA and ST outcomes reported in studies are perhaps an artefact of the protocol and analytical decisions employed in analysing accelerometer data. Finally, further research is needed to adjust current or develop new ST cut-points for the hip- and wrist-worn GT9X using triaxial data, in order for the ActiGraph GT9X to provide valid assessments of free-living SB.

Declarations

AUTHOR CONTRIBUTORSHIP: Conception and design of the study: Sophia Brady, Sally Fenton, Jet Veldhuijzen van Zanten. Data acquisition: Sophia Brady. Data analysis: Sophia Brady, Sally Fenton. Data interpretations and drafting of manuscript: all authors. Final approval of manuscript: all authors.

CHAPTER 5: RELATIONSHIPS BETWEEN PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOUR WITH CORE PATIENT- AND CLINICIAN-IMPORTANT OUTCOMES IN PEOPLE WITH RHEUMATOID ARTHRITIS

Findings from Chapters 2, 3 and 4

Rheumatoid Arthritis (RA) is an autoimmune disease, characterised by high-grade systemic inflammation (Testa et al., 2021). Symptoms include joint pain, fatigue, poor mental health and psychological wellbeing, and these can further lead to severe functional impairment and disability (Smolen et al., 2016; Testa et al., 2021). In this thesis, there is a focus on core patient- and clinician-important health outcomes, defined by Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT): pain, disease activity, functional disability, fatigue, depression, anxiety, subjective vitality and quality of life (Bartlett et al., 2012; Boers et al., 1994; Kirwan et al., 2007; Van Tuyl & Boers, 2015).

Physical activity (PA) has been advocated by the European Alliance of Associations for Rheumatology (EULAR) and American College of Rheumatology (ACR) for people with RA to manage these core OMERACT outcomes (Rausch-Osthoff et al., 2018). The majority of previous intervention studies conducted in participants with RA explored the effects of exercise or moderate to vigorous intensity PA (MVPA). However, RA-related symptoms, such as pain, fatigue and poor functional ability, are frequently reported as barriers to MVPA (Veldhuijzen Van Zanten et al., 2015). Therefore, recent intervention studies have explored the effects of increasing overall PA and reducing sedentary behaviour (SB) via interventions which aim to incorporate more PA into daily life (Swardh et al., 2020). This approach targets lifestyle PA, which may be more feasible for people with RA. Lifestyle PA interventions can include increasing engagement in all activities of daily living, incidental PA, or walking. In addition, reducing SB can also result in increased lifestyle PA, as less time spent sitting will result in increased total PA (Thomsen et al., 2017).

Therefore, ***Chapter 2*** of this thesis describes a systematic review and meta-analysis conducted to assess the efficacy of existing lifestyle PA and SB interventions at increasing PA, reducing SB, and improving OMERACT patient- and clinician-important health outcomes in people with RA. A total of

n=16 interventions were included in this review, and results revealed significant positive effects of interventions on disease activity, functional ability, fatigue, depression, leisure/light intensity PA, steps, MVPA and sedentary time. However, interventions were particularly heterogeneous with regards to measurement tools and outcomes assessed. In addition, only 1 SB intervention was identified, in which baseline differences in daily sitting time, pain and fatigue were observed between the intervention and control groups (Thomsen et al., 2017).

Research investigating the links between lifestyle PA and SB with OMERACT patient- and clinician-important health outcomes in people with RA is still in its infancy. Currently, few cross-sectional studies have assessed associations between different dimensions and elements of PA and SB (i.e., FITT, SITT) with these health outcomes. One of the conclusions of this review was that a greater number of cross-sectional and longitudinal studies, using validated and reliable measurement tools, are needed in people with RA. This will enable a greater understanding of the nuances of relationships between different dimensions of PA and SB with RA-related health. In addition, the majority of studies included in the systematic review used self-report measures of PA and SB, and single-item measures for health outcomes, such as pain. Therefore, as part of ***Chapter 3*** and ***4*** of this thesis, this limitation was addressed. This involved assessment of the reliability and comparability of multidimensional psychophysical measurement methods of pain, and device-based measures of PA and SB.

In detail, ***Chapter 3*** involved a study assessing the test-retest and inter-rater reliability of 3 modalities of Quantitative Sensory Testing (QST) in a sample of RA, as well as healthy participants and individuals with chronic lower back pain (LBP). Findings revealed that a protocol consisting of pressure-pain threshold (PPT), temporal summation (TS) and conditioned pain modulation (CPM) modalities can reliably be conducted on the forearm or leg by the three study raters involved in the research in healthy people and individuals with RA and LBP.

Chapter 4 provides details of a study assessing the comparability of free-living PA and sedentary time estimations obtained from application of different data capture methods and placement sites of the ActiGraph GT9X accelerometer in a sample of healthy participants. The hip- and wrist-worn GT9X accelerometers, with different cut-points applied, were also validated against the activPAL™ for assessments of sedentary time as a secondary aim. Conclusions indicated excellent agreement and little difference in sedentary time and PA measurements between GT9X and GT3X ActiGraph accelerometers when devices were attached to the same body site, and identical data capture methods were applied (***Chapter 4*** prelude). In addition, hip- and wrist-worn GT9X accelerometers demonstrated poor validity for quantifying sedentary time, compared to the activPAL (***Chapter 4*** Aim 2). It can therefore be reasonably assumed that the GT9X can be reliably used to assess free-living PA in populations where the GT3X has been previously validated, such as people with RA (O'Brien et al., 2020). Moreover, as the activPAL has been validated for assessment of free-living sedentary time in people with RA (O'Brien et al., 2020), ***Chapter 4*** results indicate that the activPAL and the ActiGraph GT9X can be used to quantify free-living sedentary time and PA, respectively, in people with RA.

As results of ***Chapter 3*** and ***4*** demonstrate that QST and accelerometers are reliable, they can therefore be used in future cross-sectional, longitudinal and intervention studies. Although there is some evidence of a link between lifestyle PA and SB with health in people with RA (Fenton et al., 2018b; Khoja et al., 2016; O'Brien et al., 2021; O'Leary et al., 2021; Veldhuijzen Van Zanten et al., 2020), associations between lifestyle PA and SB with core OMERACT outcomes need further exploration before the development of further interventions. To my knowledge, relationships between different dimensions (i.e., types and intensities) of PA (e.g., light intensity PA (LPA), MVPA, lifestyle PA) and SB with OMERACT outcomes have not been extensively investigated through large-scale longitudinal research in people with RA (O'Brien et al., 2021). Therefore, it is essential to collect further high-quality evidence to further understand the links between PA and SB with health in this population.

The results of methodological chapters of this thesis (***Chapters 2, 3*** and ***4***) were to inform the methods and design of a longitudinal study (**Figure 5.2**). Specifically, the QST modalities, ActiGraph GT9X and activPAL, used as part of ***Chapters 3*** and ***4***, would be further used to assess pain outcomes and PA and SB, respectively, as part of this planned longitudinal study. This would add to the existing body of research assessing the links between different dimensions of PA (i.e., LPA, MVPA, lifestyle PA) and SB with pain and other core OMERACT outcomes in people with RA. This study would have also added the unique dimension of assessing multiple mechanisms involved in RA-related pain (through QST) and had accurate device-based measures of PA and SB (through the ActiGraph GT9X and activPAL). In addition, by conducting data-prompted interviews as part of this longitudinal study, qualitative data would aid to further increase the understanding of the key determinants, facilitators and barriers people with RA experience to PA and SB.

Therefore, major aims (introduction aims 4 and 5) of the next study were:

1. To conduct a large-scale longitudinal observational study to assess the presence and magnitude of relationships between ActiGraph GT9X-assessed dimensions of PA (i.e., LPA, MVPA, lifestyle PA) and activPAL-assessed SB with QST-assessed pain and other OMERACT patient- and clinician-important health outcomes in people with RA.
2. To conduct data-prompted interviews as part of this longitudinal study, to understand the determinants, facilitators and barriers to PA and SB for people with RA.
3. To use the quantitative and qualitative data from the longitudinal study, to design and compare the initial feasibility of two interventions targeting; 1) increasing MVPA through structured exercise vs 2) increasing total PA through increasing lifestyle PA and reducing SB.

Planned Longitudinal Study

Briefly, the methods and protocol of the planned longitudinal study involved recruitment of patients with RA from Rheumatology outpatient clinics at Russells Hall Hospital (RHH), Dudley Group NHS Trust, Birmingham. Participants would be asked to visit the Clinical Research Unit at RHH on 4 occasions/visits over the 6-7 month study period, for completion of 2 study weeks (see **Figure 5.1** for a timeline of visits). Each study week would take place over 7 days, and occur 6 months apart. During Visits 1 and 3, patients would be asked to complete a pack of validated self-report questionnaires assessing the following OMERACT patient- and clinician-important health outcomes: pain, fatigue, functional ability, sleep, quality of life, mental health and psychological wellbeing and motivation. The researcher conducting the study (Sophia Brady) would also conduct PPT, TS and CPM QST assessments (validated as part of ***Chapter 3***), take a fasted blood sample, measure resting blood pressure, disease activity (using the DAS28), height, weight and body composition. In addition, the 20-metre timed walk test would be assessed, as an additional measure of functional ability.

Figure 5.1: Timeline of visits to Russells Hall Hospital as part of the longitudinal study

At the end of Visits 1 and 3, participants would then be asked to wear 1 x hip-worn ActiGraph GT9X and 1 x thigh-worn activPAL for 7 days, to obtain device-based measures of PA and SB, respectively (validated as part of ***Chapter 4***). ActiGraph-measured PA data would be used to quantify frequency, intensity and duration of PA (i.e., FITT components), including total PA accumulated as part of daily

life (i.e., lifestyle PA, accrued through activities of daily living, incidental PA etc.). In addition, participants would be given a PA diary and log book in order to give details about accelerometer wear and non-wear periods, and provide context for researcher interpretation of PA participation. A member of the research teamwould fit the monitors and provide instructions on how to wear them.

A week later, during Visits 2 and 4 (**Figure 5.1**), participants would be asked to return accelerometers and to complete questionnaires asking about pain and fatigue over the past 7 days (i.e., the days they were wearing devices). As part of Visit 2 and 4, data prompted interviews would be conducted with participants to discuss facilitators, barriers and participation in different dimensions of PA and SB. Qualitative research exploring how lifestyle PA and SB is experienced and performed in this specific population is scarce (Thomsen et al., 2015). In order for future lifestyle PA or SB interventions to be effectively designed, qualitative data is needed to give an in-depth understanding of the personal and specific determinants, barriers and facilitators towards PA and SB in individuals with RA (Larkin et al., 2017; Thomsen et al., 2015; Veldhuijzen Van Zanten et al., 2015). In this study, qualitative interviews would involve people living with RA being shown their activPAL data from the study week in order to prompt discussions and stimulate conversations regarding their experience of PA and SB, and the factors which influence their participation in these activities. Specifically, data prompted interviews would explore their perceived facilitators and barriers to increasing PA and/or reducing SB, with a focus on different types and intensities of PA (i.e., structured exercise, LPA, MVPA, lifestyle PA) and SB.

Repeating Visits 1 and 2 after 6 months (i.e. Visits 3 and 4) would allow for sufficient time to observe natural changes in behaviours and health outcomes. This is needed in order to detect relationships over time between PA and SB with OMERACT patient- and clinician-important outcomes in RA. The data gathered through results of the longitudinal observational study and associated qualitative research were intended to inform the design of subsequent interventions targeting; 1) increasing

MVPA through structured exercise; and 2) increasing total PA through increasing lifestyle PA and reducing SB. These interventions would be evaluated in terms of; 1) their feasibility and acceptability to patients; 2) their potential for improving OMERACT patient- and clinician-important outcomes.

Recruitment for the longitudinal study began in March 2020, and was due to be completed by the end of 2020. We were able to successfully recruit and complete Visit 1 and 2 assessments on 1 participant. However, due to the COVID-19 pandemic, this longitudinal study could not continue as planned. Restrictions were placed on day to day life, with people only allowed to leave their house for basic necessities, daily exercise, and for essential hospital visits. The meant that all non-COVID-related clinical research was halted throughout 2020, with further lockdowns restricting research in 2021. Therefore, the longitudinal research study previously described as the main experimental chapter as part of this thesis, was unable to proceed (**Figure 5.2**).

As the main aims of this longitudinal study were to assess relationships between different dimensions of PA (i.e., LPA, MVPA, lifestyle PA) and SB with core OMERACT outcomes using quantitative and qualitative research methods, by adjusting measurement tools, a similar study was able to be undertaken. Using self-report measurement methods, an online cross-sectional study was conducted to assess the links between different dimensions of PA (non-exercise LPA, exercise and walking) and SB with pain, fatigue, anxious and depressive symptoms and subjective vitality, and the impact of COVID-19 on these associations. The next chapter of this thesis (***Chapter 6***) presents a study conducted in patients with RA during the COVID-19 pandemic.

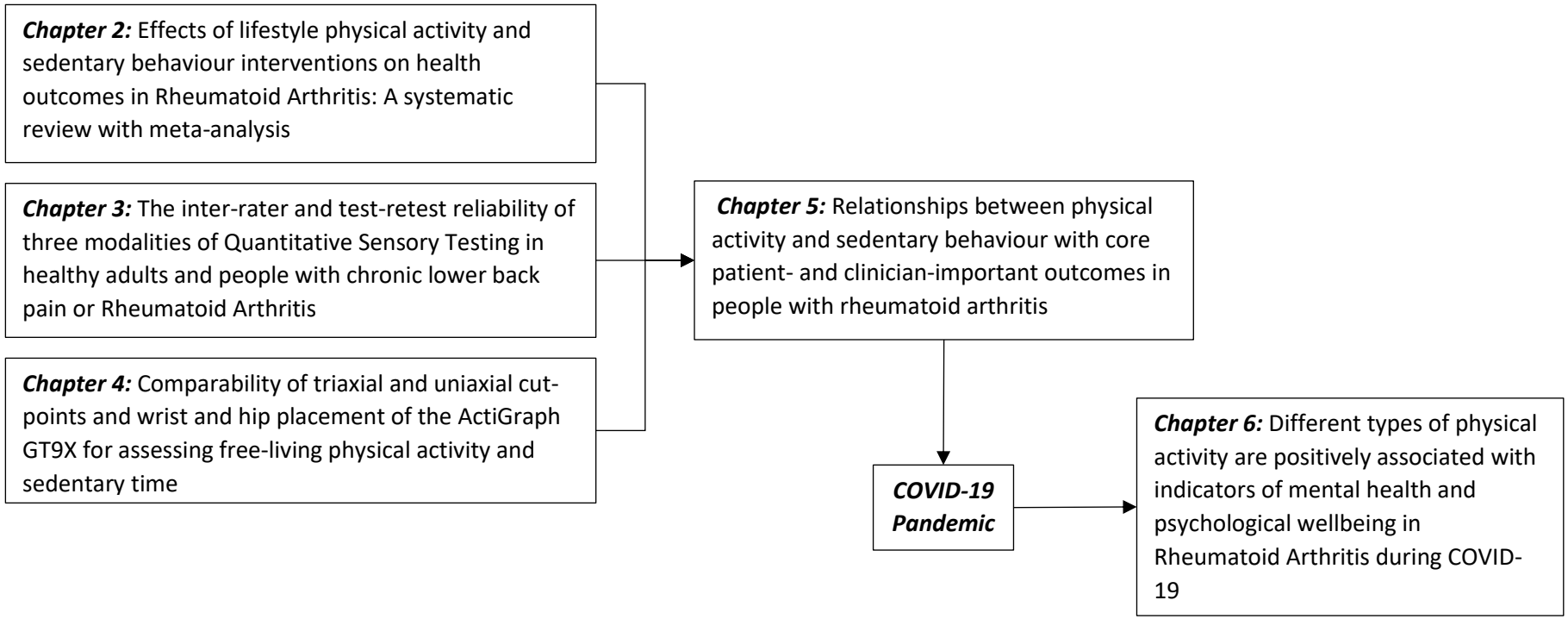


Figure 5.2: Flowchart of experimental chapters as part of this thesis

CHAPTER 6: DIFFERENT TYPES OF PHYSICAL ACTIVITY ARE POSITIVELY ASSOCIATED WITH INDICATORS OF MENTAL HEALTH AND PSYCHOLOGICAL WELLBEING IN RHEUMATOID ARTHRITIS DURING COVID-19

Abstract

Nationwide lockdowns during SARS-CoV-2 (COVID-19) can compromise mental health and psychological wellbeing and limit opportunities for physical activity (PA), particularly in clinical populations, such as people with rheumatoid arthritis (RA), who are considered at risk for COVID-19 complications. This study aimed to investigate associations between PA and sedentary time (ST) with indicators of mental health and wellbeing in RA during COVID-19 lockdown, and examine the moderation effects of self-isolating. N=345 RA patients completed an online questionnaire measuring PA (NIH-AARP Diet and Health Study Questionnaire), ST (International Physical Activity Questionnaire-Short Form), pain (McGill Pain Questionnaire and Visual Analogue Scale), fatigue (Multidimensional Fatigue Inventory), depressive and anxious symptoms (Hospital Anxiety and Depression Scale), and vitality (Subjective Vitality Scale) during the United Kingdom COVID-19 lockdown. Associations between PA and ST with mental health and wellbeing were examined using hierarchical multiple linear regressions. Light intensity PA (LPA) was significantly negatively associated with mental fatigue (β = − .11), depressive symptoms (β = − .14), and positively with vitality (β = .13). Walking was negatively related to physical fatigue (β = − .11) and depressive symptoms (β = − .12) and positively with vitality (β = .15). Exercise was negatively associated with physical (β = − .19) and general (β = − .12) fatigue and depressive symptoms (β = − .09). ST was positively associated with physical fatigue (β = .19). Moderation analyses showed that LPA was related to lower mental fatigue and better vitality in people not self-isolating, and walking with lower physical fatigue in people self-isolating. These findings show the importance of encouraging PA for people with RA during a lockdown period for mental health and wellbeing.

Keywords: Physical activity · Sedentary behaviour · Mental health · COVID-19 · Self-isolation · Rheumatoid arthritis

Introduction

SARS-CoV-2 (COVID-19) has been declared a pandemic by the World Health Organisation (WHO) (WHO, 2020a). Unprecedented nationwide restrictions were put in place to limit the spread of the virus. In the United Kingdom (UK), the general population was instructed to only leave their home for basic necessities (i.e., food shopping, medical treatment), essential work that could not be carried out at home, and once a day for exercise. People considered at increased risk of serious complications following infection were advised to "self-isolate" (i.e., limit outside contact and not leave their homes). One such at risk population is Rheumatoid Arthritis (RA), which is an autoimmune disease characterised by inflammation, pain, fatigue, and poor mental health, and commonly managed by immunosuppressive therapies (Katz, 2017b; Matcham et al., 2013; Smolen et al., 2016; Uhlig et al., 2014). Whilst restrictions were deemed necessary to contain the spread of the virus, they can negatively impact mental health and psychological wellbeing (Brooks et al., 2020), as well as behaviours which can support mental health and wellbeing such as physical activity (PA) and sedentary behaviour (Veldhuijzen Van Zanten et al., 2020).

High levels of anxiety, depression and stress have been reported in the general population during COVID-19 restrictions (Harper et al., 2020; Huang & Zhao, 2020; Lai et al., 2020; Rodriguez-Rey et al., 2020). In people living with rheumatic diseases, a United States study reported difficulty managing negative emotions, perceived increased risk of being infected, and reduced access to healthcare and medications during this pandemic (Michaud et al., 2020). Such COVID-19-related concerns have been associated with poor mental health and compromised psychological wellbeing in the general population (Harper et al., 2020; Rodriguez-Rey et al., 2020). As the mental health impact of COVID-19 is even greater in those self-isolating (Brooks et al., 2020; Meyer et al., 2020), people living with RA, a population already at risk of compromised mental health (Fiest et al., 2017), may be at even greater risk of adverse psychological impact (Veldhuijzen Van Zanten et al., 2020). Indeed, negative

consequences of COVID-19 on mental health has been reported in 73% participants with rheumatic diseases (Ziade et al., 2020). Therefore, identifying factors that could positively impact mental health during the COVID-19 pandemic is critical for establishing effective management to attenuate the negative impact of this pandemic on wellbeing.

In the general population, PA is positively associated with indicators of mental health and psychological wellbeing and reductions in PA are associated with negative mental health during COVID-19 (Meyer et al., 2020). In RA, PA is related to reduced anxiety, depression, fatigue, pain and increased vitality (Chekroud et al., 2018; Kelley et al., 2015; Rongen-van Dartel et al., 2015; Rouse et al., 2015; Stenstrom, 1994). However, COVID-19 lockdown has restricted opportunities for PA, and not surprisingly, lower PA and increased screen time have been reported (Meyer et al., 2020; Pépin et al., 2020). This may pose a significant risk for the mental health and psychological wellbeing of those who are already at risk of low levels of PA and high levels of sedentary behaviour (i.e., waking activities involving sitting/lying and energy expenditure ≤1.5 metabolic equivalents), such as people with RA (Fenton & Kitas, 2016). Indeed, during COVID-19, people with rheumatic diseases reported challenges to being active (Michaud et al., 2020), and reductions in PA and increased sedentary time (ST) have been reported in RA (Pinto et al., 2020b), which could further worsen mental health and wellbeing (Fenton et al., 2018a).

Emerging evidence during COVID-19 emphasises the benefits of PA for mental health in the general population (Lesser & Nienhuis, 2020; Rodriguez-Rey et al., 2020; Zhang et al., 2020). However, associations between PA, mental health, and psychological wellbeing during COVID-19 have not been investigated in RA. The aims of this study were to 1) explore the associations between PA and ST with indicators of mental health and wellbeing in RA during COVID-19; and 2) examine the impact of self-isolation on these associations.

Methods

Participants

People with RA were recruited through social media of the National Rheumatoid Arthritis Society (NRAS) in the UK. Inclusion criteria were a self-reported clinical diagnosis of RA and aged ≥ 18 years. Ethical approval was obtained from the University of Birmingham ethics committee (ERN_20-0475). Participants were given an information sheet, provided informed consent, and completed the online questionnaire between April 8 − April 30 2020, when the most stringent lockdown restrictions applied in the UK.

Patient and Public Involvement

This online questionnaire was developed in collaboration with NRAS.

Measures

Physical Activity and Sedentary Behaviour

PA was assessed using the National Institutes of Health-American Association of Retired Persons (NIH-AARP) Diet and Health Study questionnaire to record participation in different types of PA (Gierach et al., 2009). Participants were asked to indicate how much time they had spent during the last 7 days in 3 PA types: 1) light intensity PA (LPA) (e.g., cooking, laundry), 2) walking, and 3) exercise (e.g., tennis, cycling). Items were scored using a categorical scale with response options from "none" to "≥ 10 hours". Sedentary behaviour was assessed using the International Physical Activity Questionnaire-Short Form (IPAQ-SF), which comprises 2 questions. Participants are asked "during the last 7 days, how much time did you spend sitting on a"…. 1) weekday and 2) weekend day, to calculate overall weekly ST. Both questionnaires are reliable and valid measures of PA and ST in older adults (Cleland et al., 2018; Gierach et al., 2009).

Pain

Pain experienced during the past week was assessed with the McGill Pain Questionnaire (MPQ), comprising 15 pain descriptors (e.g., "throbbing", "tender") on a scale from 0 (none) to 3 (severe) (Melzack, 1987). Participants were also asked to rate last week's pain on a visual analogue scale (VAS), from 0 (no pain) to 10 (worst imaginable pain) (Scott & Huskisson, 1976). The MPQ demonstrated good internal reliability in this study (Cronbach α=.92), and the MPQ and VAS have been validated in RA (Van Lankveld et al., 1992).

Fatigue

Fatigue during the past week was assessed using the Multidimensional Fatigue Inventory (MFI) reflecting physical, mental and general fatigue (4 items each), rated on a scale from 1 (no, that's not true) to 5 (yes, that's true) (Smets et al., 1995). MFI is a validated fatigue measure used in RA, with good internal reliability in this study (physical: α=.78, mental: α=.86, general: α=.73) (Rupp et al., 2004).

Anxious and Depressive Symptoms

Anxious (7 items) and depressive (7 items) symptoms during the past week were measured using the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983). Participants were asked to indicate their agreement (on a scale from 0 to 3) with each statement (e.g., "I feel tense or 'wound-up'"). The HADS has previously shown good validity (Covic et al., 2009), and good internal reliability (anxious symptoms: α=.87, depressive symptoms: α=.83) in this study.

Subjective Vitality

Vitality, a measure of positive wellbeing, experienced during the past week was measured using the Subjective Vitality Scale (SVS) (Ryan & Frederick, 1997). Participants were asked to rate 6 statements (e.g., "I've been feeling energised") on a scale from 1 (not at all true) to 7 (very true). The SVS demonstrates good reliability in this study (α=.88), and has been validated in RA (Rouse et al., 2015).

Functional Disability

Participants' functional disability was determined using the Stanford Health Assessment Questionnaire disability index (HAQ-DI), comprising 8 subscales each reflecting an activity of daily living (i.e., dressing, rising, eating, walking, hygiene, grip, reach and activities) (Kirwan & Reeback, 1986). Participants were asked to rate their ability to perform specific activities (e.g., "open car doors") on a scale from "without any difficulty" to "unable to do", and indicate if they used aids/devices for each activity. An overall disability index score is calculated as an average of the 8 subscales. A higher score represents higher functional disability.

General COVID-19 Concern

Concerns regarding COVID-19 were measured as the extent to which participants were apprehensive about 1) testing positive for COVID-19, 2) a family member testing positive for COVID-19, and 3) not being able to receive arthritis-related medical care. Each item was scored on a scale from 1 (not concerned at all) to 5 (very concerned), and the average was calculated.

COVID-19 Living

COVID-19 Living reflected living circumstances during COVID-19, which included "self-isolating at home" (i.e., not leaving the house due to medical recommendation (i.e., shielding) or personal concern) and "not self-isolating" (i.e., leaving the house for basic necessities, exercise, and/or work).

Data Reduction and Statistical Analysis

Data were analysed using IBM Statistical Package for the Social Sciences (SPSS) Version 26 and checked for normality using the Kolmogorov-Smirnov test. Independent variables were PA types (LPA, walking and exercise) and ST. Dependent variables were indicators of mental health and psychological wellbeing (i.e., pain (MPQ and VAS), physical, mental and general fatigue, anxious symptoms, depressive symptoms and vitality). Covariates were age, gender, functional disability,

living alone/with others, education and general COVID-19 concern, all with known associations with dependent variables in RA. The moderator variable was COVID-19 living.

N=408 participants provided complete data for PA and ST. Participants were excluded due to implausible ST (>18hours/day (Loppenthin & Esbensen, 2015), n=26), missing covariate (n=8) or moderator data (n=6). For the dependent variables, missing data were imputed using the expectation maximisation method where participants were missing one item of a questionnaire (MPQ: n=24; MFI: n=15; HADS: n=1; SVS: n=1). Participants with >1 missing value per questionnaire were excluded (n=23). The final sample size for all statistical analysis was n=345 participants.

Differences between those self-isolating vs not self-isolating were assessed with Mann-Whitney or Chi-square tests, as appropriate. To address the primary aim, hierarchical linear regression analyses were conducted to examine associations between PA and ST with indicators of mental health and psychological wellbeing, while adjusting for potential covariates. In these hierarchical regression analyses, we explored the following sequential models:

Regression Model 1 examined the associations between the covariates (age, gender, functional disability, living alone/with others, education and general COVID-19 concern) and the indicators of mental health and psychological wellbeing as dependent variables (pain, physical, mental and general fatigue, anxious symptoms, depressive symptoms and vitality). Separate regression analyses were conducted for each indicator of mental health and wellbeing. For each regression analysis, the F-value and p-value are reported to reflect statistical significance, the R^2-value is presented to reflect the variance in the indicator of mental health and psychological wellbeing explained by all covariates combined, and standardised beta-values (β-values) are presented to reflect the direction and strength of the association between each covariate and indicator of mental health and psychological wellbeing.

Regression Model 2 explored the associations between LPA, walking, exercise or ST with each indicator of mental health and psychological wellbeing indicator, while adjusting for the covariates included in Model 1. In other words, Regression Model 2 expanded the analyses conducted in Model 1 (with only covariates as the predictors of mental health and psychological wellbeing) to include LPA, walking, exercise or ST (independent variable) as a predictor of the indicators of mental health and psychological wellbeing (dependent variable). Separate analyses were conducted for each combination of independent and dependent variable, with the covariates included in all regression models. For each regression analysis, ΔR^2 was calculated to reflect the additional variance in the dependent variable explained by including the independent variable to the model with covariates only. F- and p-values are reported to reflect statistical significance of adding the independent variable to the model, and β-values reflect the direction and strength of the association between each independent (LPA, walking, exercise or ST) and dependent variable (indicator of mental health and psychological wellbeing).

Regression Model 3 explored whether the associations between the activity-related independent variables (i.e., LPA, walking, or exercise) with the indicators of mental health and psychological wellbeing were independent of ST, and vice versa, whether the associations between ST and indicators of mental health and psychological wellbeing were independent of the levels of activity (LPA, walking and exercise). More specifically, where regression Model 2 revealed significant associations between LPA, walking or exercise with a specific dependent variable, Model 3 included both the significant PA type as well as ST as predictors for that dependent variable, while also adjusting for covariates. For each regression analysis, ΔR^2 was calculated to reflect the additional variance explained in the dependent variable compared to Model 2. F- and p-values are reported to reflect statistical significance of adding the independent variable to the model, and β-values reflected the direction and strength of the association between each independent (LPA, walking, exercise or ST) and dependent variable (indicator of mental health and psychological wellbeing).

Finally, to explore the impact of COVID-19 living situation (i.e., self-isolation vs not self-isolating) on all associations between independent (LPA, walking, exercise, ST) and dependent variables (indicators of mental health and wellbeing), moderation analyses were conducted using the PROCESS model (Hayes, 2012). In all moderation analyses, age, gender, living situation, education, general COVID-19 concern and functional disability were included as covariates.

As the majority of variables were not normally distributed, bootstrapping was employed in regression analyses. Bootstrapping is a nonparametric re-sampling procedure that does not impose the assumption of normal distribution on the data (Preacher & Hayes, 2008). Significance was interpreted based on bootstrap-generated 95% bias-corrected confidence intervals (CIs) (5000 samples). CIs also provide more information than p values, showing the possible variability of effect size, and therefore are more appropriate for determining significance of bootstrapped data (du Prel et al., 2009; Preacher & Hayes, 2008), and standardised beta coefficients (β) were used to interpret the strength of associations.

Results

Participant characteristics are reported in **Table 6.1**. The sample predominantly comprised white females, with moderate functional disability. Mann-Whitney and Chi-Square tests revealed a longer disease duration, lower levels of LPA and walking, more pain, physical fatigue, functional disability and COVID-19 concern in self-isolating participants (all p's < .05, see **Table 6.1**).

Regression Analyses

Model 1: The associations between covariates with each mental health and psychological wellbeing indicator are summarised in **Table 6.2**. As is evident from **Table 6.2**, all regression models were statistically significant. Examination of the β-values of the covariates showed that functional disability was most strongly and consistently associated with all indicators of mental health and

psychological wellbeing, with higher functional disability related to more pain, fatigue and depressive and anxious symptoms, and lower vitality.

Model 2 expanded on Model 1 with the covariates only, by adding the independent variables (LPA, walking, exercise, ST) as predictors of the indicators of mental health and psychological wellbeing in separate analyses. **Table 6.3** presents the summary findings of Model 2 regression analyses, by focussing on the additional amount of variance explained by each independent variable and the associated beta-coefficient. More detailed information on these models, including all covariates, unstandardised beta-values (B) and the 95% CI (used to assess significance of the model), are reported in **Tables 6.4, 6.5, 6.6 and 6.7**. LPA was significantly negatively associated with mental fatigue and depressive symptoms and positively with vitality. Walking was negatively related to physical fatigue and depressive symptoms and positively with vitality. Exercise was negatively associated with physical and general fatigue, and depressive symptoms. ST was positively linked to physical fatigue. No other significant associations were detected. In all these analyses, the beta-coefficients for the covariates remained broadly similar to those reported in **Table 6.2**. More detailed information about the full regression models are presented in **Tables 6.4, 6.5, 6.6 and 6.7**.

Model 3 explored whether the associations reported between the activity-related independent variables with the dependent variables were independent of levels of ST and vice versa. When adding ST as an additional predictor, all significant associations between activity-related independent variables and the indicators of mental health and psychological wellbeing observed in Model 2 remained significant, with the exception of the association between walking with physical fatigue, which no longer remained significant when ST was added to the model (β= −.08, B= −.001, 95% CI= $-3.1\cdot10^{-3}$, $1.5\cdot10^{-4}$). In both walking and exercise models, ST was significantly associated with physical fatigue (walking model: β= .18, B< .001, 95% CI= $2.8\cdot10^{-4}$, $8.3\cdot10^{-4}$; exercise model: β= .18, B< .001, 95% CI= $2.9\cdot10^{-4}$, $8.2\cdot10^{-4}$).

Moderation Analyses

Moderation analyses revealed that COVID-19 living situation only moderated the associations between LPA with mental fatigue and vitality, and walking with physical fatigue. More LPA was significantly associated with lower mental fatigue and better vitality in those who were not self-isolating (mental fatigue model: β= −.26, 95% CI= −.42, −.09; vitality model: β= .31, 95% CI= .15, .47), but not in those who were self-isolating (mental fatigue model: β= −.04, 95% CI= −.16, .10; vitality model: β= .03, 95% CI= −.09, .16). Walking was associated with lower physical fatigue in people who were self-isolating (β= −.22, 95% CI= −.35, −.08), but not in those not self-isolating (β= −.02, 95% CI= −.15, .12). For more detailed information of moderation analysis results, please see **Table 6.8**.

Table 6.1: Descriptive statistics with p-values for total sample and sample stratified by COVID-19 Living status

	Self-isolating at home (n=230)	***Not self-isolating (n=115)***	***All participants (n=345)***	***p value***
Demographic Information				
Age (years)	51.53 ± 11.82	51.37 ± 11.58	51.48 ± 11.73	.805
Gender (n= female (%))	214 (93.0)	107 (93.0)	321 (93.0)	1.000
Ethnicity (n= white (%))	224 (97.4)	110 (96.5)	334 (96.8)	.640
RA duration (years from diagnosis)	11.51 ± 10.51	8.52 ± 8.08 [a]	10.52 ± 9.87	.012
Independent Variables				
LPA (minutes/week)	150 ± 420	300 ± 420 [a]	300 ± 420	.033
Walking (minutes/week)	60 ± 150	240 ± 420 [a]	90 ± 300	<.001
Exercise (minutes/week)	0 ± 30	0 ± 30	0 ± 30	.820
Sedentary Time (minutes/week)	3360 ± 1823	3360 ± 1980	3360 ± 1680	.088
Outcomes				
Pain (MPQ)	15.34 ± 9.30	13.12 ± 9.02 [a]	14.60 ± 9.26	.024
Pain (VAS rating)	4.57 ± 2.62	4.10 ± 2.58	4.41 ± 2.62	.095
Physical Fatigue	15.25 ± 4.06	14.06 ± 4.15 [a]	14.85 ± 4.12	.007
Mental Fatigue	12.04 ± 5.00	11.94 ± 5.01	12.01 ± 5.00	.860
General Fatigue	15.88 ± 3.63	15.31 ± 4.21	15.69 ± 3.83	.383
Anxious symptoms	9.39 ± 4.58	8.34 ± 4.51	9.04 ± 4.58	.055
Depressive symptoms	7.39 ± 3.94	6.86 ± 4.04	7.21 ± 3.98	.146
Subjective Vitality	2.56 ± 1.20	2.74 ± 1.32	2.62 ± 1.24	.266
Covariates				
Functional Disability (HAQ-DI)	1.33 ± 0.77	0.99 ± 0.69 [a]	1.22 ± 0.76	<.001
Living Situation (n= living alone (%))	43 (18.7)	17 (14.8)	60 (17.4)	.366
Education (n= higher education (%))	119 (51.7)	61 (53.0)	180 (52.2)	.819
General COVID-19 Concern	3.91 ± 0.90	3.62 ± 0.90 [a]	3.81 ± 0.91	.003

Note: [a]= significantly different from self-isolating at home with p<.05. Differences were examined using Mann-Whitney U and Chi-Square tests, as appropriate.

Values are reported as means ± SD, except for PA and SB variables which show Medians ± IQR.

Living situation was characterised as living with others (i.e., partner, family) or living alone. Education was characterised as higher education (university degree, doctorate) or secondary education (GCSE/O-level, A-level/GCE).

RA= Rheumatoid Arthritis, LPA= light non-exercise physical activity, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale, HAQ-DI= Health Assessment Questionnaire-Disability Index, SD= standard deviation, IQR= interquartile rage.

Table 6.2: Model 1 Regression Analyses for all covariates with each indicator of mental health and psychological wellbeing (dependent variable)

	Pain (MPQ)	***Pain (VAS rating)***	***Physical Fatigue***	***Mental Fatigue***	***General Fatigue***	***Anxious Symptoms***	***Depressive Symptoms***	***Vitality***
	R^2= .383	R^2= .330	R^2= .348	R^2= .140	R^2= .260	R^2= .207	R^2= .219	R^2= .176
	F= 34.91	F= 27.73	F= 30.09	F= 9.15	F= 19.78	F= 14.67	F= 15.77	F= 12.05
	p<.001	p<.001	p<.001	p<.001	p<.001	p<.001	p<.001	p<.001
	β	**β**	**β**	**β**	**β**	**β**	**β**	**β**
Age	-.11 [a]	-.08	-.11 [a]	-.27 [a]	-.21 [a]	-.25 [a]	-.16 [a]	.18 [a]
Gender	-.01	-.02	.05	-.06	-.03	-.01	.01	-.03
Education	-.03	.01	.05	-.04	.03	-.10 [a]	-.04	.06
Living Situation	.02	.04	.01	-.10 [a]	.01	.02	-.09	.07
Concern	.15 [a]	.08	.09	.10	.04	.32 [a]	.15 [a]	-.08
Functional Disability	.55 [a]	.55 [a]	.55 [a]	.23 [a]	.46 [a]	.14 [a]	.38 [a]	-.34 [a]

Note: Model 1: Regressions included all covariates (age, gender, education, living situation, general COVID-19 concern and functional disability) as predictors for each indicator of mental health and wellbeing in separate analyses.

R^2 represents the variance explained in the dependent variable by all covariates together. Statistical information about each model is presented by the F-value and the p-value, with β representing the standardised beta coefficient of each covariate.

[a]: Significant associations between covariates with indicators of mental health and wellbeing derived using bootstrapped 95% CI, Concern= General COVID-19 Concern, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale, CI= Confidence Interval.

Table 6.3: Summary Model 2 Regression Analyses for Light non-exercise Physical Activity, Walking, Exercise and Sedentary Time with dependent variables adjusting for covariates

	Pain (MPQ)		***Pain (VAS rating)***		***Physical Fatigue***		***Mental Fatigue***		***General Fatigue***		***Anxious Symptoms***		***Depressive Symptoms***		***Vitality***	
	β	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**	**β**	**ΔR²**
LPA	-.02	.000	-.06	.003	-.08	.006	-.11 [a]	.011 [b]	-.07	.005	-.04	.002	-.14 [a]	.018 [b]	.13 [a]	.016 [b]
Walking	.04	.002	.06	.003	-.11 [a]	.010 [b]	-.01	.000	-.06	.003	.02	.000	-.12 [a]	.012 [b]	.15 [a]	.020 [b]
Exercise	.04	.001	.01	.000	-.19 [a]	.033 [b]	-.04	.001	-.12 [a]	.013 [b]	-.04	.001	-.09 [a]	.008	.07	.005
Sedentary Time	-.08	.006	-.02	.000	.19 [a]	.032 [b]	.04	.001	.07	.005	-.04	.001	.08	.006	-.09	.008

Note: [a] = significant β (standardised beta) coefficients derived using bootstrapped 95% CI (see Tables 6.4, 6.5, 6.6 and 6.7); [b]= p<.05 for ΔR^2 values determining the significance of the overall model.

Model 2: expanded Model 1 (covariates age, gender, education, living situation, general COVID-19 concern and functional disability only) by adding LPA, walking, exercise and ST as individual predictors of the dependent variables.

β represents the standardised beta coefficient and ΔR^2 represents the proportion of the variance that is explained by the addition of the predictor independent variable to the model relative to Model 1. Beta coefficients for covariates were broadly similar as those reported in Table 6.1. Therefore, to improve readability, betas are only reported for the associations between the independent variables (LPA, walking, exercise, ST) and the dependent variables. Full information of all Model 2 regressions, including the 95% CI intervals to determine significance, F-statistics with associated p-values and the β for each covariate are reported in Tables 6.4, 6.5, 6.6 and 6.7.

LPA= light intensity physical activity, Concern= General COVID-19 Concern, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale, CI= Confidence Interval

Discussion

This is the first study to show associations between activity behaviours and indicators of mental health and psychological wellbeing in people with RA during COVID-19. LPA and walking were associated with lower physical and mental fatigue and depressive symptoms, and higher vitality. Exercise was related to lower physical and general fatigue and fewer depressive symptoms, and ST was related to higher physical fatigue. In addition, COVID-19 living situation moderated some associations between LPA and walking with physical and mental fatigue and vitality.

The finding that LPA and walking were associated with higher vitality is in line with epidemiological research demonstrating a relationship between LPA with wellbeing in older adults (Buman et al., 2010). My results point to the importance of LPA and walking for wellbeing in RA during COVID-19. From a behaviour change perspective, encouraging non-exercise LPA (e.g., household chores) and walking may be perceived as more feasible and accessible for RA, a population experiencing significant disease-related barriers to PA (Veldhuijzen Van Zanten et al., 2015).

PA was associated with lower depressive symptoms, aligned with previous arthritis research (Fenton et al., 2018b; Kelley et al., 2015). Associations were of a similar, though opposite, magnitude to associations between COVID-19 concerns and depressive symptoms. Fear or concern about the virus has been related to depression (Brooks et al., 2020), and my findings suggest that PA counteracts this negative impact on depressive symptoms in RA, in line with findings in college students and older adults (Carriedo et al., 2020; Zhang et al., 2020). Importantly, associations between PA with depressive symptoms and vitality were independent of functional disability in this study and others (Fenton et al., 2018b). Thus, activity at any intensity should be promoted in all people with RA to improve mental health and wellbeing, regardless of functional disability.

PA was associated with lower and ST with higher physical and mental fatigue, in line with interventions promoting PA and reducing ST improving fatigue in RA (Katz et al., 2018; Rongen-van Dartel et al., 2015; Thomsen et al., 2017). My findings emphasise the importance of the multidimensional aspects of fatigue in RA. Specifically, LPA negatively associated with mental fatigue, whereas walking, exercise, and ST were related to physical fatigue. My results suggest that different PA types could be related to different aspects of fatigue in people living with RA. Aligned with present findings, exercise interventions have been reported to be particularly effective for physical fatigue in RA (Rupp et al., 2004).

PA can lead to improved mental health and wellbeing during lockdown situations through several pathways. For example, PA can distract from negative thoughts and worries (Mikkelsen et al., 2017), being active can have an immediate positive effect on mood (Mikkelsen et al., 2017), outdoor environment can induce mental stimulation (Lesser & Nienhuis, 2020), and PA can provide structure when daily routine is disrupted due to lockdown, with the resulting sense of control improving wellbeing (Ryan & Deci, 2017). Thus, recommendations promoting PA may offer an avenue to support clinical populations to cope with the impact of COVID-19 on their mental health (Carriedo et al., 2020; Matias et al., 2020). Given that maintaining PA during COVID-19 is associated with better mental health in older adults (Carriedo et al., 2020), the present results imply that increasing PA may positively impact mental health and wellbeing in RA during COVID-19. Given the duration of the pandemic and recent return of restrictions, longitudinal studies during COVID-19 are needed to understanding how changes in PA contribute to better mental health and wellbeing in RA. This can inform guidance on management of wellbeing during these difficult times, not just in RA, but also other clinical populations.

PA and ST were not associated with anxious symptoms in this study. There are mixed findings related to anxiety and PA during COVID-19 (Antunes et al., 2020; Lesser & Nienhuis, 2020; Zhang et al.,

2020). The impact of PA on anxiety during COVID-19 may be influenced by prior activity levels. Indeed, research suggests inactive people who increased PA during COVID-19 reported lower anxiety compared to those who became less active, but these associations were not seen in people who were classed as active prior to COVID-19 (Lesser & Nienhuis, 2020). Thus, perhaps instead of looking at absolute values, it might be more important to examine changes in PA during a pandemic in relation to anxiety.

People with RA in this study did not report associations between PA and ST with pain, which is in contrast to previous observational studies of RA (Hakkinen et al., 2001). This may be due to the possible bi-directional association between PA with mental health and wellbeing (Lwin et al., 2020), and also the association between pain and functional disability (Luyster et al., 2011), which may have affected levels of PA observed depending on COVID-19 living status. Specifically, correlational analysis revealed walking was negatively related to pain (data not reported), but this association was no longer significant when adjusting for functional disability. Given my findings suggested that people with RA who were self-isolating had higher levels of pain and functional disability, and lower levels of LPA and walking, compared to those leaving the house, it could be assumed that individuals with the least pain and disability were more likely to leave the house for PA. This lower variability in pain among those who were accruing some form of PA (through leaving the house), could mean associations between PA and pain are less likely to be observed.

The lower levels of PA observed among people with RA who were self-isolating is in agreement with previous studies (Meyer et al., 2020). Moderation analysis showed LPA was only related to mental fatigue and vitality in those not self-isolating, whereas walking was only associated with physical fatigue in those self-isolating. As those self-isolating did significantly less walking than those not self-isolating, this could suggest that walking specifically, should be encouraged among individuals self-

isolating to reduce physical fatigue. However, as few significant moderation effects of COVID-19 living were observed, these findings should be interpreted with care.

Except for physical fatigue, ST was not associated with indicators of mental health or wellbeing, contrasting previous COVID-19 research in the general population (Meyer et al., 2020). My measure of ST reflected total ST, and did not differentiate between different sedentary behaviours; e.g., sitting while being intellectually stimulated, e.g., during work, is suggested to have less negative impact health and wellbeing compared to sitting watching TV or using electronic devices (Saidj et al., 2014). Consequently, future studies are required to understand the specific role of different sedentary behaviours for mental health and wellbeing in RA, both during and beyond the pandemic.

The current study included a large sample of people with RA during stringent lockdown conditions. Therefore, it was only possible to collect self-report data for all outcomes, which should be acknowledged as a limitation. In addition, the associations reported are cross-sectional, so causality cannot be inferred. Therefore, exploring these associations over time during COVID-19 will help to better understand the implications of this pandemic on the link between activity behaviours and mental health and wellbeing, and how PA can support mental health and wellbeing throughout the pandemic.

In summary, PA, specifically LPA and walking, was positively associated with mental health and psychological wellbeing in RA during COVID-19. These findings support recommendations from different governments to encourage PA during lockdown restrictions, to attenuate the negative impact of a pandemic on mental health and wellbeing. Given the known barriers for PA in RA (Veldhuijzen Van Zanten et al., 2015), and the reported additional barriers experienced during COVID-19 (Michaud et al., 2020), these findings emphasise the importance of appropriate support and recommendations for PA in people with RA, and potentially other clinical populations,

particularly in those self-isolating, during a pandemic to maintain mental health and psychological wellbeing.

Declarations

AUTHOR CONTRIBUTORSHIP: Conception and design of the study: Sophia Brady, Sally Fenton, George Metsios, Ailsa Bosworth, George Kitas, Jet Veldhuijzen van Zanten. Data acquisition: Sophia Brady, Sally Fenton, Ailsa Bosworth, Jet Veldhuijzen van Zanten. Data analysis: Sophia Brady, Sally Fenton, Jet Veldhuijzen van Zanten. Data interpretations and drafting of manuscript: all authors. Final approval of manuscript: all authors.

FUNDING: This work was completed as part of a PhD studentship supported by the Medical Research Council (MRC)-Versus Arthritis Centre for Musculoskeletal Ageing Research (CMAR) (grant number: MR/K00414X/1).

COMPLIANCE WITH ETHICAL STANDARDS: CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

ETHICS APPROVAL: This study was performed in line with the principles of the Declaration of Helsinki. Ethical approval was obtained from the University of Birmingham ethics committee (ERN_20-0475).

CONSENT TO PARTICIPATE AND CONSENT TO PUBLISH: Informed consent was obtained from all individual participants included in the study. The authors affirm that human research participants provided informed consent for publication.

ACKNOWLEDGEMENTS: The authors would like to thank the National Rheumatoid Arthritis Society (NRAS) for their support in the development of this study and the recruitment for this study. In addition, the authors would like to thank all participants for their time to complete to questionnaire pack.

Table 6.4: Model 2 Regression Analyses for Light non-exercise PA (min/week) with indicators of mental health and psychological wellbeing

	Pain (MPQ)			***Pain (VAS rating)***			***Physical Fatigue***			***Mental Fatigue***			***General Fatigue***			***Anxious Symptoms***			***Depressive Symptoms***			***Vitality***		
	R^2= .383			R^2= .333			R^2= .354			R^2= .151			R^2= .264			R^2= .208			R^2= .237			R^2= .192		
	ΔR^2= .000			ΔR^2= .003			ΔR^2= .006			ΔR^2= .011			ΔR^2= .005			ΔR^2= .002			ΔR^2= .018			ΔR^2= .016		
	F= 0.23, p=.63			F= 1.66, p=.20			F= 2.88, p=.09			F= 4.40, p=.04			F= 2.14, p=.15			F= 0.77, p=.38			F= 8.03, p=.01			F= 6.69, p=.01		
	β	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**
Age	-.11[a]	-0.08[a]	-0.15, -0.01	-.08	-0.02	-0.04, 0.00	-.10[a]	-0.03[a]	-0.07, 0.00	-.25[a]	-0.11[a]	-0.15, -0.07	-.21[a]	-0.07[a]	-0.10, -0.03	-.24[a]	-0.09[a]	-0.13, -0.05	-.14[a]	-0.05[a]	-0.08, -0.02	.16[a]	0.02[a]	0.01, 0.03
Gender	-.01	-0.26	-3.29, 2.69	-.02	-0.19	-1.19, 0.75	.06	0.97	-0.50, 2.42	-.05	-1.04	-2.85, 0.67	-.03	-0.42	-1.73, 0.95	-.01	-0.15	-1.65, 1.35	.03	0.39	-1.17, 1.89	-.05	-0.23	-0.65, 0.15
Education	-.03	-0.58	-2.19, 1.05	.01	0.08	-0.40, 0.54	.06	0.51	-0.23, 1.26	-.03	-0.32	-1.33, 0.71	.04	0.29	-0.43, 1.02	-.09	-0.86	-1.77, 0.04	-.03	-0.23	-0.99, 0.52	.05	0.12	-0.13, 0.37
Living Situation	.03	0.60	-1.59, 2.73	.05	0.37	-0.23, 0.94	.03	0.30	-0.56, 1.19	-.08	-1.12	-2.46, 0.19	.02	0.23	-0.65, 1.14	.03	0.32	-0.84, 1.42	-.07	-0.73	-1.71, 0.25	.04	0.14	-0.19, 0.45
Concern	.15[a]	1.57[a]	0.74, 2.39	.09	0.25	-0.02, 0.50	.09	0.40	-0.02, 0.80	.10	0.54	-0.04, 1.09	.04	0.16	-0.27, 0.57	.32[a]	1.60[a]	1.11, 2.07	.15[a]	0.66[a]	0.20, 1.07	-.08	-0.11	-0.25, 0.03
Functional Disability	.55[a]	6.74[a]	5.67, 7.81	.54[a]	1.86[a]	1.56, 2.15	.54[a]	2.95[a]	2.46, 3.44	.22[a]	1.43[a]	0.77, 2.09	.46[a]	2.31[a]	1.82, 2.82	.13[a]	0.79[a]	0.18, 1.40	.36[a]	1.90[a]	1.34, 2.45	-.33[a]	-0.54[a]	-0.71, -0.38
LPA	-.02	-0.00	$-5.2\cdot10^{-3}$, $3.2\cdot10^{-3}$	-.06	-0.00	$-1.9\cdot10^{-3}$, $3.8\cdot10^{-4}$	-.08	-0.00	$-3.4\cdot10^{-3}$, $2.8\cdot10^{-4}$	-.11[a]	-0.00[a]	$-5.2\cdot10^{-3}$, $-1.5\cdot10^{-4}$	-.07	-0.00	$-3.2\cdot10^{-3}$, $4.1\cdot10^{-4}$	-.04	-0.00	$-3.3\cdot10^{-3}$, $1.4\cdot10^{-3}$	-.14[a]	-0.00[a]	$-4.8\cdot10^{-3}$, $-6.4\cdot10^{-4}$	.13[a]	0.00[a]	$1.9\cdot10^{-4}$, $1.4\cdot10^{-3}$

Note: Significant associations (=[a]) between LPA with indicators of mental health and well-being were interpreted using bootstrapped 95% CI. Bootstrapping was used to compute 95% CI which produces an unstandardised B-coefficient and corresponding unstandardised 95% CI. The standardised beta-value (β) is reported (and in the main analyses) to allow to facilitate the interpretation regarding the strength and direction of all associations reported.

Model 2: expanded Model 1 (covariates age, gender, education, living situation, general COVID-19 concern and functional disability only) by adding LPA as a predictor of the separate indicators of mental health and psychological wellbeing.

R^2 represents the variance explained on the dependent variable (pain, fatigue, anxious and depressive symptoms and vitality) by LPA (independent variable) and all covariates together. ΔR^2 indicates proportion of the variance that is explained by the addition of the LPA to the Model 1 (covariates only). The F value and p statistic represent the f-statistic related to the ΔR^2, indicating the significance of the model when adding the additional predictor (i.e., significance of the ΔR^2). β represents the standardised beta coefficient.

B= unstandardised beta coefficient, 95% CI= 95% Confidence Interval [lower limit, upper limit], LPA= light intensity physical activity, Concern= General COVID-19 Concern, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale.

The associations between covariates and indicators of mental health and well-being were unchanged from Model 1 ($\Delta\beta<0.05$). There were minor significance changes for some associations that were borderline significant in Model 1: between age and physical fatigue, education and anxious symptoms and living situation and mental fatigue.

Table 6.5: Model 2 Regression Analyses for Walking PA (min/week) with dependent variables

	Pain (MPQ)			***Pain (VAS rating)***			***Physical Fatigue***			***Mental Fatigue***			***General Fatigue***			***Anxious Symptoms***			***Depressive Symptoms***			***Vitality***		
	R^2= .384			R^2= .333			R^2= .358			R^2= .140			R^2= .263			R^2= .207			R^2= .231			R^2= .196		
	ΔR^2= .002			ΔR^2= .003			ΔR^2= .010			ΔR^2= .000			ΔR^2= .003			ΔR^2= .000			ΔR^2= .012			ΔR^2= .020		
	F= 0.85, p=.36			F= 1.59, p=.21			F= 5.26, p=.02			F= 0.02, p=.89			F= 1.40, p=.24			F= 0.19, p=.66			F= 5.36, p=.02			F= 8.33, p=.004		
	β	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**
Age	-.11[a]	-0.09[a]	-0.16, -0.02	-.09	-0.02	-0.04, 0.00	-.10[a]	-0.03[a]	-0.07, 0.00	-.27[a]	-0.11[a]	-0.15, -0.07	-.21[a]	-0.07[a]	-0.10, -0.04	-.25[a]	-0.10[a]	-0.13, -0.06	-.15[a]	-0.05[a]	-0.08, -0.02	.17[a]	0.02[a]	0.01, 0.03
Gender	-.01	-0.29	-3.29, 2.59	-.02	-0.23	-1.19, 0.71	.05	0.81	-0.70, 2.31	-.06	-1.24	-3.08, 0.50	-.04	-0.54	-1.85, 0.85	-.01	-0.22	-1.81, 1.29	.01	0.15	-1.48, 1.70	-.03	-0.15	-0.57, 0.23
Education	-.03	-0.62	-2.26, 1.02	.01	0.05	-0.41, 0.51	.05	0.43	-0.29, 1.17	-.04	-0.43	-1.41, 0.56	.03	0.23	-0.50, 0.94	-.10[a]	-0.90[a]	-1.77, -0.01	-.04	-0.36	-1.09, 0.40	.06	0.16	-0.09, 0.41
Living Situation	.02	0.48	-1.58, 2.63	.04	0.28	-0.32, 0.90	.02	0.19	-0.66, 1.03	-.10[a]	-1.37[a]	-2.65, -0.12	.01	0.12	-0.71, 1.01	.02	0.22	-0.87, 1.27	-.09	-0.95	-1.94, 0.01	.06	0.20	-0.12, 0.52
Concern	.15[a]	1.58[a]	0.80, 2.39	.09	0.25	-0.02, 0.51	.08	0.37	-0.04, 0.77	.10	0.53	-0.05, 1.10	.03	0.15	-0.28, 0.55	.32[a]	1.60[a]	1.12, 2.06	.14[a]	0.62[a]	0.20, 1.05	-.07	-0.10	-0.23, 0.04
Functional Disability	.57[a]	6.90[a]	5.83, 8.00	.56[a]	1.94[a]	1.64, 2.23	.52[a]	2.83[a]	2.31, 3.35	.23[a]	1.49[a]	0.81, 2.21	.45[a]	2.26[a]	1.75, 2.78	.14[a]	0.86[a]	0.24, 1.48	.34[a]	1.80[a]	1.28, 2.34	-.30[a]	-0.49[a]	-0.67, -0.33
Walking	.04	0.00	$-2.3\cdot10^{-3}$, $5.2\cdot10^{-3}$	.06	0.00	$-3.7\cdot10^{-4}$, $1.7\cdot10^{-3}$	-.11[a]	-0.00[a]	$-3.7\cdot10^{-3}$, $-2.9\cdot10^{-4}$	-.01	0.00	$-2.7\cdot10^{-3}$, $2.4\cdot10^{-3}$	-.06	-0.00	$-2.9\cdot10^{-3}$, $7.1\cdot10^{-4}$	.02	0.00	$-1.6\cdot10^{-3}$, $2.6\cdot10^{-3}$	-.12[a]	-0.00[a]	$-3.8\cdot10^{-3}$, $-2.6\cdot10^{-4}$	.15[a]	0.00[a]	$2.5\cdot10^{-4}$, $1.4\cdot10^{-3}$

Note: Significant associations (=[a]) between Walking with indicators of mental health and well-being were interpreted using bootstrapped 95% CI. Bootstrapping was used to compute 95% CI which produces an unstandardised B-coefficient and corresponding unstandardised 95% CI. The standardised beta-value (β) is reported (and in the main analyses) to allow to facilitate the interpretation regarding the strength and direction of all associations reported.

Model 2: expanded Model 1 (covariates age, gender, education, living situation, general COVID-19 concern and functional disability only) by adding Walking as a predictor of the separate indicators of mental health and psychological wellbeing.

R^2 represents the variance explained on the dependent variable (pain, fatigue, anxious and depressive symptoms and vitality) by Walking (independent variable) and all covariates together. ΔR^2 indicates proportion of the variance that is explained by the addition of the Walking to the Model 1 (covariates only). The F value and p statistic represent the f-statistic related to the ΔR^2, indicating the significance of the model when adding the additional predictor (i.e., significance of the ΔR^2). β represents the standardised beta coefficient.

B= unstandardised beta coefficient, 95% CI= 95% Confidence Interval [lower limit, upper limit], Concern= General COVID-19 Concern, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale.

The associations between covariates and indicators of mental health and well-being were unchanged from Model 1 ($\Delta\beta<0.05$). There were minor significance changes for some associations that were borderline significant in Model 1: between age and physical fatigue, education and anxious symptoms and living situation and mental fatigue.

Table 6.6: Model 2 Regression Analyses for Exercise PA (min/week) with dependent variables

	Pain (MPQ)			***Pain (VAS rating)***			***Physical Fatigue***			***Mental Fatigue***			***General Fatigue***			***Anxious Symptoms***			***Depressive Symptoms***			***Vitality***		
	R^2= .384			R^2= .330			R^2= .381			R^2= .141			R^2= .273			R^2= .208			R^2= .227			R^2= .181		
	ΔR^2= .001			ΔR^2= .000			ΔR^2= .033			ΔR^2= .001			ΔR^2= .013			ΔR^2= .001			ΔR^2= .008,			ΔR^2= .005		
	F= 0.65, p=.42			F= 0.04, p=.84			F= 18.00, p<.001			F= 0.48, p=.49			F= 6.14, p=.01			F= 0.62, p=.43			F= 3.55, p=.06			F= 2.11, p=.15		
	β	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**
Age	-.11 [a]	-0.08 [a]	-0.15, -0.01	-.08	-0.02	-0.04, 0.00	-.12 [a]	-0.04 [a]	-0.07, -0.01	-.27 [a]	-0.11 [a]	-0.15, -0.07	-.22 [a]	-0.07 [a]	-0.11, -0.04	-.25 [a]	-0.10 [a]	-0.14, -0.06	-.16 [a]	-0.05 [a]	-0.09, -0.02	.18 [a]	0.02 [a]	0.01, 0.03
Gender	-.01	-0.27	-3.25, 2.65	-.02	-0.25	-1.16, 0.70	.04	0.73	-0.74, 2.29	-.06	-1.27	-3.12, 0.49	-.04	-0.60	-1.92, 0.83	-.01	-0.26	-1.78, 1.26	.01	0.13	-1.52, 1.74	-.03	-0.15	-0.58, 0.22
Education	-.04	-0.69	-2.35, 0.90	.01	0.04	-0.43, 0.49	.07	0.58	-0.16, 1.30	-.04	-0.40	-1.42, 0.63	.04	0.31	-0.40, 1.03	-.10	-0.87	-1.76, 0.00	-.04	-0.28	-1.05, 0.47	.06	0.14	-0.11, 0.39
Living Situation	.02	0.51	-1.54, 2.67	.04	0.30	-0.30, 0.89	.01	0.15	-0.72, 1.02	-.10 [a]	-1.38 [a]	-2.64, -0.09	.01	0.10	-0.77, 1.00	.02	0.23	-0.92, 1.31	-.09	-0.99	-1.96, -0.06	.07	0.22	-0.11, 0.52
Concern	.15 [a]	1.57 [a]	0.78, 2.39	.08	0.24	-0.03, 0.49	.08	0.38	-0.03, 0.77	.10	0.53	-0.05, 1.10	.04	0.15	-0.28, 0.56	.31 [a]	1.59 [a]	1.11, 2.06	.15 [a]	0.64 [a]	0.20, 1.06	-.08	-0.11	-0.24, 0.04
Functional Disability	.56 [a]	6.85 [a]	5.73, 7.94	.55 [a]	1.89 [a]	1.59. 2.19	.51 [a]	2.79 [a]	2.30, 3.28	.22 [a]	1.46 [a]	0.79, 2.18	.44 [a]	2.23 [a]	1.74, 2.72	.13 [a]	0.77 [a]	0.13, 1.39	.36 [a]	1.87 [a]	1.34, 2.41	-.33 [a]	-0.54 [a]	-0.71, -0.38
Exercise	.04	0.00	$-3.6\cdot10^{-3}$, $1.0\cdot10^{-2}$	.01	0.00	$-1.6\cdot10^{-3}$, $2.0\cdot10^{-3}$	-.19 [a]	-0.01 [a]	$-9.7\cdot10^{-3}$, $-2.9\cdot10^{-3}$	-.04	-0.00	$-6.0\cdot10^{-3}$, $4.0\cdot10^{-3}$	-.12 [a]	-0.00 [a]	$-7.0\cdot10^{-3}$, $-1.8\cdot10^{-4}$	-.04	-0.00	$-5.5\cdot10^{-3}$, $2.3\cdot10^{-3}$	-.09 [a]	-0.00 [a]	$-5.5\cdot10^{-3}$, $-2.9\cdot10^{-4}$	.07	0.00	$-3.1\cdot10^{-4}$, $1.9\cdot10^{-3}$

Note: Significant associations (= [a]) between Exercise with indicators of mental health and well-being were interpreted using bootstrapped 95% CI. Bootstrapping was used to compute 95% CI which produces an unstandardised B-coefficient and corresponding unstandardised 95% CI. The standardised beta-value (β) is reported (and in the main analyses) to allow to facilitate the interpretation regarding the strength and direction of all associations reported.

Model 2: expanded Model 1 (covariates age, gender, education, living situation, general COVID-19 concern and functional disability only) by adding Exercise as a predictor of the separate indicators of mental health and psychological wellbeing.

R^2 represents the variance explained on the dependent variable (pain, fatigue, anxious and depressive symptoms and vitality) by Exercise (independent variable) and all covariates together. ΔR^2 indicates proportion of the variance that is explained by the addition of the Exercise to the Model 1 (covariates only). The F value and p statistic represent the f-statistic related to the ΔR^2, indicating the significance of the model when adding the additional predictor (i.e., significance of the ΔR^2). β represents the standardised beta coefficient.

B= unstandardised beta coefficient, 95% CI= 95% Confidence Interval [lower limit, upper limit], Concern= General COVID-19 Concern, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale.

The associations between covariates and indicators of mental health and well-being were unchanged from Model 1 ($\Delta\beta<0.05$). There were minor significance changes for some associations that were borderline significant in Model 1: between age and physical fatigue, education and anxious symptoms and living situation and mental fatigue.

Table 6.7: Model 2 Regression Analyses for Sedentary Time (min/week) with dependent variables

	Pain (MPQ)			***Pain (VAS rating)***			***Physical Fatigue***			***Mental Fatigue***			***General Fatigue***			***Anxious Symptoms***			***Depressive Symptoms***			***Vitality***		
	R^2= .389			R^2= .330			R^2= .380			R^2= .141			R^2= .265			R^2= .208			R^2= .225			R^2= .184		
	ΔR^2= .006			ΔR^2= .000			ΔR^2= .032			ΔR^2= .001			ΔR^2= .005			ΔR^2= .001			ΔR^2= .006			ΔR^2= .008		
	F= 3.43, p=.07			F= 0.12, p=.73			F= 17.42, p<.001			F= 0.51, p=.48			F= 2.20, p=.14			F= 0.58, p=.45			F= 2.74, p=.10			F= 3.12, p=.08		
	β	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**	**β**	**B**	**95% CI**
Age	-.13[a]	-0.10[a]	-0.17, -0.03	-.09	-0.02	-0.04, 0.00	-.07	-0.02	-0.06, 0.01	-.26[a]	-0.11[a]	-0.15, -0.07	-.20[a]	-0.07[a]	-0.10, -0.03	-.25[a]	-0.10[a]	-0.14, -0.06	-.14[a]	-0.05[a]	-0.08, -0.02	.16[a]	0.02[a]	0.01, 0.03
Gender	-.01	-0.40	-3.16, 2.40	-.02	-0.25	-1.22, 0.70	.06	0.93	-0.57, 2.48	-.06	-1.22	-3.11, 0.49	-.03	-0.49	-1.80, 0.90	-.01	-0.24	-1.82, 1.30	.01	0.23	-1.35, 1.79	-.04	-0.18	-0.59, 0.21
Education	-.03	-0.49	-2.07, 1.10	.01	0.05	-0.40, 0.53	.04	0.30	-0.41, 1.04	-.05	-0.46	-1.50, 0.56	.02	0.18	-0.55, 0.89	-.10	-0.87	-1.75, -0.01	-.05	-0.41	-1.16, 0.36	.07	0.18	-0.08, 0.43
Living Situation	.00	0.11	-2.00, 2.12	.04	0.28	-0.35, 0.88	.05	0.56	-0.27, 1.41	-.10	-1.27	-2.58, -0.04	.02	0.25	-0.65, 1.15	.01	0.14	-0.97, 1.25	-.08	-0.81	-1.75, 0.13	.05	0.16	-0.17, 0.46
Concern	.15[a]	1.57[a]	0.73, 2.36	.08	0.24	-0.02, 0.50	.09	0.39	-0.01, 0.78	.10	0.53	-0.05, 1.13	.04	0.16	-0.28, 0.57	.32[a]	1.60[a]	1.10, 2.07	.15[a]	0.64[a]	0.20, 1.07	-.08	-0.11	-0.24, 0.03
Functional Disability	.57[a]	6.97[a]	5.88, 8.04	.55[a]	1.89[a]	1.58, 2.19	.51[a]	2.79[a]	2.29, 3.27	.22[a]	1.46[a]	0.76, 2.14	.45[a]	2.27[a]	1.77, 2.79	.14[a]	0.87[a]	0.25, 1.47	.36[a]	1.88[a]	1.35, 2.42	-.33[a]	-0.53[a]	-0.70, -0.37
Sedentary Time	-.08	-0.00	$-1.2\cdot10^{-3}$, $5.9\cdot10^{-5}$	-.02	-0.00	$-2.2\cdot10^{-4}$, $1.6\cdot10^{-4}$	.19[a]	0.00[a]	$3.3\cdot10^{-4}$, $8.4\cdot10^{-4}$	.04	0.00	$-2.6\cdot10^{-4}$, $5.4\cdot10^{-4}$	.07	0.00	$-4.6\cdot10^{-5}$, $4.8\cdot10^{-4}$	-.04	0.00	$-5.0\cdot10^{-4}$, $2.2\cdot10^{-4}$	.08	0.00	$-4.3\cdot10^{-5}$, $5.4\cdot10^{-4}$	-.09	-0.00	$-1.8\cdot10^{-4}$, $1.1\cdot10^{-5}$

Note: Significant associations (=[a]) between Sedentary Time with indicators of mental health and well-being were interpreted using bootstrapped 95% CI. Bootstrapping was used to compute 95% CI which produces an unstandardised B-coefficient and corresponding unstandardised 95% CI. The standardised beta-value (β) is reported (and in the main analyses) to allow to facilitate the interpretation regarding the strength and direction of all associations reported.

Model 2: expanded Model 1 (covariates age, gender, education, living situation, general COVID-19 concern and functional disability only) by adding Sedentary Time as a predictor of the separate indicators of mental health and psychological wellbeing.

R^2 represents the variance explained on the dependent variable (pain, fatigue, anxious and depressive symptoms and vitality) by Sedentary Time (independent variable) and all covariates together. ΔR^2 indicates proportion of the variance that is explained by the addition of the Sedentary Time to the Model 1 (covariates only). The F value and p statistic represent the f-statistic related to the ΔR^2, indicating the significance of the model when adding the additional predictor (i.e., significance of the ΔR^2). β represents the standardised beta coefficient.

B= unstandardised beta coefficient, 95% CI= 95% Confidence Interval [lower limit, upper limit], Concern= General COVID-19 Concern, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale.

The associations between covariates and indicators of mental health and well-being were unchanged from Model 1 (Δβ<0.05). There were minor significance changes for some associations that were borderline significant in Model 1: between age and physical fatigue, education and anxious symptoms and living situation and mental fatigue.

Table 6.8: Moderation Analysis results to explore effects of COVID-19 living situation (self-isolating vs not self-isolating) on all associations between independent and dependent variables

	Pain (MPQ)			***Pain (VAS rating)***			***Physical Fatigue***			***Mental Fatigue***			***General Fatigue***			***Anxious Symptoms***			***Depressive Symptoms***			***Vitality***		
	β	95 % CI		β	95 % CI		β	95 % CI		β	95 % CI		β	95 % CI		β	95 % CI		β	95 % CI		β	95 % CI	
		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit
LPA	-.06	-0.24	0.12	-.03	-0.21	0.15	-.02	-0.20	0.17	-.22 [a]	-0.43	-0.01	-.15	-0.35	0.04	-.11	-0.32	0.09	-.10	-0.30	0.10	.28 [a]	0.08	0.48
Walking	-.03	-0.21	0.15	.01	-0.17	0.20	.20 [a]	0.01	0.39	.19	-0.02	0.41	.06	-0.14	0.26	.04	-0.17	0.25	.05	-0.15	0.26	-.02	-0.23	0.19
Exercise	-.04	-0.24	0.15	.15	-0.05	0.35	.01	-0.19	0.21	.05	-0.18	0.28	-.05	-0.26	0.16	-.08	-0.30	0.15	-.06	-0.28	0.16	-.01	-0.23	0.22
Sedentary Time	-.05	-0.24	0.15	-.12	-0.32	0.08	.04	-0.15	0.24	.21	-0.02	0.44	.05	-0.16	0.26	-.14	-0.36	0.08	.00	-0.22	0.22	-.09	-0.31	0.14

Note: Significant associations (= [a]) between LPA, walking, exercise and sedentary time (independent variables) with indicators of mental health and wellbeing (dependent variables) were derived using bootstrapped 95% CI

Moderation Analysis: COVID-19 living situation (self-isolating vs not self-isolating) was added as a moderator variable to Model 2 regressions (LPA, walking, exercise or ST, adjusted for all covariates), using the PROCESS model in SPSS, to explore if COVID-19 living situation moderated the associations between the independent and dependent variables, i.e., if self-isolation status affected associations between activity behaviour and indicator of mental health and psychological wellbeing.

Results are reported using Model 2 regressions (adjusted for covariates: age, gender, living situation, education, general COVID-19 concern and functional disability). β coefficients represent the degree of change in the outcome dependant variable for every 1 unit change of the independent variable. β coefficients and 95% CIs were calculated using z scores of all independent and dependant variables in order to standardise output values.

LPA= light non-exercise physical activity, MPQ= McGill Pain Questionnaire, VAS= Visual Analogue Scale, β= standardised beta coefficient, 95% CI= 95% confidence intervals.

CHAPTER 7: GENERAL DISCUSSION

Summary of Findings

The studies comprising this thesis contribute to the existing knowledge regarding the associations between different dimensions and elements of physical activity (PA) and sedentary behaviour (SB) with core Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT), namely, patient- and clinician important health outcomes (i.e., pain, disease activity, functional ability, fatigue, depression, anxiety, subjective vitality, and quality of life (Bartlett et al., 2012; Boers et al., 1994; Van Tuyl & Boers, 2015)) in Rheumatoid Arthritis (RA). An illustrated overview of the chapters and key findings of this thesis is provided in **Figure 7.1**.

The following section highlights the main findings of this thesis and suggests some future directions in this area of research as a consequence of the findings.

1. **Existing lifestyle PA interventions are effective at increasing PA, reducing SB and improving core OMERACT health outcomes in people with RA, whilst further SB interventions are needed to confirm their effects;**

In healthy individuals and people living with musculoskeletal conditions, interventions which focus on incorporating PA into daily life have been increasingly advocated by researchers and participants alike (Khoja et al., 2016). Results from the systematic review and meta-analysis conducted in ***Chapter 2*** revealed that lifestyle PA and SB interventions had significant positive effects on disease activity, daily steps, moderate to vigorous PA (MVPA), leisure/light intensity PA, sedentary time, functional ability, depression, and fatigue in people with RA. Therefore, this review provided evidence that interventions targeting lifestyle PA and SB (i.e., all PA that is accumulated as part of daily life, such as incidental PA, home-based PA, walking, or reducing SB) offer a promising approach to improve OMERACT patient- and clinician-important outcomes in people with RA.

Lifestyle PA interventions focused on increasing specific FITT dimensions such as MVPA, steps and leisure/light PA were particularly effective at improving OMERACT outcomes and should therefore be recommended and built upon in future interventions. PA interventions were also effective at improving disease activity, clinicians should recommend lifestyle PA as a form of RA disease management.

Compared to PA, only 1 intervention was identified as part of the systematic review which focused on reducing SB in people with RA (Thomsen et al., 2017). As well as the lack of SB interventions in existing literature, few observational studies have investigated the associations between sedentary time and core OMERACT outcomes in RA (Fenton et al., 2017; Fenton et al., 2018b; Hammam et al., 2019; O'Leary et al., 2021). Longitudinal findings have reported that a bi-directional relationship exists between activPAL-assessed sedentary time with pain and fatigue, and suggest that the replacement of sedentary time with standing may be an effective avenue to improve these outcomes (O'Brien et al., 2021). The single SB intervention conducted in people with RA did demonstrate significant improvements in OMERACT outcomes, but these results had the potential to be underpowered (Thomsen et al., 2017). In addition, this SB intervention had significant imbalances between intervention groups, and consequently scored as high risk in risk of bias assessment. Results regarding the efficacy of SB intervention should therefore be interpreted with caution. All in all, findings from the systematic review and meta-analyses highlight the need for further SB interventions that are sufficiently powered to fully understand what role SB has in RA, and if interventions targeting SB can be effective.

However, prior to the development of SB interventions, additional research is required to establish the relationships between SB and core OMERACT outcomes (Fenton et al., 2017; Greene et al., 2006; Huffman et al., 2014; Khoja et al., 2016; O'Leary et al., 2021). Studies which explore the amounts and patterns of SB, and the links to specific RA outcomes will improve understanding regarding the role

of different dimensions of SB (i.e., SITT elements: e.g., sedentary bouts, sedentary interruptions, context of SB, (Tremblay et al., 2010)) and how they may impact RA outcomes. This will, in turn, help to inform more targeted and effective SB dimension-specific interventions, which may have greater potential to improve core OMERACT outcomes.

2. **Lifestyle PA and SB interventions were heterogeneous in terms of selection and measurement of OMERACT outcomes. More methodological consistency across studies is required;**

The research comprising this thesis was focused on patient- and clinician-important RA outcomes defined by OMERACT (Bartlett et al., 2012; Boers et al., 1994; Van Tuyl & Boers, 2015). The OMERACT initiative created these core outcome sets in order to provide a consensus between researchers on optimal outcomes and outcome measures for use in RA clinical trials and allow pooling of data in meta-analyses (Boers et al., 2014).

Despite this, findings from ***Chapter 2*** indicate that researchers continue to target varied outcomes and use different measurement tools in lifestyle PA and SB interventions for RA participants. For example, the most common measure of pain in interventions included in the review was a visual analogue scale (VAS) (Scott & Huskisson, 1976). Pain is a notoriously difficult construct to measure, and many self-report tools, including the VAS, are single-item and only ask about overall pain severity or intensity (Scott & Huskisson, 1976). As such, they are unable to quantify the multidimensionality of pain (e.g., assessing sensory and affective aspects) nor the central and peripheral mechanisms present in RA-related pain processing (Walsh & McWilliams, 2014). In contrast, psychophysical measurement methods, such as Quantitative Sensory Testing (QST), may give a more comprehensive overview of the multiple pain mechanisms present in RA (McWilliams & Walsh, 2017).

In addition to pain, measures used to assess other OMERACT outcomes also varied between studies. For example, whilst multidimensional measures of fatigue are validated, reliable and available for use in research (e.g., multidimensional assessment of fatigue (MAF), multidimensional fatigue inventory (MFI) (Hewlett et al., 2007; Smets et al., 1995)), many studies continue to use a single-item VAS which is unable to give a comprehensive overview of different aspects of fatigue. Single-item scales also have the drawback that they have limited sensitivity, lack a measure of internal reliability and don't capture factorial or content validity (McIver & Carmines, 1981).

Future studies in people with RA should therefore seek to use OMERACT-approved multidimensional measures for all OMERACT outcomes. This would enable researchers to elucidate optimally effective interventions impacting such outcomes, and create beneficial patient guidance to improve RA-related health. In addition, this will increase between-study methodological consistency, enable further subgroup analysis and data pooling, and permit direct comparison between interventions. The lack of consistency in measures between existing studies limits the ability to calculate mean differences, and assess clinical significance of meta-analysis findings. As a result, the systematic review within this thesis could not include clinically meaningful recommendations for many of the core OMERACT outcomes.

Still, whilst OMERACT provides a useful framework to inform research in RA, this and other initiatives should seek to develop more specific outcome sets for different subcategories of RA (i.e., early RA, advanced disease etc.). Indeed, there may be important differences in aspects of health that are deemed as most important in these sub-populations. Initiatives should also encourage researchers to use a greater variety of reliable and valid multidimensional measurement tools to measure these outcomes, and continue to develop further core outcomes which encompass other patient- and clinician-important aspects of health in people with RA.

3. **QST psychophysical measurement methods can be reliably used to quantify multiple parameters of pain in healthy adults and in patients with lower back pain (LBP) and RA (using a protocol consisting of pressure-pain threshold (PPT) and temporal summation (TS));**

Results from ***Chapter 3*** concluded that PPT and TS (calculated as a wind-up difference) can be reliably used to measure multiple parameters of pain processing in healthy people, and individuals with LBP and RA, when conducted on both the forearm and lower leg. However, the study detailed in ***Chapter 3*** also established that further research is needed to confirm the reliability of conditioned pain modulation (CPM) using larger sample sizes and in different musculoskeletal populations. Previous studies have highlighted the methodological advantages of QST compared to self-report questionnaires (Arendt-Nielsen et al., 2015; Middlebrook et al., 2020; Pavlakovic & Petzke, 2010). Both peripheral and central pain mechanisms are thought to contribute towards pain sensitivity in RA (Walsh & McWilliams, 2014). QST conducted at different anatomical sites can provide information about the relative contribution of these pain mechanisms. This is important as QST can provide novel data through being able to characterise underlying RA-related pain processing pathways (McWilliams & Walsh, 2017). QST may also aid in the development of population- or patient-specific prescriptions to target and minimise pain.

The present findings encourage future studies to use QST in order to obtain multidimensional quantitative information about pain processing in people with RA. By using QST in future PA and SB studies in RA, this may enable researchers to identify more specific links and targets (i.e., central vs peripheral pain mechanisms) for subsequent behavioural interventions.

4. **Lifestyle PA and SB interventions were heterogeneous in terms of selection and measurement of PA and SB. More methodological consistency across studies is required;**

The significant heterogeneity between different interventions in terms of outcomes assessed and measures employed may have contributed towards the varied meta-analysis results in ***Chapter 2***. The majority of studies employed self-report questionnaires to measure PA and/or SB in existing lifestyle PA and SB interventions. Interestingly, most studies which used device-based tools to assess PA and/or SB demonstrated significant improvements in PA, SB, and OMERACT outcomes post-intervention. In contrast, studies employing self-report methods generally reported fewer significant effects in response to intervention. Patients with RA have previously shown to over-estimate PA and underestimate sedentary time using self-report measures (Yu et al., 2015). Nevertheless, devices are more "objective" than self-report, and so may be more likely to detect changes in PA and/or SB following intervention owing to greater validity/accuracy. This inaccuracy may translate into why self-reported PA/SB outcomes are not providing consistent effects as a result of lifestyle PA and SB interventions.

However, existing interventions which used device-based measures of PA and SB, such as accelerometers, were still very different, in terms of accelerometer processing decisions used, measurement and wear periods, and outcomes chosen to reflect PA and/or SB. This emphasises the need for a standardised approach to assessing PA and SB outcomes so it can be possible to compare, contrast and conduct meta-analyses stemming from the findings of interventions. Devices which have been validated and demonstrate high reliability in people with RA include the ActiGraph GT3X (for measuring free-living PA) and the activPAL (for measuring free-living sedentary time) (O'Brien et al., 2020). Therefore, I recommend these devices to be used as valid and reliable measures of PA and SB in future interventions in people with RA.

Accelerometers can be superior to self-report as they can reliably capture the intensity of free-living behaviours (i.e., ${}_{F}I_{TT}$ and ${}_{S}I_{TT}$). These dimensions are particularly important in people with RA as these behaviours have demonstrated differing associations with health outcomes (Fenton et al., 2017;

Khoja et al., 2016). Notwithstanding this recommendation, the study described in ***Chapter 6*** used self-report measures to quantify and distinguish between the different types and intensities of exercise and non-exercise PA and SB. Self-report questionnaires can provide context to behaviours and activities by investigating participation in specific types of PA or SB. These are important components of the FITT and SITT acronyms (i.e., $_{FIT}$T and $_{SIT}$T) which device-based measures are unable to record (Troiano et al., 2014). Types and context of activity are important dimensions of PA and SB which may need to be accurately measured in studies targeting specific behaviours (e.g., reducing TV viewing or increasing engagement in leisure time PA). In addition, self-report measures are less burdensome for participants, can more easily quantify behaviours in large-scale epidemiological studies, and are cheaper to administer than device-based measures (Sylvia et al., 2014). In studies involving large sample sizes, such as the study described in ***Chapter 6***, it is not viable or feasible to use device-based measures of PA.

As described in ***Chapter 6***, self-reported non-exercise PA and walking showed independent associations with OMERACT health outcomes. This highlights the contribution of self-report measures to assess specific dimensions of PA and SB. Therefore, findings of this thesis include that a combination of validated and reliable device-based (to assess time in PA/SB) and self-report measures (to explore the relative contributions of non-exercise types of PA and SB) should be recommended to quantify PA and SB in future studies in people with RA. As the ActiGraph GT3X and activPAL have been validated in people with RA (O'Brien et al., 2020), these are two such device-based measures that I would recommend for use in future studies in this population. In addition to this, there is a need to develop a reliable and valid self-report measure capable of capturing the context of SB and PA in RA. Another approach to categorise type and context of free-living PA could be the use of wearable cameras to complement existing accelerometry measures (Doherty et al., 2013). However, as this method is still relatively novel, wearable cameras would need to be validated and deemed reliable specifically in people with RA.

5. **Accelerometer model and placement site specific cut-points must be applied to ActiGraph accelerometers in order to give comparable estimations of free-living PA and sedentary time;**

Where accelerometers are used to measure PA and sedentary time, analytical methods should consider the specific model (i.e., uniaxial or triaxial) and placement site (i.e., wrist- or hip-worn) of the device. Findings from ***Chapter 4*** demonstrated that uniaxial vs triaxial and hip vs wrist cut-points are not comparable for estimation of free-living sedentary time, light intensity PA (LPA) and MVPA, when applied to the ActiGraph GT9X in healthy adults. Cut-points may not be appropriate for employment to different ActiGraph monitors to the ones they were specifically designed for, and application to different models or placement sites could lead to activity intensity misclassification (Migueles et al., 2017; Rhudy et al., 2020).

The findings from ***Chapter 4*** point to the importance of ensuring comparisons between PA and sedentary time estimates taken from different accelerometer models and placement sites are valid. Participants in the study described in ***Chapter 4*** were young, healthy adults. Therefore the findings described above need to be replicated in the RA population to make these conclusions in people with RA.

Nevertheless, findings suggested that given the greater granularity of the data which can be collected with triaxial accelerometers (vs uniaxial), future studies should seek to employ triaxial accelerometers, and ensure data is analysed using triaxial cut-points calibrated for a specific device and placement site (hip vs wrist) (Migueles et al., 2017). Despite these recommendations that triaxial model and placement site specific cut-points be employed to accelerometers, large-scale epidemiological studies, such as National Health and Nutrition Examination Survey (NHANES) (Matthews et al., 2008), continue to employ early uniaxial hip-specific accelerometer cut-points to data obtained from wrist-worn triaxial accelerometers to estimate PA and sedentary time (Troiano et

al., 2008). This could have a significant impact on the accuracy of the PA and sedentary time data reported in such studies, which are frequently referenced as “comparator studies/cohorts” in smaller scale research (Koster et al., 2016; O'Brien et al., 2020; Rhudy et al., 2020).

As well as using triaxial models with corresponding cut-points, I recommend the use of devices that have been specifically validated in people with RA (e.g., ActiGraph and activPAL, (O'Brien et al., 2020)). The more consistent use of valid and reliable measures of PA and SB will enable researchers to more accurately pool data, strengthen the evidence regarding the engagement with PA and SB in RA, and elucidate the links between different dimensions of PA and SB with other aspects of health. This may help the development of specific clinical recommendations for PA and SB for people with RA. Therefore, a future recommendation of this thesis is the use of the ActiGraph and activPAL to quantify total and intensities of free-living PA and sedentary time, respectively, in people with RA. As mentioned earlier, researchers should also include self-report measures to explore the relative contributions of non-exercise types of PA and SB, which have demonstrated independent associations with OMERACT health outcomes in this thesis.

6. **Different intensities and types of PA and SB display differing associations with core OMERACT outcomes in people with RA during the COVID-19 pandemic.**

Results reported in ***Chapter 6*** revealed that non-exercise LPA, walking, exercise, and sedentary time had complex, individual, and distinct relationships with OMERACT indicators of mental health and wellbeing in people with RA. With non-exercise LPA and walking having the most statistically significant findings, this indicated that lower-intensity, non-exercise types of PA could offer effective targets in future interventions to improve mental health and wellbeing in RA. To my knowledge, no interventions have specifically targeted LPA in people with RA, and limited longitudinal studies have been conducted exploring the relative contributions of different intensities and types of PA for health in people with RA (O'Brien et al., 2021). Nonetheless, ***Chapter 6*** findings and other cross-sectional

studies have recently established there is a link between LPA and health in RA (Fenton et al., 2017; Khoja et al., 2016). This is important as non-exercise behaviours (such as SB and LPA) encompass the majority of waking hours in people with RA (Hammam et al., 2019; Summers et al., 2019). Further, the disease-related barriers faced by people with RA indicate that interventions targeting PA of a lower intensity nature may be more acceptable and feasible in this population (Fenton et al., 2017; Larkin et al., 2017; Veldhuijzen Van Zanten et al., 2015).

Recommendations and official guidance regarding PA and SB engagement during the COVID-19 pandemic were scarce. This was particularly true at the start of the pandemic when the study comprising ***Chapter 6*** was conducted. At this time, the risks associated with developing COVID-19 for people with clinical conditions were assumed to be significant due to the immunosuppressive nature of RA. In addition, those deemed most at risk, including many people with RA, were advised to self-isolate. As a consequence, PA engagement was particularly low in the study described in ***Chapter 6***. As PA has demonstrated positive effects on OMERACT health outcomes in findings of this thesis, it is no surprise that decreased non-exercise LPA, walking, exercise and increased sedentary time during the COVID-19 pandemic were related to poor health outcomes in this study. I recommend that longitudinal studies are conducted during the COVID-19 pandemic to determine the direction and causality of these associations. Nevertheless, one of the main implications of ***Chapter 6*** was that population-specific PA recommendations and guidance are needed during pandemics, in order to mitigate and limit the detrimental effects of low PA and increased SB to health.

Summary and Recommendations

The key findings of this thesis are summarised in **Figure 7.1**. Together, findings from ***Chapter 2 – 6*** in this thesis, provide an insight into the nuanced relationships between different dimensions and elements of PA and SB with health in people living with RA, including during the COVID-19 pandemic.

In doing so, thesis findings highlight key considerations regarding measurement methodologies (***Chapter 3*** and ***4*** findings- **Figure 7.1**).

All in all, results of this thesis (**Figure 7.1**) point to the recommendation that targeting non-exercise types of PA, such as lifestyle PA, non-exercise LPA, could represent effective avenues for future interventions to improve core OMERACT patient- and clinician-important outcomes in people with RA. Interventions should continue to focus on specific frequencies, intensities and types of PA and SB (i.e., FITT and SITT elements) using a combination of device-based and self-report measures. However, experimental research is required to confirm the present findings, establish cause and effect, and better understand the patterns of non-exercise PA that might be most beneficial for health in people with RA. Based on what is known from longitudinal findings (O'Brien et al., 2021), national PA and SB guidelines (Bull et al., 2020), and conclusions from this thesis, I encourage the development of experimental studies which replace sedentary time with LPA behaviours (i.e., “move more”) to target OMERACT outcomes.

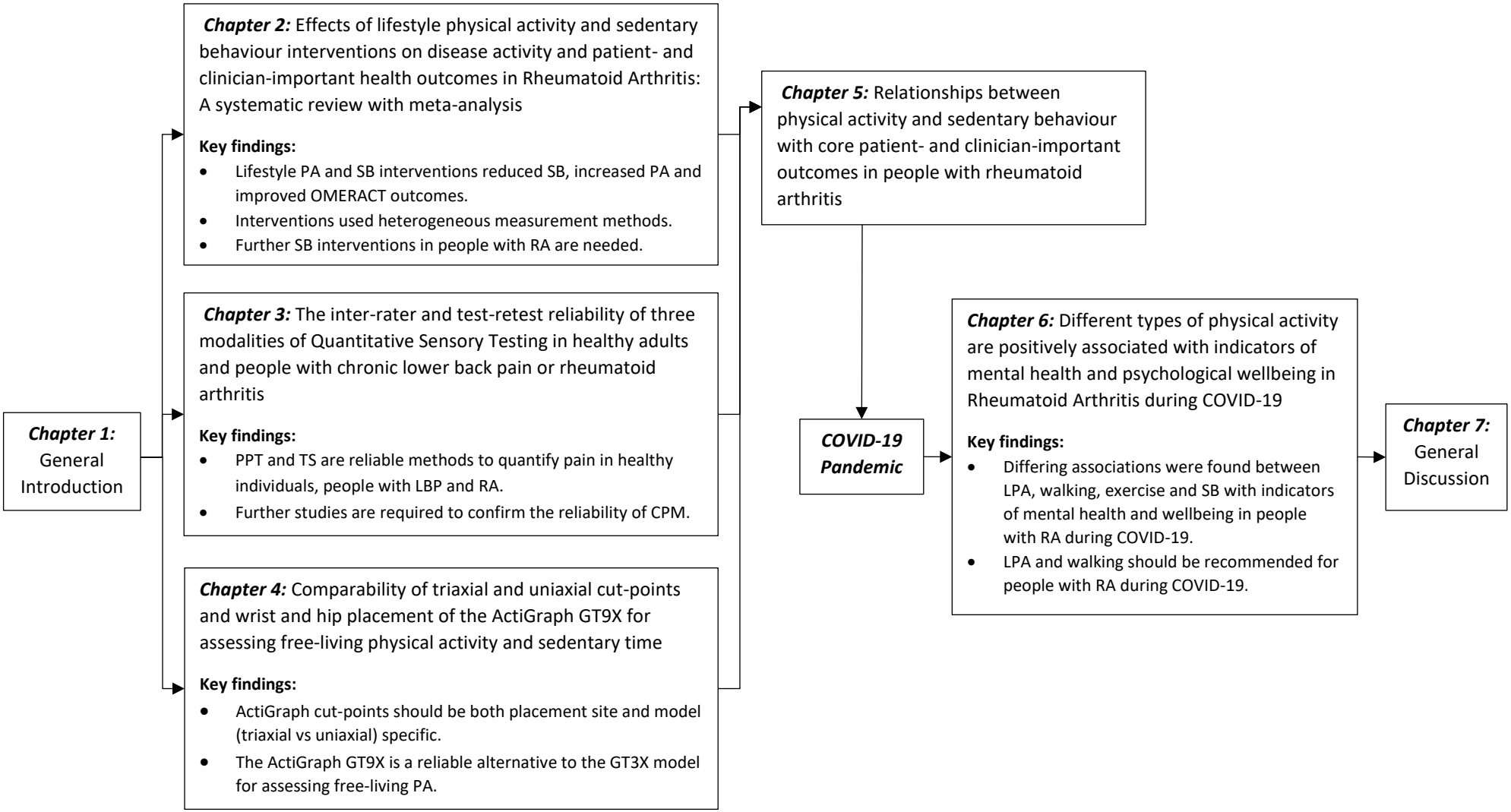


Figure 7.1: *Flowchart of all chapters as part of this thesis*

Note: RA= Rheumatoid Arthritis, PA= Physical Activity, SB= Sedentary Behaviour, LBP= Lower Back Pain, QST= Quantitative Sensory Testing, PPT= Pressure-pain Threshold, TS= Temporal Summation, CPM= Conditioned Pain Modulation, LPA= Light intensity Physical Activity.

Reflections throughout PhD

Over the past 3+ years of completing this PhD, I have developed as both a researcher and an individual. I have been taught how to critique and analyse many forms of research and data, and these are skills which I currently and will continue to use both professionally and in all areas of my life.

Through conducting a systematic review, I have gained an appreciation of the need for openness and transparency in research. Research should be available and reproducible to all, and this was highlighted to me through the various stages, from registration to meta-analyses. The many steps involved in completing the review emphasised to me the complexity and many stages involved in timelines of research and clinical trials. I also learnt the importance of teamwork through completing this review. Although completing a PhD is an independent task and I have found it easy to get lost in my own "research bubble", conducting a systematic review requires a team of personnel. These collaborations as part of conducting the review have improved my patience, communication skills and encouraged me to work more collaboratively in my subsequent projects. Although the systematic review was initially daunting, it was one of my favourite projects as part of my PhD as it appealed to my organised and structured way of working. I now work in a more systematic manner, which will prove useful in my future career.

The methodological chapters of my thesis (i.e., reliability studies as part of ***Chapter 3*** and ***4***) have given me an understanding of the need for accurate and reliable measurement tools, and the statistical processes involved in determining reliability and validity. If measurement tools or processes have not been proven reliable and valid, this can limit interpretation of the research they are subsequently employed in. I believe I am now equipped with enhanced statistical knowledge on how to analyse reliability and validity, and I feel confident that I will use these skills in research I will conduct as part of my continuing academic career.

I have also gained invaluable knowledge on the processes involved in setting up clinical research studies. I have created and had approved ethical applications and amendments at both a university and NHS level. In addition, I have gained experience with all steps involved in setting up and conducting small-scale clinical studies, as well as invaluable experience working with a variety of research staff, other students, clinicians, academic researchers and patients. Through working with such a multidisciplinary team, and being required to do face-to-face recruitment of patients, this has greatly improved my confidence and communication skills.

The COVID-19 pandemic provided a unique event whereby the main research study as part of this thesis could not proceed. This period of uncertainty taught me resilience and mental strength. Through the innovative thinking of my supervisors and quick and decisive action, I was able to continue my research and minimise the COVID-related disruption. One of my proudest achievements from this PhD was the ability to quickly set up and conduct a clinically-important study investigating PA, SB and health in people with RA during the COVID-19 pandemic. By publishing these findings, I was able to contribute to the limited evidence and guidance for people with RA during this unique global event.

This PhD equipped me with the skills and expertise required for conducting research in clinical and disadvantaged populations. My motivation has been tested throughout, and my resilience and dedication is proof to me that PA and SB research is something I am truly passionate about. Consequently, I hope to continue in the field of academic research, and further increase my own and others knowledge on the importance of PA and SB for health in other populations.

Conclusion

Taken in their totality, the results of this thesis indicate that although interventions targeting lifestyle PA and SB in people with RA are effective at improving core OMERACT outcomes, more consistency is required in regard to the outcomes targeted and the measurement tools employed to assess those outcomes (***Chapter 2***). Future research in RA should use reliable and valid measures which provide novel and multidimensional information about outcomes (i.e., QST (***Chapter 3***)) and PA and SB (i.e., ActiGraph and activPAL) (***Chapter 4***). Self-report tools can also be employed in conjunction with device-based measures of PA and sedentary time, to provide information about non-exercise PA and the contexts/activities in which PA and SB are realised (***Chapter 6***). Finally, the present findings suggest that interventions which target certain types and/or intensities of non-exercise PA (e.g., lifestyle PA, non-exercise LPA, walking) or SB may be effective at improving OMERACT patient- and clinician-important outcomes in people with RA (***Chapters 2*** and ***6***). Therefore, as a next step, researchers should develop interventions using reliable and valid measures (i.e., QST, ActiGraph, activPAL) for people with RA targeting: 1) increasing non-exercise types of PA (e.g., lifestyle PA, light non-exercise PA, walking); and 2) reducing SB, in order to target core patient- and clinician-important outcomes as defined by OMERACT.

Appendix 3.1: Quantitative Sensory Testing Script

1. Introduce myself and my background and PhD area- how this study links in to PhD. Emphasise the impact this can have on RA patients, may influence treatment and recommendations.

   "Hello, thank you for coming today. I'm Sophia and I'm a PhD student, studying physical activity and sedentary behaviour in rheumatoid arthritis patients. As part of my preliminary study, I am testing the reliability of some assessments of **how we feel pain** and this is why you are here today. As an RA patient, I hope that the results of my study will **influence treatment and recommendations** to have a positive impact on your quality of life."

2. Explain the study (3 assessments which I will explain in detail).

   "This study will involve 3 short assessments which will measure **how you feel pain**. I will do the 3 assessments now and these assessments will then be repeated by my colleague once you feel ready to do them again and then I will repeat them once more at a later date."

   "The assessments have been done many times before, by myself and colleagues in Nottingham. They are **safe**, and patients have not reported **any long term side effects**. The pain you will feel should be **mild and transient** but you can **stop** or **withdraw** at any point and do not need to give a reason. "

3. Confirm that volunteer meets criteria and has read and understands the participant information sheet, obtain consent.

   Meet criteria
   Read and understands participant information sheet
   Given consent

Pressure-Pain Threshold (PPT)

"The idea of this test is to look at your pressure to pain threshold. We are not looking at how much pain you can tolerate, simply at **what point you start to feel pain**. The pain you feel will only be **fleeting**, as the test will be stopped as soon as you indicate that you have started to feel pain. You will hold this push button in your dominant hand and I will start to apply a graded pressure to the front of your lower leg (tibialis anterior muscle). You will feel pressure as the probe is pressed down and the pressure will be gradually increased. **As soon as the pressure starts to change to pain**, you should press the button and I will withdraw the probe."

"We will do a **practice test** on your non-dominant lower leg, to let you know how it feels. I will then do the same **three** times on your **dominant** lower leg. There will be a break in between each test."

PPT Procedure in detail:

Ensure patient is not experiencing more than their normal level of pain before applying stimuli. Ask participant to **close their eyes** during testing.

1. **Mark testing site** as 5cm below base of knee cap just off to the side (on the muscle, not bone) on both legs.
2. Do practice test on non-dominant leg.
3. Begin Test Site: Record three measurements

Apply the probe perpendicular to the skin and start the force application. Apply graded pressure.

When the patient starts to feel that the pressure has changed to pain, he/she presses the button. This gives an audible indication. At this point, the probe is withdrawn and the pain threshold **recorded**.

Manually record PPT level as well as via automated data collection unit.

Temporal Summation (TS):

The patient will retain the same relaxing siting position while the examiner will apply the pen, which features a retractable blunt needle, in repetitive manner (once per second for ten seconds) on the front of the individual's upper leg, approx. 5cm above the patella. The individual will feel light pricking with every application and will be asked to specify the intensity of pain, on a 0 to 10 scale, on the first and on the last time. The given scores will be noted on a piece of paper.

Participant preparation:

- Explain the procedure to the participant:

"This is a test of your ability to detect a sensation of '**Sharpness'** or 'stinging'. For this test, we will use a weighted **blunt** needle. This will be **pressed gently against your skin** on your upper-leg. The stimulator is only applied for 1 second, so the sensation of sharpness will be **temporary**. The pinprick stimulator is designed **not to puncture** your skin and is disinfected before we use it on every participant."

- Participant should **close their eyes** during definitive testing
- Ensure the participant is lying comfortably
- Identify and **mark** test site: Upper-leg (quadriceps, 5cm from superior to the mid-point of the patella)

Training session:

- Perform one measure on the arm for training purposes only.

"I will perform one measure on your **arm just as a practice** to show you what the procedure is like"

"Once the experiment starts, I will apply a stimulator to **the front of your dominant leg**. Please rate the pain or sharpness you experience from 0-10 where 0 indicates no pain or sharpness and 10 indicates the most intense pain or sharpness imaginable. I will then apply the same stimulator at the same site 10 times repeatedly at a rate of 1/second. After completing the 10 pinpricks, please rate the **average** pain or sharpness you experienced during all 10 pinpricks, from 0-10 where 0 indicates no pain or sharpness and 10 indicates the most intense pain or sharpness imaginable. I will then **repeat** the procedure."

1. Apply pinprick perpendicular to the skin of the quadriceps, 5cm above the patella
2. Ask the participant to rate the pain or sharpness they experience from 0-10 where 0 indicates no pain or sharpness and 10 indicates the most intense pain or sharpness imaginable.
3. Record the rating.
4. Apply the same stimulator at the same site 10 times repeatedly at a rate of 1/second.
5. After completing the 10 times pinprick test, ask the participant how painful was the average stimulation. From 0-10 where 0 indicates no pain or sharpness and 10 indicates the most intense pain or sharpness imaginable.
6. Repeat the procedure.
7. The wind-up will be calculated as the average rating of the 2 procedures

Conditioned Pain Modulation (CPM):

For the purposes of CPM, a manual blood pressure sphygmomanometer will be used in combination with the electronic algometer mentioned above. Before the use of the sphygmomanometer, the algometer will be used to assess the individual's PPT at the lower-leg in a procedure identical to the one described above. The participants are expected to feel pain at a lower pressure during the second PPT measurement.

"The final test is called Conditioning Pain Modulation (CPM), and it involves **simultaneously** applying pressure to your upper arm via a blood pressure cuff and repeating the pain threshold test on your lower leg."

"The computer records your pain threshold of each test and this will be compared with other volunteers. If, for any reason, you want to stop, let me know straight away."

"Do you have any questions?"

"I will first repeat the pain-pressure threshold test that I did before on your dominant lower leg, to act as a **comparison** before I do it again with the blood pressure cuff on."

PPT CPM procedure in detail:

- Do PPT once more to act as a comparison for CPM

"I am now going to apply continuous pressure to your lower-leg as before, but with the aid of a blood pressure cuff and ask you to do some hand grip exercises. I would like you to let me know when you would rate the pain as **4/10**. I would describe this as **uncomfortable but bearable** for a **short amount of time**."

"I will then **immediately** re-test your pain threshold on your lower-leg with the cuff still inflated. As before, as soon as the pressure starts to change the pain, press the button and I will withdraw the probe, the cuff will also **deflate**."

Operator Instructions:

- Read out verbal instructions.
- Wrap cuff around the contralateral arm to the leg being tested.
- Set systolic pressure to 270mmHg and do not exceed. After target pressure is reached, ask participant to rate sensation in arm from 0-10.
- Repeat handgrip ≥10 times until 4 is reached on numerical rating scale (NRS). Ask for NRS rating every five handgrips.
- Once NRS 4 achieved, apply probe in same manner as before to lower-leg test site.
- Once participant presses button, withdraw probe and immediately release cuff.
- Manually record PPT level and systolic pressure value where the NRS target (≥4) was reached.
- Calculate CPM effect: difference in average threshold values (in kPa) of the two test stimuli (with conditioning – without conditioning). Positive value = efficient CPM.

Before testing: Sterilise probe head (prior to each patient)

BIBLIOGRAPHY

Aadland, E., & Ylvisaker, E. (2015). Reliability of the Actigraph GT3X+ Accelerometer in Adults under Free-Living Conditions. *Plos One, 10*(8), e0134606. doi:10.1371/journal.pone.0134606

Aguilar-Farías, N., Brown, W. J., & Peeters, G. M. (2014). ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. *J Sci Med Sport, 17*(3), 293-299. doi:10.1016/j.jsams.2013.07.002

Aguilar-Farias, N., Peeters, G. M. E. E. G., Brychta, R. J., Chen, K. Y., & Brown, W. J. (2019). Comparing ActiGraph equations for estimating energy expenditure in older adults. *Journal of Sports Sciences, 37*(2), 188-195. doi:10.1080/02640414.2018.1488437

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., O'Brien, W. L., Bassett, D. R., Jr., Schmitz, K. H., Emplaincourt, P. O., Jacobs, D. R., Jr., & Leon, A. S. (2011). 2011 Compendium of Physical Activities: A Second Update of Codes and MET Values. *Medicine & Science in Sports & Exercise, 43*(8), 1575-1581.

Aletaha, D., Neogi, T., Silman, A. J., Funovits, J., Felson, D. T., Bingham, C. O., Birnbaum, N. S., Burmester, G. R., Bykerk, V. P., Cohen, M. D., Combe, B., Costenbader, K. H., Dougados, M., Emery, P., Ferraccioli, G., Hazes, J. M. W., Hobbs, K., Huizinga, T. W. J., Kavanaugh, A., Kay, J., Kvien, T. K., Laing, T., Mease, P., Ménard, H. A., Moreland, L. W., Naden, R. L., Pincus, T., Smolen, J. S., Stanislawska-Biernat, E., Symmons, D., Tak, P. P., Upchurch, K. S., Vencovský, J., Wolfe, F., & Hawker, G. (2010). 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis, 69*(9), 1580. doi:10.1136/ard.2010.138461

Aletaha, D., Smolen, J. S., & Ward, M. M. (2006). Measuring function in rheumatoid arthritis: Identifying reversible and irreversible components. *Arthritis & Rheumatism, 54*(9), 2784-2792. doi:10.1002/art.22052

Almutairi, K., Nossent, J., Preen, D., Keen, H., & Inderjeeth, C. (2021). The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatology International, 41*(5), 863-877. doi:10.1007/s00296-020-04731-0

Amagasa, S., Machida, M., Fukushima, N., Kikuchi, H., Takamiya, T., Odagiri, Y., & Inoue, S. (2018). Is objectively measured light-intensity physical activity associated with health outcomes after adjustment for moderate-to-vigorous physical activity in adults? A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 15*(1), 65. doi:10.1186/s12966-018-0695-z

Antunes, R., Frontini, R., Amaro, N., Salvador, R., Matos, R., Morouco, P., & Rebelo-Goncalves, R. (2020). Exploring Lifestyle Habits, Physical Activity, Anxiety and Basic Psychological Needs in a Sample of Portuguese Adults during COVID-19. *International Journal of Environmental Research and Public Health, 17*(12). doi:10.3390/ijerph17124360

Arendt-Nielsen, L., Egsgaard, L. L., Petersen, K. K., Eskehave, T. N., Graven-Nielsen, T., Hoeck, H. C., & Simonsen, O. (2015). A mechanism-based pain sensitivity index to characterize knee osteoarthritis patients with different disease stages and pain levels. *Eur J Pain, 19*(10), 1406-1417. doi:10.1002/ejp.651

Arendt-Nielsen, L., & Yarnitsky, D. (2009). Experimental and clinical applications of quantitative sensory testing applied to skin, muscles and viscera. *Pain, 10*, 556-572. doi:10.1016/j.jpain.2009.02.002

Arendt-Nielsen, L., Morlion, B., Perrot, S., Dahan, A., Dickenson, A., Kress, H., Wells, C., Bouhassira, D., & Mohr Drewes, A. (2018). Assessment and manifestation of central sensitisation across different chronic pain conditions. *European Journal of Pain, 22*(2), 216-241.

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., Luthra, H. S., & et al. (1988). The American Rheumatism

Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatism, 31*(3), 315-324. doi:10.1002/art.1780310302

Arthritis Research UK. (2018). State of Musculoskeletal Health 2017: Arthritis & other musculoskeletal conditions in numbers.

Arts, E. E. A., Fransen, J., den Broeder, A. A., Popa, C. D., & van Riel, P. L. C. M. (2015). The effect of disease duration and disease activity on the risk of cardiovascular disease in rheumatoid arthritis patients. *Ann Rheum Dis, 74*(6), 998-1003. doi:10.1136/annrheumdis-2013-204531

Ayhan, F., Gül, S., Uyar, S., Erdem, R., & Borman, P. (2014). The Decreased Sensory Thresholds in Rheumatoid Hand: Comparisons with Osteoarthritic and Normal Hands. *Journal of Physical Medicine & Rehabilitation Sciences, 17*(3).

Balboa-Castillo, T., León-Muñoz, L. M., Graciani, A., Rodríguez-Artalejo, F., & Guallar-Castillón, P. (2011). Longitudinal association of physical activity and sedentary behavior during leisure time with health-related quality of life in community-dwelling older adults. *Health and Quality of Life Outcomes, 9*, 47-47. doi:10.1186/1477-7525-9-47

Balchin, C., Tan, A. L., Wilson, O., McKenna, J., & Stavropoulos-Kalinoglou, A. (2021). Participation in physical activity decreased more in people with rheumatoid arthritis than the general population during the COVID-19 lockdown: a cross-sectional study. *Rheumatology International*, 1-10. doi:10.1007/s00296-021-05054-4

Balestroni, G., & Bertolotti, G. (2012). EuroQol-5D (EQ-5D): an instrument for measuring quality of life. *Monaldi Arch Chest Dis, 78*(3), 155-159. doi:10.4081/monaldi.2012.121

Balsamo, S., Diniz, L. R., dos Santos-Neto, L. L., & da Mota, L. M. (2014). Exercise and fatigue in rheumatoid arthritis. *Isr Med Assoc J, 16*(1), 57-60.

Barber, C. E. H., Mosher, D. P., Ahluwalia, V., Zummer, M., Marshall, D. A., Choquette, D., Lacaille, D., Bombardier, C., Lyddiatt, A., Chandran, V., Khodyakov, D., Dao, E., & Barnabe, C. (2017). Development of a Canadian core clinical dataset to support high-quality care for canadian patients with rheumatoid arthritis. *Journal of Rheumatology, 44*(12), 1813-1822.

Bartlett, S. J., Bykerk, V. P., Cooksey, R., Choy, E. H., Alten, R., Christensen, R., Furst, D. E., Guillemin, F., Halls, S., Hewlett, S., Leong, A. L., Lyddiatt, A., March, L., Montie, P., Orbai, A. M., Pohl, C., Scholte Voshaar, M., Woodworth, T. G., & Bingham, C. O. (2015). Feasibility and domain validation of rheumatoid arthritis (RA) flare core domain set: Report of the OMERACT 2014 RA Flare Group plenary. *Journal of Rheumatology, 42*(11), 2185-2189.

Bartlett, S. J., Hewlett, S., Bingham, C. O., 3rd, Woodworth, T. G., Alten, R., Pohl, C., Choy, E. H., Sanderson, T., Boonen, A., Bykerk, V., Leong, A. L., Strand, V., Furst, D. E., & Christensen, R. (2012). Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. *Ann Rheum Dis, 71*(11), 1855-1860. doi:10.1136/annrheumdis-2011-201201

Bisset, L. M., Evans, K., & Tuttle, N. (2015). Reliability of 2 protocols for assessing pressure pain threshold in healthy young adults. *J Manipulative Physiol Ther, 38*(4), 282-287. doi:10.1016/j.jmpt.2015.03.001

Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Ann Intern Med, 162*(2), 123-132. doi:10.7326/m14-1651

Bize, R., Johnson, J. A., & Plotnikoff, R. C. (2007). Physical activity level and health-related quality of life in the general adult population: A systematic review. *Preventive Medicine, 45*(6), 401-415. doi:10.1016/j.ypmed.2007.07.017

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research, 8*(2), 135-160.

Blumenthal, J. A., Babyak, M. A., Moore, K. A., Craighead, W. E., Herman, S., Khatri, P., Waugh, R., Napolitano, M. A., Forman, L. M., Appelbaum, M., Doraiswamy, P. M., & Krishnan, K. R.

(1999). Effects of Exercise Training on Older Patients With Major Depression. *Archives of Internal Medicine, 159*(19), 2349-2356. doi:10.1001/archinte.159.19.2349

Boers, M., Kirwan, J. R., Wells, G., Beaton, D., Gossec, L., d'Agostino, M. A., Conaghan, P. G., Bingham, C. O., 3rd, Brooks, P., Landewé, R., March, L., Simon, L. S., Singh, J. A., Strand, V., & Tugwell, P. (2014). Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. *Journal of Clinical Epidemiology, 67*(7), 745-753. doi:10.1016/j.jclinepi.2013.11.013

Boers, M., Tugwell, P., Felson, D. T., van Riel, P. L., Kirwan, J. R., Edmonds, J. P., Smolen, J. S., Khaltaev, N., & Muirden, K. D. (1994). World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl, 41*, 86-89.

Boers, M., Verhoeven, A. C., Markusse, H. M., van de Laar, M. A., Westhovens, R., van Denderen, J. C., van Zeben, D., Dijkmans, B. A., Peeters, A. J., Jacobs, P., van den Brink, H. R., Schouten, H. J., van der Heijde, D. M., Boonen, A., & van der Linden, S. (1997). Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *Lancet, 350*(9074), 309-318. doi:10.1016/s0140-6736(97)01300-7

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Thériault, G. (1983). A method to assess energy expenditure in children and adults. *Am J Clin Nutr, 37*(3), 461-467. doi:10.1093/ajcn/37.3.461

Brady, S. M., Fenton, S. A. M., Metsios, G. S., Bosworth, A., Duda, J. L., Kitas, G. D., & Veldhuijzen Van Zanten, J. J. C. S. (2021). Different types of physical activity are positively associated with indicators of mental health and psychological wellbeing in rheumatoid arthritis during COVID-19. *Rheumatology International, 41*(2), 335-344. doi:10.1007/s00296-020-04751-w

Brady, S. M., Veldhuijzen Van Zanten, J. J. C. S., Duda, J. L., Kitas, G. D., & Fenton, S. A. M. (2019). Physical Activity/Inactivity: Objective Measurement of. In M. Gellman (Ed.), *Encyclopedia of Behavioral Medicine* (pp. 1-4). New York, NY: Springer New York.

Brage, S., Wedderkopp, N., Franks, P. W., Andersen, L. B., & Froberg, K. (2003). Reexamination of validity and reliability of the CSA monitor in walking and running. *Medicine & Science in Sports & Exercise, 35*(8), 1447-1454. doi:10.1249/01.Mss.0000079078.62035.Ec

Brodin, N., Eurenius, E., Jensen, I., Nisell, R., Opava, C. H., Algebrandt, M., Almin, I., Andersson, B., Bertholds, G., Forsberg, C., Haglund, E., Holmen-Andersson, A. M., Hultman, A., Lennartsson, C., & Norman, E. (2008). Coaching patients with early rheumatoid arthritis to healthy physical activity: A multicenter, randomized, controlled study. *Arthritis Care & Research, 59*(3), 325-331.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet, 395*(10227), 912-920. doi:10.1016/s0140-6736(20)30460-8

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., Lambert, E., Leitzmann, M., Milton, K., Ortega, F. B., Ranasinghe, C., Stamatakis, E., Tiedemann, A., Troiano, R. P., van der Ploeg, H. P., Wari, V., & Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine, 54*(24), 1451. doi:10.1136/bjsports-2020-102955

Buman, M. P., Hekler, E. B., Haskell, W. L., Pruitt, L., Conway, T. L., Cain, K. L., Sallis, J. F., Saelens, B. E., Frank, L. D., & King, A. C. (2010). Objective light-intensity physical activity associations with rated health in older adults. *Am J Epidemiol, 172*(10), 1155-1165. doi:10.1093/aje/kwq249

Burckhardt, C. S. (1984). The use of the McGill Pain Questionnaire in assessing arthritis pain. *Pain, 19*(3), 305-314. doi:10.1016/0304-3959(84)90007-1

Burckhardt, C. S., & Anderson, K. L. (2003). The Quality of Life Scale (QOLS): reliability, validity, and utilization. *Health and Quality of Life Outcomes, 1*, 60. doi:10.1186/1477-7525-1-60

Calabro, M. A., Lee, J., Saint-Maurice, P. F., Yoo, H., & Welk, G. J. (2014). Validity of physical activity monitors for assessing lower intensity activity in adults. *International Journal of Behavioral Nutrition and Physical Activity, 11*(119), 1-9.

Carlson, J. A., Tuz-Zahra, F., Bellettiere, J., Ridgers, N. D., Steel, C., Bejarano, C., LaCroix, A. Z., Rosenberg, D. E., Greenwood-Hickman, M. A., Jankowska, M. M., & Natarajan, L. (2021). Validity of Two Awake Wear-Time Classification Algorithms for activPAL in Youth, Adults, and Older Adults. *Journal for the measurement of physical behaviour, 4*(2), 151-162. doi:10.1123/jmpb.2020-0045

Carriedo, A., Cecchini, J. A., Fernandez-Rio, J., & Mendez-Gimenez, A. (2020). COVID-19, Psychological Well-being and Physical Activity Levels in Older Adults During the Nationwide Lockdown in Spain. *Am J Geriatr Psychiatry*. doi:10.1016/j.jagp.2020.08.007

Carson, V., Wong, S. L., Winkler, E., Healy, G. N., Colley, R. C., & Tremblay, M. S. (2014). Patterns of sedentary time and cardiometabolic risk among Canadian adults. *Preventive Medicine, 65*, 23-27. doi:10.1016/j.ypmed.2014.04.005

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports (Washington, D.C. : 1974), 100*(2), 126-131.

Castañeda-Babarro, A., Arbillaga-Etxarri, A., Gutiérrez-Santamaría, B., & Coca, A. (2020). Physical Activity Change during COVID-19 Confinement. *International Journal of Environmental Research and Public Health, 17*(18), 6878.

Cathcart, S., Winefield, A. H., Rolan, P., & Lushington, K. (2009). Reliability of temporal summation and diffuse noxious inhibitory control. *Pain Research and Management, 14*(6), 433-438.

Chan, C. S., Slaughter, S. E., Jones, C. A., Ickert, C., & Wagg, A. S. (2017). Measuring Activity Performance of Older Adults Using the activPAL: A Rapid Review. *Healthcare (Basel, Switzerland), 5*(4). doi:10.3390/healthcare5040094

Chandarana, P. C., Eals, M., Steingart, A. B., Bellamy, N., & Allen, S. (1987). The Detection of Psychiatric Morbidity and Associated Factors in Patients with Rheumatoid Arthritis. *The Canadian Journal of Psychiatry, 32*(5), 356-361. doi:10.1177/070674378703200506

Chang, C.-L., Chiu, C.-M., Hung, S.-Y., Lee, S.-H., Lee, C.-S., Huang, C.-M., & Chou, C.-L. (2009). The relationship between quality of life and aerobic fitness in patients with rheumatoid arthritis. *Clinical Rheumatology, 28*(6), 685-691. doi:10.1007/s10067-009-1132-0

Chastin, S. F., Dontje, M. L., Skelton, D. A., Cukic, I., Shaw, R. J., Gill, J. M. R., Greig, C. A., Gale, C. R., Deary, I. J., Der, G., Dall, P. M., & Seniors, U. S. P. t. (2018). Systematic comparative validation of self-report measures of sedentary time against an objective measure of postural sitting (activPAL). *International Journal of Behavioral Nutrition and Physical Activity, 15*(1), 21. doi:10.1186/s12966-018-0652-x

Chastin, S. F., Egerton, T., Leask, C., & Stamatakis, E. (2015). Meta-analysis of the relationship between breaks in sedentary behavior and cardiometabolic health. *Obesity (Silver Spring), 23*(9), 1800-1810. doi:10.1002/oby.21180

Chastin, S. F. M., Fitzpatrick, N., Andrews, M., & DiCroce, N. (2014). Determinants of Sedentary Behavior, Motivation, Barriers and Strategies to Reduce Sitting Time in Older Women: A Qualitative Investigation. *International Journal of Environmental Research and Public Health, 11*(1), 773-791.

Chekroud, S. R., Gueorguieva, R., Zheutlin, A. B., & Paulus, M. (2018). Association between physical exercise and mental health in 1·2 million individuals in the USA between 2011 and 2015: a cross-sectional study. *The Lancet Psychiatry, 5*(9), 739-746. doi:10.1016/s2215-0366(18)30227-x

Chesterton, L. S., Sim, J., Wright, C. C., & Foster, N. E. (2007). Interrater reliability of algometry in measuring pressure pain thresholds in healthy humans, using multiple raters. *The clinical Journal of Pain, 23*(9), 760-766.

Chmelo, E., Nicklas, B., Davis, C., Miller, G. D., Legault, C., & Messier, S. (2013). Physical activity and physical function in older adults with knee osteoarthritis. *Journal of Physical Activity & Health, 10*(6), 777-783. doi:10.1123/jpah.10.6.777

Choi, L., Liu, Z., Matthews, C., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise, 43*(2), 357-364. doi:10.1249/MSS.0b013e3181ed61a3

Choi, L., Ward, S. C., Schnelle, J. F., & Buchowski, M. S. (2012). Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. *Medicine & Science in Sports & Exercise, 44*(10), 2009-2016. doi:10.1249/MSS.0b013e318258cb36

Christie, A., Jamtvedt, G., Dahm, K. T., Moe, R. H., Haavardsholm, E. A., & 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. doi:10.2522/ptj.20070039

Chung, S.-C., Um, B.-Y., & Kim, H.-S. (1992). Evaluation of pressure pain threshold in head and neck muscles by electronic algometer: intrarater and interrater reliability. *CRANIO®, 10*(1), 28-34.

Cimmino, M. A., Salvarani, C., Macchioni, P., Montecucco, C., Fossaluzza, V., Mascia, M. T., Punzi, L., Davoli, C., Filippini, D., & Numo, R. (2000). Extra-articular manifestations in 587 Italian patients with rheumatoid arthritis. *Rheumatology International, 19*(6), 213-217. doi:10.1007/pl00006853

Cleland, C., Ferguson, S., Ellis, G., & Hunter, R. (2018). Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. *BMC Medical Research Methodology, 18*(176). doi:10.1186/s12874-018-0642-3

Clevenger, K. A., Pfeiffer, K. A., & Montoye, A. H. K. (2020a). Cross-generational comparability of hip- and wrist-worn ActiGraph GT3X+, wGT3X-BT, and GT9X accelerometers during free-living in adults. *Journal of Sports Sciences, 38*(24), 2794-2802. doi:10.1080/02640414.2020.1801320

Clevenger, K. A., Pfeiffer, K. A., & Montoye, A. H. K. (2020b). Cross-Generational Comparability of Raw and Count-Based Metrics from ActiGraph GT9X and wGT3X-BT Accelerometers during Free-Living in Youth. *Measurement in Physical Education and Exercise Science, 24*(3), 194-204. doi:10.1080/1091367X.2020.1773827

Cojocaru, M., Cojocaru, I. M., Silosi, I., Vrabie, C. D., & Tanasescu, R. (2010). Extra-articular Manifestations in Rheumatoid Arthritis. *Maedica, 5*(4), 286-291.

Cooney, J. K., Law, R. J., Matschke, V., Lemmey, A. B., Moore, J. P., Ahmad, Y., Jones, J. G., Maddison, P., & Thom, J. M. (2011). Benefits of exercise in rheumatoid arthritis. *Journal of Aging Research, 2011*, 681640. doi:10.4061/2011/681640

Corrêa, J. B., Costa, L. O., de Oliveira, N. T., Sluka, K. A., & Liebano, R. E. (2015). Central sensitization and changes in conditioned pain modulation in people with chronic nonspecific low back pain: a case-control study. *Exp Brain Res, 233*(8), 2391-2399. doi:10.1007/s00221-015-4309-6

Courtney, C. A., Kavchak, A. E., Lowry, C. D., & O'Hearn, M. A. (2010). Interpreting joint pain: quantitative sensory testing in musculoskeletal management. *J Orthop Sports Phys Ther, 40*(12), 818-825. doi:10.2519/jospt.2010.3314

Covic, T., Pallant, J., Tennant, A., Cox, S., Emery, P., & Conaghan, P. G. (2009). Variability in depression prevalence in early rheumatoid arthritis: a comparison of the CES-D and HAD-D Scales. *Bmc Musculoskeletal Disorders, 10*. doi:10.1186/1471-2474-10-18

Cramp, F., Hewlett, S., Almeida, C., Kirwan, J. R., Choy, E. H., Chalder, T., Pollock, J., & Christensen, R. (2013). Non-pharmacological interventions for fatigue in rheumatoid arthritis. *Cochrane Database of Systematic Reviews*(8), Cd008322. doi:10.1002/14651858.CD008322.pub2

Cramp, F., Thomas, R., Haase, A. M., Domaille, M., Manns, S., Swales, C., Hurfurt, J., & Walsh, N. E. (2020). Promoting engagement in physical activity in early rheumatoid arthritis: A proof-of-concept intervention study. *Musculoskeletal Care, 14*.

Crouter, S. E., DellaValle, D. M., Haas, J. D., Frongillo, E. A., & Bassett, D. R. (2013). Validity of ActiGraph 2-regression model, Matthews cut-points, and NHANES cut-points for assessing free-living physical activity. *Journal of Physical Activity & Health, 10*(4), 504-514. doi:10.1123/jpah.10.4.504

da Rocha Castelar Pinheiro, G., Khandker, R. K., Sato, R., Rose, A., & Piercy, J. (2013). Impact of rheumatoid arthritis on quality of life, work productivity and resource utilisation: an observational, cross-sectional study in Brazil. *Clinical & Experimental Rheumatology, 31*, 334-340.

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., Breedveld, F. C., Vliet Vlieland, T. P., & Hazes, J. M. (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, 48*(9), 2415-2424.

de Rezende, L. F., Rey-López, J. P., Matsudo, V. K., & do Carmo Luiz, O. (2014). Sedentary behavior and health outcomes among older adults: a systematic review. *Bmc Public Health, 14*, 333. doi:10.1186/1471-2458-14-333

Demmelmaier, I., Asenlof, P., Bergman, P., Nordgren, B., & Opava, C. H. (2017). Pain rather than self-reported sedentary time explains variation in perceived health and activity limitation in persons with rheumatoid arthritis: a cross sectional study in Sweden. *Rheumatology International, 37*(6), 923-930. doi:10.1007/s00296-016-3641-x

Diaz, K. M., Krupka, D. J., Chang, M. J., Kronish, I. M., Moise, N., Goldsmith, J., & Schwartz, J. E. (2018). Wrist-based cut-points for moderate- and vigorous-intensity physical activity for the Actical accelerometer in adults. *Journal of Sports Sciences, 36*(2), 206-212. doi:10.1080/02640414.2017.1293279

DiPietro, L., Jin, Y., Talegawkar, S., & Matthews, C. E. (2018). The Joint Associations of Sedentary Time and Physical Activity With Mobility Disability in Older People: The NIH-AARP Diet and Health Study. *J Gerontol A Biol Sci Med Sci, 73*(4), 532-538. doi:10.1093/gerona/glx122

Doherty, A. R., Kelly, P., Kerr, J., Marshall, S., Oliver, M., Badland, H., Hamilton, A., & Foster, C. (2013). Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. *International Journal of Behavioral Nutrition and Physical Activity, 10*(1), 22. doi:10.1186/1479-5868-10-22

Dowd, K. P., Harrington, D. M., & Donnelly, A. E. (2012). Criterion and concurrent validity of the activPAL™ professional physical activity monitor in adolescent females. *Plos One, 7*(10), e47633. doi:10.1371/journal.pone.0047633

Dowd, K. P., Szeklicki, R., Minetto, M. A., Murphy, M. H., Polito, A., Ghigo, E., van der Ploeg, H., Ekelund, U., Maciaszek, J., Stemplewski, R., Tomczak, M., & Donnelly, A. E. (2018). A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. *International Journal of Behavioral Nutrition and Physical Activity, 15*(1), 15. doi:10.1186/s12966-017-0636-2

Drossaers-Bakker, K. W., de Buck, M., van Zeben, D., Zwinderman, A. H., Breedveld, F. C., & Hazes, J. M. (1999). Long-term course and outcome of functional capacity in rheumatoid arthritis: the effect of disease activity and radiologic damage over time. *Arthritis & Rheumatism, 42*(9), 1854-1860. doi:10.1002/1529-0131(199909)42:9<1854::Aid-anr9>3.0.Co;2-f

du Prel, J.-B., Hommel, G., Röhrig, B., & Blettner, M. (2009). Confidence interval or p-value?: part 4 of a series on evaluation of scientific publications. *Deutsches Arzteblatt international, 106*(19), 335-339. doi:10.3238/arztebl.2009.0335

Duncan, M. J., Eyre, E. L. J., Cox, V., Roscoe, C. M. P., Faghy, M. A., Tallis, J., & Dobell, A. (2020). Cross-validation of Actigraph derived accelerometer cut-points for assessment of sedentary behaviour and physical activity in children aged 8-11 years. *Acta Paediatrica, 109*(9), 1825-1830. doi:10.1111/apa.15189

Duncan, S., Stewart, T., Bo Schneller, M., Godbole, S., Cain, K., & Kerr, J. (2018). Convergent validity of ActiGraph and Actical accelerometers for estimating physical activity in adults. *Plos One, 13*(6), e0198587. doi:10.1371/journal.pone.0198587

Duvivier, B. M., Schaper, N. C., Bremers, M. A., van Crombrugge, G., Menheere, P. P., Kars, M., & Savelberg, H. H. (2013). Minimal intensity physical activity (standing and walking) of longer duration improves insulin action and plasma lipids more than shorter periods of moderate to vigorous exercise (cycling) in sedentary subjects when energy expenditure is comparable. *Plos One, 8*(2), e55542. doi:10.1371/journal.pone.0055542

Dwyer, T., Ponsonby, A. L., Ukoumunne, O. C., Pezic, A., Venn, A., Dunstan, D., Barr, E., Blair, S., Cochrane, J., Zimmet, P., & Shaw, J. (2011). Association of change in daily step count over five years with insulin sensitivity and adiposity: population based cohort study. *BMJ (Clinical research ed.), 342*, c7249. doi:10.1136/bmj.c7249

Edwards, R. R., Wasan, A. D., Bingham, C. O., 3rd, Bathon, J., Haythornthwaite, J. A., Smith, M. T., & Page, G. G. (2009). Enhanced reactivity to pain in patients with rheumatoid arthritis. *Arthritis Research & Therapy, 11*(3), R61-R61. doi:10.1186/ar2684

Edwardson, C. L., Winkler, E. A. H., Bodicoat, D. H., Yates, T., Davies, M. J., Dunstan, D. W., & Healy, G. N. (2017). Considerations when using the activPAL monitor in field-based research with adult populations. *Journal of Sport and Health Science, 6*(2), 162-178. doi:10.1016/j.jshs.2016.02.002

Ekelund, U., Steene-Johannessen, J., Brown, W. J., Fagerland, M. W., Owen, N., Powell, K. E., Bauman, A., & Lee, I. M. (2016). Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet, 388*(10051), 1302-1310. doi:10.1016/s0140-6736(16)30370-1

Ekelund, U., Tarp, J., Fagerland, M. W., Johannessen, J. S., Hansen, B. H., Jefferis, B. J., Whincup, P. H., Diaz, K. M., Hooker, S., Howard, V. J., Chernofsky, A., Larson, M. G., Spartano, N., Vasan, R. S., Dohrn, I.-M., Hagströmer, M., Edwardson, C., Yates, T., Shiroma, E. J., Dempsey, P., Wijndaele, K., Anderssen, S. A., & Lee, I. M. (2020). Joint associations of accelero-meter measured physical activity and sedentary time with all-cause mortality: a harmonised meta-analysis in more than 44 000 middle-aged and older individuals. *British Journal of Sports Medicine, 54*(24), 1499-1506. doi:10.1136/bjsports-2020-103270

Ekelund, U., Tarp, J., Steene-Johannessen, J., Hansen, B. H., Jefferis, B., Fagerland, M. W., Whincup, P., Diaz, K. M., Hooker, S. P., Chernofsky, A., Larson, M. G., Spartano, N., Vasan, R. S., Dohrn, I.-M., Hagströmer, M., Edwardson, C., Yates, T., Shiroma, E., Anderssen, S. A., & Lee, I. M. (2019). Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ (Clinical research ed.), 366*, l4570. doi:10.1136/bmj.l4570

Ellis, K., Kerr, J., Godbole, S., Staudenmayer, J., & Lanckriet, G. (2016). Hip and Wrist Accelerometer Algorithms for Free-Living Behavior Classification. *Medicine and Science in Sports and Exercise, 48*(5), 933-940. doi:10.1249/MSS.0000000000000840

Englbrecht, M., Kruckow, M., Araujo, E., Rech, J., & Schett, G. (2013). The interaction of physical function and emotional well-being in rheumatoid arthritis--what is the impact on disease activity and coping? *Seminars in Arthritis and Rheumatism, 42*(5), 482-491. doi:10.1016/j.semarthrit.2012.09.003

Evenson, K. R., Sotres-Alvarez, D., Deng, Y. U., Marshall, S. J., Isasi, C. R., Esliger, D. W., & Davis, S. (2015). Accelerometer adherence and performance in a cohort study of US Hispanic adults.

Medicine & Science in Sports & Exercise, 47(4), 725-734. doi:10.1249/MSS.0000000000000478

Ewald, B., Attia, J., & McElduff, P. (2014). How many steps are enough? Dose-response curves for pedometer steps and multiple health markers in a community-based sample of older Australians. *Journal of Physical Activity & Health, 11*(3), 509-518. doi:10.1123/jpah.2012-0091

Fabio Antonaci, M. (1998). Pressure algometry in healthy subjects: inter-examiner variability. *Scand J Rehab Med, 30*(3), 8.

Feito, Y., Bassett, D. R., & Thompson, D. L. (2012). Evaluation of activity monitors in controlled and free-living environments. *Medicine & Science in Sports & Exercise, 44*(4), 733-741. doi:10.1249/MSS.0b013e3182351913

Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical Inference for Coefficient Alpha. *Applied Psychological Measurement, 11*(1), 93-103. doi:10.1177/014662168701100107

Feldthusen, C., Dean, E., Forsblad-d'Elia, H., & Mannerkorpi, K. (2016). Effects of person-centered physical therapy on fatigue-related variables in persons with rheumatoid arthritis: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation, 97*(1), 26-36.

Feng, C., Wang, H., Lu, N., Chen, T., He, H., Lu, Y., & Tu, X. M. (2014). Log-transformation and its implications for data analysis. *Shanghai Arch Psychiatry, 26*(2), 105-109. doi:10.3969/j.issn.1002-0829.2014.02.009

Fenton, S. A. M., Duda, J. L., Veldhuijzen Van Zanten, J. J. C. S., Metsios, G. S., & Kitas, G. D. (2020a). Theory-informed interventions to promote physical activity and reduce sedentary behaviour in rheumatoid arthritis: A critical review of the literature. *Mediterranean Journal of Rheumatology, 31*(1), 19-41.

Fenton, S. A. M., & Kitas, G. D. (2016). Rheumatoid arthritis: Sedentary behaviour in RA - a new research agenda. *Nat Rev Rheumatol, 12*(12), 698-700. doi:10.1038/nrrheum.2016.179

Fenton, S. A. M., Ntoumanis, N., Duda, J. L., Metsios, G. S., Rouse, P. C., Yu, C., Kitas, G. D., & Veldhuijzen Van Zanten, J. J. C. S. (2020b). Diurnal patterns of sedentary time in rheumatoid arthritis: associations with cardiovascular disease risk. *RMD Open, 6*(2), e001216. doi:10.1136/rmdopen-2020-001216

Fenton, S. A. M., Veldhuijzen Van Zanten, J. J. C. S., Duda, J. L., Metsios, G. S., & Kitas, G. D. (2018a). Sedentary behaviour in rheumatoid arthritis: definition, measurement and implications for health. *Rheumatology (Oxford), 57*(2), 213-226. doi:10.1093/rheumatology/kex053

Fenton, S. A. M., Veldhuijzen Van Zanten, J. J. C. S., Kitas, G. D., Duda, J. L., Rouse, P. C., Yu, C. A., & Metsios, G. S. (2017). Sedentary behaviour is associated with increased long-term cardiovascular risk in patients with rheumatoid arthritis independently of moderate-to-vigorous physical activity. *Bmc Musculoskeletal Disorders, 18*(1). doi:10.1186/s12891-017-1473-9

Fenton, S. A. M., Veldhuijzen Van Zanten, J. J. C. S., Metsios, G. S., Rouse, P. C., Yu, C. A., Kitas, G. D., & Duda, J. L. (2018b). Autonomy support, light physical activity and psychological well-being in Rheumatoid Arthritis: A cross-sectional study. *Mental Health and Physical Activity, 14*, 11-18. doi:10.1016/j.mhpa.2017.12.002

Fernandes, C., Pidal-Miranda, M., Samartin-Veiga, N., & Carrillo-de-la-Peña, M. T. (2019). Conditioned pain modulation as a biomarker of chronic pain: a systematic review of its concurrent validity. *Pain, 160*(12), 2679-2690.

Fiest, K. M., Hitchon, C. A., Bernstein, C. N., Peschken, C. A., Walker, J. R., Graff, L. A., Zarychanski, R., Abou-Setta, A., Patten, S. B., Sareen, J., Bolton, J., & Marrie, R. A. (2017). Systematic Review and Meta-analysis of Interventions for Depression and Anxiety in Persons With Rheumatoid Arthritis. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases, 23*(8), 425-434. doi:10.1097/RHU.0000000000000489

Fingleton, C., Smart, K., Moloney, N., Fullen, B. M., & Doody, C. (2015). Pain sensitization in people with knee osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis and Cartilage, 23*(7), 1043-1056.

Foundation, B. H. (2017). *Physical Inactivity and Sedentary Behaviour Report 2017*. Retrieved from United Kingdom:

Fransen, J., Langenegger, T., Michel, B. A., Stucki, G., & Arthritis, f. t. m. o. t. S. C. Q. M. i. R. (2000). Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index. *Rheumatology, 39*(3), 321-327. doi:10.1093/rheumatology/39.3.321

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine & Science in Sports & Exercise, 30*(5), 777-781. doi:10.1097/00005768-199805000-00021

Fries, J. F., Spitz, P., Kraines, R. G., & Holman, H. R. (1980). Measurement of patient outcome in arthritis. *Arthritis & Rheumatism, 23*(2), 137-145. doi:10.1002/art.1780230202

Gabriel, S. E., Crowson, C. S., & O'Fallon, W. M. (1999). Comorbidity in arthritis. *Journal of Rheumatology, 26*(11), 2475-2479.

Garner, S., Fenton, T., Martin, L., Creaser, C., Johns, C., & Barnabe, C. (2018). Personalized diet and exercise recommendations in early rheumatoid arthritis: a feasibility trial. *Musculoskeletal Care, 16*(1), 167-172.

Geber, C., Klein, T., Azad, S., Birklein, F., Gierthmühlen, J., Huge, V., Lauchart, M., Nitzsche, D., Stengel, M., & Valet, M. (2011). Test–retest and interobserver reliability of quantitative sensory testing according to the protocol of the German Research Network on Neuropathic Pain (DFNS): a multi-centre study. *PAIN®, 152*(3), 548-556.

Geletka, B. J., O'Hearn, M. A., & Courtney, C. A. (2012). Quantitative sensory testing changes in the successful management of chronic low back pain. *The Journal of manual & manipulative therapy, 20*(1), 16-22. doi:10.1179/2042618611Y.0000000014

Georgopoulos, V., Akin-Akinyosoye, K., Zhang, W., McWilliams, D. F., Hendrick, P., & Walsh, D. A. (2019). Quantitative sensory testing and predicting outcomes for musculoskeletal pain, disability, and negative affect: a systematic review and meta-analysis. *Pain, 160*(9), 1920-1932. doi:10.1097/j.pain.0000000000001590

Gerecz-Simon, E. M., Trunks, E. R., Heale, J. A., Kean, W. F., & Buchanan, W. W. (1989). Measurement of pain threshold in patients with rheumatoid arthritis, osteoarthritis, ankylosing spondylitis, and healthy controls. *Clinical Rheumatology, 8*(4), 467-474. doi:10.1007/BF02032098

Gettings, L. (2010). Psychological well-being in rheumatoid arthritis: a review of the literature. *Musculoskeletal Care, 8*(2), 99-106. doi:https://doi.org/10.1002/msc.171

Gierach, G. L., Chang, S. C., Brinton, L. A., Lacey, J. V., Jr., Hollenbeck, A. R., Schatzkin, A., & Leitzmann, M. F. (2009). Physical activity, sedentary behavior, and endometrial cancer risk in the NIH-AARP Diet and Health Study. *International Journal of Cancer, 124*(9), 2139-2147. doi:10.1002/ijc.24059

Gilbert, A. L., Lee, J., Ehrlich-Jones, L., Semanik, P. A., Song, J., Pellegrini, C. A., Pinto, P. D., Dunlop, D. D., & Chang, R. W. (2018). A randomized trial of a motivational interviewing intervention to increase lifestyle physical activity and improve self-reported function in adults with arthritis. *Seminars in Arthritis and Rheumatism, 47*(5), 732-740.

Gilbert, A. L., Lee, J., Ma, M., Semanik, P., DiPietro, L., Dunlop, D., & Chang, R. (2016). Comparison of Subjective and Objective Measures of Sedentary Behavior Using the Yale Physical Activity Survey and Accelerometry in Patients With Rheumatoid Arthritis. *Journal of Physical Activity & Health, 13*(4), 371-376. doi:10.1123/jpah.2015-0176

Giles, J. T., Bartlett, S. J., Andersen, R. E., Fontaine, K. R., & Bathon, J. M. (2008). Association of body composition with disability in rheumatoid arthritis: impact of appendicular fat and lean tissue mass. *Arthritis & Rheumatism, 59*(10), 1407-1415. doi:10.1002/art.24109

Giraudet-le Quintrec, J. S., Mayoux-Benhamou, A., Ravaud, P., Champion, K., Dernis, E., Zerkak, D., Ouslimani, A., Courpied, J. P., Revel, M., Kahan, A., & Dougados, M. (2007). Effect of a collective educational program for patients with rheumatoid arthritis: a prospective 12-month randomized controlled trial. *Journal of Rheumatology, 34*(8), 1684-1691.

Gooberman-Hill, R., Woolhead, G., MacKichan, F., Ayis, S., Williams, S., & Dieppe, P. (2007). Assessing chronic joint pain: Lessons from a focus group study. *Arthritis Care & Research, 57*(4), 666-671. doi:10.1002/art.22681

Gossec, L. (2018). Chapter 5 - Monitoring of Disease and Treatment of Patients With Rheumatic Disease. In F. Atzeni, I. F. Masala, D. Aletaha, M. Lee, & X. Baraliakos (Eds.), *Handbook of Systemic Autoimmune Diseases* (Vol. 15, pp. 97-125): Elsevier.

Grant, P. M., Granat, M. H., Thow, M. K., & Maclaren, W. M. (2010). Analyzing free-living physical activity of older adults in different environments using body-worn activity monitors. *Journal of Ageing & Physical Activity, 18*(2), 171-184. doi:10.1123/japa.18.2.171

Graven-Nielson, T., Vaegter, H. B., Finocchietti, S., Handberg, G., & Arendt-Nielsen, L. (2015). Assessment of musculoskeletal pain sensitivity and temporal summation by cuff pressure algometry: a reliability study. *Pain, 156*(11), 2193-2202.

Green, A. N., McGrath, R., Martinez, V., Taylor, K., Paul, D. R., & Vella, C. A. (2014). Associations of objectively measured sedentary behavior, light activity, and markers of cardiometabolic health in young women. *Eur J Appl Physiol, 114*(5), 907-919. doi:10.1007/s00421-014-2822-0

Greene, B. L., Haldeman, G. F., Kaminski, A., Neal, K., Lim, S. S., & Conn, D. L. (2006). Factors affecting physical activity behavior in urban adults with arthritis who are predominantly African-American and female. *Physical Therapy, 86*(4), 510-519.

Greenwood-Hickman, M. A., Renz, A., & Rosenberg, D. E. (2015). Motivators and Barriers to Reducing Sedentary Behavior Among Overweight and Obese Older Adults. *The Gerontologist, 56*(4), 660-668. doi:10.1093/geront/gnu163

Guidelines, A. C. o. R. S. o. R. A. (2002). Guidelines for the management of rheumatoid arthritis: 2002 Update. *Arthritis & Rheumatism, 46*(2), 328-346. doi:10.1002/art.10148

Gulati, M., Farah, Z., & Mouyis, M. (2018). Clinical features of rheumatoid arthritis. *Medicine, 46*(4), 211-215. doi:10.1016/j.mpmed.2018.01.008

Hagstromer, M., Ainsworth, B. E., Oja, P., & Sjostrom, M. (2010). Comparison of a subjective and an objective measure of physical activity in a population sample. *Journal of Physical Activity & Health, 7*(4), 541-550. doi:10.1123/jpah.7.4.541

Haider, S., Sedlak, M., Kapan, A., Grabovac, I., Lamprecht, T., Erlacher, L., Quittan, M., Fenzl, K. H., & Dorner, T. E. (2020). Factors Associated with Objectively Measured Physical Activity in Patients with Seropositive Rheumatoid Arthritis. *International Journal of Environmental Research and Public Health, 17*(23), 9008. doi:10.3390/ijerph17239008

Hajna, S., White, T., Brage, S., van Sluijs, E. M. F., Westgate, K., Jones, A. P., Luben, R., Khaw, K.-T., Wareham, N. J., & Griffin, S. J. (2018). Descriptive epidemiology of changes in objectively measured sedentary behaviour and physical activity: six-year follow-up of the EPIC-Norfolk cohort. *International Journal of Behavioral Nutrition and Physical Activity, 15*(1), 122. doi:10.1186/s12966-018-0746-5

Hakkinen, A., Sokka, T., Kotaniemi, A., & 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. doi:10.1002/1529-0131(200103)44:3<515::AID-ANR98>3.0.CO;2-5

Hammam, N., Ezeugwu, V. E., Rumsey, D. G., Manns, P. J., & Pritchard-Wiart, L. (2019). Physical activity, sedentary behavior, and long-term cardiovascular risk in individuals with rheumatoid arthritis. *Phys Sportsmed, 47*(4), 463-470. doi:10.1080/00913847.2019.1623995

Harper, C., Satchell, L., Fido, D., & Latzman, R. (2020). Functional fear predicts public health compliance in the COVID-19 pandemic. *International Journal of Mental Health Addiction*. doi: 10.1007/s11469-020-00281-5

Harvey, J. A., Chastin, S. F., & Skelton, D. A. (2015). How Sedentary are Older People? A Systematic Review of the Amount of Sedentary Behavior. *Journal of Ageing & Physical Activity, 23*(3), 471-487. doi:10.1123/japa.2014-0164

Hattori, Y., Katayama, M., Kida, D., & Kaneko, A. (2018). Hospital Anxiety and Depression Scale Score Is an Independent Factor Associated With the EuroQoL 5-Dimensional Descriptive System in Patients With Rheumatoid Arthritis. *Journal of Clinical Rheumatology, 24*(6), 308-312. doi:10.1097/rhu.0000000000000735

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care & Research, 63 Suppl 11*, S240-252. doi:10.1002/acr.20543

Hayes, A. F. (2012). PROCESS: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling.

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence Based Nursing, 18*(3), 66-67. doi:10.1136/eb-2015-102129

Healy, G. N., Clark, B. K., Winkler, E. A., Gardiner, P. A., Brown, W. J., & Matthews, C. E. (2011). Measurement of adults' sedentary time in population-based studies. *Am J Prev Med, 41*(2), 216-227. doi:10.1016/j.amepre.2011.05.005

Heiberg, T., & Kvien, T. K. (2002). Preferences for improved health examined in 1,024 patients with rheumatoid arthritis: pain has highest priority. *Arthritis & Rheumatism, 47*(4), 391-397. doi:10.1002/art.10515

Helmrich, S. P., Ragland, D. R., & Paffenbarger, R. S., Jr. (1994). Prevention of non-insulin-dependent diabetes mellitus with physical activity. *Medicine and Science in Sports and Exercise, 26*(7), 824-830. doi:10.1249/00005768-199407000-00003

Hernández-Hernández, M. V., & Díaz-González, F. (2017). Role of physical activity in the management and assessment of rheumatoid arthritis patients. *Reumatologia Clinica, 13*(4), 214-220. doi:10.1016/j.reumae.2016.04.014

Hernández-Hernández, V., Ferraz-Amaro, I., & Díaz-González, F. (2014). Influence of disease activity on the physical activity of rheumatoid arthritis patients. *Rheumatology (Oxford), 53*(4), 722-731. doi:10.1093/rheumatology/ket422

Hewlett, S., Ambler, N., Almeida, C., Cliss, A., Hammond, A., Kitchen, K., Knops, B., Pope, D., Spears, M., Swinkels, A., & Pollock, J. (2011). Self-management of fatigue in rheumatoid arthritis: a randomised controlled trial of group cognitive-behavioural therapy. *Ann Rheum Dis, 70*(6), 1060-1067. doi:10.1136/ard.2010.144691

Hewlett, S., Chalder, T., Choy, E., Cramp, F., Davis, B., Dures, E., Nicholls, C., & Kirwan, J. (2010). Fatigue in rheumatoid arthritis: time for a conceptual model. *Rheumatology, 50*(6), 1004-1006. doi:10.1093/rheumatology/keq282

Hewlett, S., Hehir, M., & Kirwan, J. R. (2007). Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. *Arthritis & Rheumatism, 57*(3), 429-439. doi:10.1002/art.22611

Hibbing, P. R., Bassett, D. R., & Crouter, S. E. (2020). Modifying Accelerometer Cut-Points Affects Criterion Validity in Simulated Free-Living for Adolescents and Adults. *Research Quarterly for Exercise and Sport, 91*(3), 514-524. doi:10.1080/02701367.2019.1688227

Higgins, J. P., Savović, J., Page, M. J., Elbers, R. G., & Sterne, J. A. C. (2021a). Chapter 8: Assessing risk of bias in a randomized trial. In J. P. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, P.

M.J., & V. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021)*. Chichester (UK): Cochrane.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. (2021b). *Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021)*. Chichester (UK): John Wiley & Sons.

Higgins, J. P., White, I. R., & Anzures-Cabrera, J. (2008). Meta-analysis of skewed data: combining results reported on log-transformed or raw scales. *Stat Med, 27*(29), 6072-6092. doi:10.1002/sim.3427

Hildebrand, M., VT, V. A. N. H., Hansen, B. H., & Ekelund, U. (2014). Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Medicine & Science in Sports & Exercise, 46*(9), 1816-1824. doi:10.1249/mss.0000000000000289

Ho, R. C., Fu, E. H., Chua, A. N., Cheak, A. A., & Mak, A. (2011). Clinical and psychosocial factors associated with depression and anxiety in Singaporean patients with rheumatoid arthritis. *International Journal of Rheumatic Diseases, 14*(1), 37-47. doi:10.1111/j.1756-185X.2010.01591.x

Hogan, T. P., Benjamin, A., & Brezinski, K. L. (2000). Reliability Methods: A Note on the Frequency of Use of Various Types. *Educational and Psychological Measurement, 60*(4), 523-531. doi:10.1177/00131640021970691

Hogeweg, J. A., Kuis, W., Oostendorp, R. A. B., & Helders, P. J. M. (1996). The Influence of Site of Stimulation, Age, and Gender on Pain Threshold in Healthy Children. *Physical Therapy, 76*(12), 1331-1339. doi:10.1093/ptj/76.12.1331

Hsieh, P. H., Wu, O., Geue, C., McIntosh, E., McInnes, I. B., & Siebert, S. (2020). Economic burden of rheumatoid arthritis: a systematic review of literature in biologic era. *Ann Rheum Dis, 79*(6), 771-777. doi:10.1136/annrheumdis-2019-216243

Huang, Y., & Zhao, N. (2020). Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: a web-based cross-sectional survey. *Psychiatry Res, 288*. doi: 10.1016/j.psychres.2020.112954

Huffman, K. M., Pieper, C., Hall, K. S., & St Clair, E. W. (2014). Self-efficacy for Exercise, More Than Disease-Related Factors, Is Associated With Objectively Assessed Exercise Time and Sedentary Behaviour in Rheumatoid Arthritis. *Scandinavian Journal of Rheumatology, 44*(2), 106-110. doi:10.3109/03009742.2014.931456

Hurkmans, E., Van Der Giesen, F. J., Vlieland, T. P. M. V., Schoones, J., & Van Den Ende, E. C. H. M. (2009). Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. *Cochrane Database of Systematic Reviews*(4). doi:10.1002/14651858.CD006853.pub2

Hurst, N. P., Kind, P., Ruta, D., Hunter, M., & Stubbings, A. (1997). Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EuroQol (EQ-5D). *British Journal of Rheumatology, 36*(5), 551-559. doi:10.1093/rheumatology/36.5.551

Imamura, M., Chen, J., Matsubayashi, S. R., Targino, R. A., Alfieri, F. M., Bueno, D. K., & Hsing, W. T. (2013). Changes in pressure pain threshold in patients with chronic nonspecific low back pain. *Spine, 38*(24), 2098-2107. doi:10.1097/01.brs.0000435027.50317.d7

Iversen, M. D., Frits, M., von Heideken, J., Cui, J., Weinblatt, M., & Shadick, N. A. (2017). Physical Activity and Correlates of Physical Activity Participation Over Three Years in Adults With Rheumatoid Arthritis. *Arthritis Care & Research, 69*(10), 1535-1545. doi:10.1002/acr.23156

Joharatnam, N., McWilliams, D. F., Wilson, D., Wheeler, M., Pande, I., & Walsh, D. A. (2015). A cross-sectional study of pain sensitivity, disease-activity assessment, mental health, and fibromyalgia status in rheumatoid arthritis. *Arthritis Research & Therapy, 17*, 11. doi:10.1186/s13075-015-0525-5

Jurakić, D., Pedišić, Ž., & Greblo, Z. (2010). Physical activity in different domains and health-related quality of life: a population-based study. *Quality of Life Research, 19*(9), 1303-1309. doi:10.1007/s11136-010-9705-6

Kanecki, K., Tyszko, P., Wisłowska, M., & Łyczkowska-Piotrowska, J. (2013). Preliminary report on a study of health-related quality of life in patients with rheumatoid arthritis. *Rheumatology International, 33*(2), 429-434. doi:10.1007/s00296-012-2421-5

Katz, P. (2017a). Causes and consequences of fatigue in rheumatoid arthritis. *Curr Opin Rheumatol, 29*(3), 269-276. doi:10.1097/bor.0000000000000376

Katz, P. (2017b). Fatigue in Rheumatoid Arthritis. *Curr Rheumatol Rep, 19*(5), 25. doi:10.1007/s11926-017-0649-5

Katz, P., Margaretten, M., Gregorich, S., & Trupin, L. (2018). Physical Activity to Reduce Fatigue in Rheumatoid Arthritis: A Randomized Controlled Trial. *Arthritis Care & Research, 70*(1), 1-10. doi:10.1002/acr.23230

Kelley, G. A., Kelley, K. S., & Hootman, J. M. (2015). Effects of exercise on depression in adults with arthritis: A systematic review with meta-analysis of randomized controlled trials. *Arthritis Research & Therapy, 17*(1). doi:10.1186/s13075-015-0533-5

Kelly, L. A., McMillan, D. G., Anderson, A., Fippinger, M., Fillerup, G., & Rider, J. (2013). Validity of actigraphs uniaxial and triaxial accelerometers for assessment of physical activity in adults in laboratory conditions. *BMC Med Phys, 13*(1), 5. doi:10.1186/1756-6649-13-5

Kennedy, D., Kemp, H., Ridout, D., Yarnitsky, D., & Rice, A. (2016). Reliability of conditioned pain modulation: a systematic review. *Pain, 157*(11), 2410-2419.

Khoja, S. S., Almeida, G. J., Chester Wasko, M., Terhorst, L., & Piva, S. R. (2016). Association of Light-Intensity Physical Activity with Lower Cardiovascular Disease Risk Burden in Rheumatoid Arthritis. *Arthritis Care & Research, 68*(4), 424-431. doi:10.1002/acr.22711

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy, 65*(23), 2276-2284. doi:10.2146/ajhp070364

Kirwan, J. R. (1995). The effect of glucocorticoids on joint destruction in rheumatoid arthritis. The Arthritis and Rheumatism Council Low-Dose Glucocorticoid Study Group. *N Engl J Med, 333*(3), 142-146. doi:10.1056/nejm199507203330302

Kirwan, J. R., Minnock, P., Adebajo, A., Bresnihan, B., Choy, E., De Wit, M., Hazes, M., Richards, P., Saag, K., & Suarez-Almazor, M. (2007). Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *Journal of Rheumatology, 34*(5), 1174-1177.

Kirwan, J. R., & Reeback, J. S. (1986). Stanford Health Assessment Questionnaire Modified to Assess Disability in British Patients With Rheumatoid Arthritis. *British Journal of Rheumatology, 25*(2), 206-209. doi:10.1093/rheumatology/25.2.206

Knittle, K., De Gucht, V., Hurkmans, E., Peeters, A., Ronday, K., Maes, S., & Vlieland, T. V. (2015). Targeting motivation and self-regulation to increase physical activity among patients with rheumatoid arthritis: a randomised controlled trial. *Clinical Rheumatology, 34*(2), 231-238.

Kong, J.-T., Johnson, K. A., Balise, R. R., & Mackey, S. (2013). Test-retest reliability of thermal temporal summation using an individualized protocol. *The Journal of Pain, 14*(1), 79-88.

Konttinen, Y. T., Honkanen, V. E., Grönblad, M., Keinonen, M., Santavirta, N., & Santavirta, S. (1992). The relation of extraarticular tenderness to inflammatory joint disease and personality in patients with rheumatoid arthritis. *Journal of Rheumatology, 19*(6), 851-855.

Kosek, E., Ekholm, J., & Nordemar, R. (1993). A comparison of pressure pain thresholds in different tissues and body regions. Long-term reliability of pressure algometry in healthy volunteers. *Scand J Rehabil Med, 25*(3), 117-124.

Koster, A., Shiroma, E. J., Caserotti, P., Matthews, C., Chen, K. Y., Glynn, N. W., & Harris, T. B. (2016). Comparison of Sedentary Estimates between activPAL and Hip- and Wrist-Worn ActiGraph.

Medicine & Science in Sports & Exercise, 48(8), 1514-1522. doi:10.1249/mss.0000000000000924

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. S. (2011). Validation of wearable monitors for assessing sedentary behavior. *Medicine & Science in Sports & Exercise, 43*(8), 1561-1567. doi:10.1249/MSS.0b013e31820ce174

Kozey-Keadle, S., Shiroma, E. J., Freedson, P. S., & Lee, I. M. (2014). Impact of accelerometer data processing decisions on the sample size, wear time and physical activity level of a large cohort study. *Bmc Public Health, 14*, 1210. doi:10.1186/1471-2458-14-1210

Kramer, H. R., Fontaine, K. R., Bathon, J. M., & Giles, J. T. (2012). Muscle density in rheumatoid arthritis: associations with disease features and functional outcomes. *Arthritis & Rheumatism, 64*(8), 2438-2450. doi:10.1002/art.34464

Kuster, R. P., Hagströmer, M., Baumgartner, D., & Grooten, W. J. A. (2021). Concurrent and discriminant validity of ActiGraph waist and wrist cut-points to measure sedentary behaviour, activity level, and posture in office work. *Bmc Public Health, 21*(1), 345-345. doi:10.1186/s12889-021-10387-7

Lai, J., Ma, S., Wang, Y., Cai, Z., Hu, J., Wei, N., Wu, J., Du, H., Chen, T., Li, R., Tan, H., Kang, L., Yao, L., Huang, M., Wang, H., Wang, G., Liu, Z., & Hu, S. (2020). Factors Associated With Mental Health Outcomes Among Health Care Workers Exposed to Coronavirus Disease 2019. *JAMA Netw Open, 3*(3), e203976. doi:10.1001/jamanetworkopen.2020.3976

Landewé, R. B., Boers, M., Verhoeven, A. C., Westhovens, R., van de Laar, M. A., Markusse, H. M., van Denderen, J. C., Westedt, M. L., Peeters, A. J., Dijkmans, B. A., Jacobs, P., Boonen, A., van der Heijde, D. M., & van der Linden, S. (2002). COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis & Rheumatism, 46*(2), 347-356. doi:10.1002/art.10083

Lange, E., Gjertsson, I., & Mannerkorpi, K. (2020). Long-time follow up of physical activity level among older adults with rheumatoid arthritis. *European Review of Aging and Physical Activity, 17*(1). doi:10.1186/s11556-020-00242-w

Larkin, L., & Kennedy, N. (2014). Correlates of physical activity in adults with rheumatoid arthritis: a systematic review. *Journal of Physical Activity & Health, 11*(6), 1248-1261. doi:10.1123/jpah.2012-0194

Larkin, L., Kennedy, N., Fraser, A., & Gallagher, S. (2017). 'It might hurt, but still it's good': People with rheumatoid arthritis beliefs and expectations about physical activity interventions. *Journal of Health Psychology, 22*(13), 1678-1690.

Larkin, L., Nordgren, B., Purtill, H., Brand, C., Fraser, A., & N, K. (2016). Criterion Validity of the activPAL Activity Monitor for Sedentary and Physical Activity Patterns in People Who Have Rheumatoid Arthritis. *Physical Therapy, 96*(7), 1093-1101.

Law, R. J., Markland, D. A., Jones, J. G., Maddison, P. J., & Thom, J. M. (2013). Perceptions of issues relating to exercise and joint health in rheumatoid arthritis: a UK-based questionnaire study. *Musculoskeletal Care, 11*(3), 147-158. doi:10.1002/msc.1037

Lee, D. M., & Weinblatt, M. E. (2001). Rheumatoid arthritis. *Lancet, 358*(9285), 903-911. doi:10.1016/s0140-6736(01)06075-5

Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet, 380*(9838), 219-229. doi:10.1016/s0140-6736(12)61031-9

Lee, S. J., & Kavanaugh, A. (2003). Pharmacological treatment of established rheumatoid arthritis. *Best Pract Res Clin Rheumatol, 17*(5), 811-829. doi:10.1016/S1521-6942(03)00048-2

Lee, Y. C., Bingham, C. O., Edwards, R. R., Marder, W., Phillips, K., Bolster, M., & Clauw, D. J. (2018). Association Between Pain Sensitization and Disease Activity in Patients With Rheumatoid Arthritis: A Cross-Sectional Study. *Arthritis Care & Research, 70*(2), 197-204.

Lee, Y. C., Chibnik, L. B., Lu, B., Wasan, A. D., Edwards, R. R., Fossel, A. H., Helfgott, S. M., Solomon, D. H., Clauw, D. J., & Karlson, E. W. (2009). The relationship between disease activity, sleep, psychiatric distress and pain sensitivity in rheumatoid arthritis: a cross-sectional study. *Arthritis Research & Therapy, 11*(5), R160. doi:10.1186/ar2842

Lee, Y. C., Cui, J., Lu, B., Frits, M. L., Iannaccone, C. K., Shadick, N. A., Weinblatt, M. E., & Solomon, D. H. (2011). Pain persists in DAS28 rheumatoid arthritis remission but not in ACR/EULAR remission: a longitudinal observational study. *Arthritis Research & Therapy, 13*(3), R83. doi:10.1186/ar3353

Leeger-Aschmann, C. S., Schmutz, E. A., Zysset, A. E., Kakebeeke, T. H., Messerli-Bürgy, N., Stülb, K., Arhab, A., Meyer, A. H., Munsch, S., Jenni, O. G., Puder, J. J., & Kriemler, S. (2019). Accelerometer-derived physical activity estimation in preschoolers – comparison of cut-point sets incorporating the vector magnitude vs the vertical axis. *Bmc Public Health, 19*(1), 513. doi:10.1186/s12889-019-6837-7

Leenders, N. Y., Sherman, W. M., Nagaraja, H. N., & Kien, C. L. (2001). Evaluation of methods to assess physical activity in free-living conditions. *Medicine & Science in Sports & Exercise, 33*(7), 1233-1240. doi:10.1097/00005768-200107000-00024

Lesser, I. A., & Nienhuis, C. P. (2020). The Impact of COVID-19 on Physical Activity Behavior and Well-Being of Canadians. *International Journal of Environmental Research and Public Health, 17*(11). doi:10.3390/ijerph17113899

Lewis, G. N., Luke, H., Rice, D. A., Rome, K., & McNair, P. J. (2012a). Reliability of the conditioned pain modulation paradigm to assess endogenous inhibitory pain pathways. *Pain Research and Management, 17*(2), 98-102.

Lewis, G. N., Rice, D. A., & McNair, P. J. (2012b). Conditioned pain modulation in populations with chronic pain: a systematic review and meta-analysis. *The Journal of Pain, 13*(10), 936-944.

Li, L. C., Feehan, L. M., Xie, H., Lu, N., Shaw, C., Gromala, D., Avina-Zubieta, J. A., Koehn, C., Hoens, A. M., English, K., Tam, J., Therrien, S., Townsend, A. F., Noonan, G., & Backman, C. L. (2020). Efficacy of a Physical Activity Counseling Program With Use of a Wearable Tracker in People With Inflammatory Arthritis: A Randomized Controlled Trial. *Arthritis Care & Research, 72*(12), 1755-1765.

Linde, L., SØRensen, J. A. N., ØStergaard, M., HØRslev-Petersen, K. I. M., & Hetland, M. L. (2008). Health-Related Quality of Life: Validity, Reliability, and Responsiveness of SF-36, EQ-15D, EQ-5D, RAQoL, and HAQ in Patients with Rheumatoid Arthritis. *Journal of Rheumatology, 35*(8), 1528.

Liu, S., Brooks, D., Thomas, S., Eysenbach, G., & Nolan, R. P. (2015). Lifesource XL-18 pedometer for measuring steps under controlled and free-living conditions. *Journal of Sports Sciences, 33*(10), 1001-1006. doi:10.1080/02640414.2014.981847

Lofgren, M., Opava, C. H., Demmelmaier, I., Friden, C., Lundberg, I. E., Nordgren, B., & Kosek, E. (2018a). Long-term, health-enhancing physical activity is associated with reduction of pain but not pain sensitivity or improved exercise-induced hypoalgesia in persons with rheumatoid arthritis. *Arthritis Research & Therapy, 20*(1), 262.

Lofgren, M., Opava, C. H., Demmelmaier, I., Friden, C., Lundberg, I. E., Nordgren, B., & Kosek, E. (2018b). Pain sensitivity at rest and during muscle contraction in persons with rheumatoid arthritis: a substudy within the Physical Activity in Rheumatoid Arthritis 2010 study. *Arthritis Research & Therapy, 20*(1), 48.

Loppenthin, K., & Esbensen, B. A. (2015). Sleep quality and correlates of poor sleep in patients with rheumatoid arthritis. *Clinical Rheumatology, 34*, 2029-2039. doi:10.1007/s10067-015-2875-4

Loprinzi, P. D., & Smith, B. (2017). Comparison Between Wrist-Worn and Waist-Worn Accelerometry. *Journal of Physical Activity & Health, 14*(7), 539-545. doi:10.1123/jpah.2016-0211

Luyster, F. S., Chasens, E. R., Wasko, M. C. M., & Dunbar-Jacob, J. (2011). Sleep quality and functional disability in patients with rheumatoid arthritis. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine, 7*(1), 49-55.

Luzak, A., Heier, M., Thorand, B., Laxy, M., Nowak, D., Peters, A., Schulz, H., & Group, K. O.-S. (2017). Physical activity levels, duration pattern and adherence to WHO recommendations in German adults. *Plos One, 12*(2), e0172503-e0172503. doi:10.1371/journal.pone.0172503

Lwin, M. N., Serhal, L., Holroyd, C., & Edwards, C. J. (2020). Rheumatoid Arthritis: The Impact of Mental Health on Disease: A Narrative Review. *Rheumatology and Therapy, 7*(3), 457-471. doi:10.1007/s40744-020-00217-4

Machin, A. R., Babatunde, O., Haththotuwa, R., Scott, I., Blagojevic-Bucknall, M., Corp, N., Chew-Graham, C. A., & Hider, S. L. (2020). The association between anxiety and disease activity and quality of life in rheumatoid arthritis: a systematic review and meta-analysis. *Clinical Rheumatology, 39*(5), 1471-1482. doi:10.1007/s10067-019-04900-y

Madsen, S. G., Danneskiold-Samsøe, B., Stockmarr, A., & Bartels, E. M. (2016). Correlations between fatigue and disease duration, disease activity, and pain in patients with rheumatoid arthritis: a systematic review. *Scandinavian Journal of Rheumatology, 45*(4), 255-261. doi:10.3109/03009742.2015.1095943

Maini, R. N., Breedveld, F. C., Kalden, J. R., Smolen, J. S., Furst, D., Weisman, M. H., St.Clair, E. W., Keenan, G. F., van der Heijde, D., Marsters, P. A., & Lipsky, P. E. (2004). Sustained improvement over two years in physical function, structural damage, and signs and symptoms among patients with rheumatoid arthritis treated with infliximab and methotrexate. *Arthritis & Rheumatism, 50*(4), 1051-1065. doi:10.1002/art.20159

Manns, P. J., Dunstan, D. W., Owen, N., & Healy, G. N. (2012). Addressing the nonexercise part of the activity continuum: a more realistic and achievable approach to activity programming for adults with mobility disability? *Physical Therapy, 92*(4), 614-625. doi:10.2522/ptj.20110284

Manresa, J. A. B., Fritsche, R., Vuilleumier, P. H., Oehler, C., Mørch, C. D., Arendt-Nielsen, L., Andersen, O. K., & Curatolo, M. (2014). Is the conditioned pain modulation paradigm reliable? A test-retest assessment using the nociceptive withdrawal reflex. *PLoS ONE, 9*(6).

Manresa, J. A. B., Neziri, A. Y., Curatolo, M., Arendt-Nielsen, L., & Andersen, O. K. (2011). Test–retest reliability of the nociceptive withdrawal reflex and electrical pain thresholds after single and repeated stimulation in patients with chronic low back pain. *European journal of applied physiology, 111*(1), 83-92.

Marcuzzi, A., Wrigley, P. J., Dean, C. M., Adams, R., & Hush, J. M. (2017). The long-term reliability of static and dynamic quantitative sensory testing in healthy individuals. *Pain, 158*(7), 1217-1223.

Marcuzzi, A., Wrigley, P. J., Dean, C. M., Graham, P. L., & Hush, J. M. (2018). From acute to persistent low back pain: a longitudinal investigation of somatosensory changes using quantitative sensory testing-an exploratory study. *Pain Reports, 3*(2), e641. doi:10.1097/pr9.0000000000000641

Margaretten, M., Julian, L., Katz, P., & Yelin, E. (2011). Depression in patients with rheumatoid arthritis: description, causes and mechanisms. *International Journal of Clinical Rheumtology, 6*(6), 617-623. doi:10.2217/ijr.11.6

Martel, M. O., Wasan, A. D., & Edwards, R. R. (2013). Sex differences in the stability of conditioned pain modulation (CPM) among patients with chronic pain. *Pain Medicine, 14*(11), 1757-1768.

Martin, A., Fitzsimons, C., Jepson, R., Saunders, D. H., van der Ploeg, H. P., Teixeira, P. J., Gray, C. M., & Mutrie, N. (2015). Interventions with potential to reduce sedentary time in adults: systematic review and meta-analysis. *British Journal of Sports Medicine, 49*(16), 1056-1063. doi:10.1136/bjsports-2014-094524

Martin, J., Krc, K., Mitchell, E., Eng, J., & Noble, J. (2012). Pedometer accuracy in slow walking older adults. *International Journal of Therapeutic Rehabilitation, 19*(7), 387-393.

Mâsse, L. C., Fuemmeler, B. F., Anderson, C. B., Matthews, C. E., Trost, S. G., Catellier, D. J., & Treuth, M. (2005). Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Medicine & Science in Sports & Exercise, 37*(11 Suppl), S544-554. doi:10.1249/01.mss.0000185674.09066.8a

Matcham, F., Norton, S., Scott, D. L., Steer, S., & Hotopf, M. (2016). Symptoms of depression and anxiety predict treatment response and long-term physical health outcomes in rheumatoid arthritis: secondary analysis of a randomized controlled trial. *Rheumatology, 55*(2), 268-278. doi:10.1093/rheumatology/kev306

Matcham, F., Rayner, L., Steer, S., & Hotopf, M. (2013). The Prevalence of Depression in Rheumatoid Arthritis: A Systematic Review and Meta-Analysis. *Rheumatology (Oxford), 52*(12), 2136-2148. doi:10.1093/rheumatology/ket169.

Matcham, F., Scott, I. C., Rayner, L., Hotopf, M., Kingsley, G. H., Norton, S., Scott, D. L., & Steer, S. (2014). The impact of rheumatoid arthritis on quality-of-life assessed using the SF-36: a systematic review and meta-analysis. *Seminars in Arthritis and Rheumatism, 44*(2), 123-130. doi:10.1016/j.semarthrit.2014.05.001

Matias, T., Dominski, F. H., & Marks, D. F. (2020). Human needs in COVID-19 isolation. *J Health Psychol, 25*(7), 871-882. doi:10.1177/1359105320925149

Matthews, C. E. (2005). Calibration of Accelerometer Output for Adults. *Medicine & Science in Sports & Exercise, 37*(11), S512-S522.

Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., & Troiano, R. P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003-2004. *Am J Epidemiol, 167*(7), 875-881. doi:10.1093/aje/kwm390

Maugeri, G., Castrogiovanni, P., Battaglia, G., Pippi, R., D'Agata, V., Palma, A., Di Rosa, M., & Musumeci, G. (2020). The impact of physical activity on psychological health during Covid-19 pandemic in Italy. *Heliyon, 6*(6), e04315. doi:10.1016/j.heliyon.2020.e04315

McIver, J., & Carmines, E. G. (1981). *Unidimensional scaling* (Vol. 24): Sage.

McKenna, S. G., Donnelly, A., Esbensen, B. A., Comber, L., Ng, W. L., Anjum, A. M., Fraser, A., & Kennedy, N. M. (2021). The feasibility of an exercise intervention to improve sleep (time, quality and disturbance) in people with rheumatoid arthritis: a pilot RCT. *Rheumatology International, 41*(2), 297-310. doi:10.1007/s00296-020-04760-9

McMullan, I. I., Bunting, B. P., McDonough, S. M., Tully, M. A., & Casson, K. (2020). The association between light intensity physical activity with gait speed in older adults (≥ 50 years). A longitudinal analysis using the English Longitudinal Study of Ageing (ELSA). *Aging Clinical and Experimental Research, 32*(11), 2279-2285. doi:10.1007/s40520-019-01453-7

McTiernan, A., Kooperberg, C., White, E., Wilcox, S., Coates, R., Adams-Campbell, L. L., Woods, N., & Ockene, J. (2003). Recreational Physical Activity and the Risk of Breast Cancer in Postmenopausal Women: The Women's Health Initiative Cohort Study. *Journal of the American Medical Association, 290*(10), 1331-1336. doi:10.1001/jama.290.10.1331

McWilliams, D. F., & Walsh, D. A. (2017). Pain mechanisms in rheumatoid arthritis. *Clinical & Experimental Rheumatology, 35*, S94-S101.

Melzack, R. (1987). The Short-Form McGill Pain Questionnaire. *Pain, 30*(2), 191-197. doi:10.1016/0304-3959(87)91074-8

Merskey, H., & Bogduk, N. (1994). *Classification of Chronic Pain*. Seattle, USA: IASP Publications.

Metsios, G. S., & Kitas, G. D. (2018). Physical activity, exercise and rheumatoid arthritis: Effectiveness, mechanisms and implementation. *Best Pract Res Clin Rheumatol, 32*(5), 669-682. doi:10.1016/j.berh.2019.03.013

Metsios, G. S., Moe, R. H., & Kitas, G. D. (2020). Exercise and inflammation. *Best Pract Res Clin Rheumatol, 34*(2), 101504. doi:10.1016/j.berh.2020.101504

Metsios, G. S., Stavropoulos-Kalinoglou, A., & Kitas, G. D. (2015). The role of exercise in the management of rheumatoid arthritis. *Expert Rev Clin Immunol, 11*(10), 1121-1130. doi:10.1586/1744666X.2015.1067606

Metsios, G. S., Stavropoulos-Kalinoglou, A., Panoulas, V., Wilson, M., A., N., & Koutedakis, Y. (2009). Association of physical inactivity with increased cardiovascular risk in patients with rheumatoid arthritis. *Eur J Cardiovasc Prev Rehabil, 16*, 188-194.

Metsios, G. S., Stavropoulos-Kalinoglou, A., Treharne, G. J., Nevill, A. M., Sandoo, A., Panoulas, V. F., Toms, T. E., Koutedakis, Y., & Kitas, G. D. (2011). Disease activity and low physical activity associate with number of hospital admissions and length of hospitalisation in patients with rheumatoid arthritis. *Arthritis Research & Therapy, 13*(3), R108. doi:10.1186/ar3390

Meyer, J., McDowell, C., Lansing, J., Brower, C., Smith, L., Tully, M., & Herring, M. (2020). Changes in Physical Activity and Sedentary Behavior in Response to COVID-19 and Their Associations with Mental Health in 3052 US Adults. *International Journal of Environmental Research and Public Health, 17*(18), E6469.

Micalos, P. S., Drinkwater, E. J., Cannon, J., Arendt-Nielsen, L., & Marino, F. E. (2009). Reliability of the nociceptive flexor reflex (RIII) threshold and association with pain threshold. *European journal of applied physiology, 105*(1), 55.

Michaud, K., Wipfler, K., Shaw, Y., & Simon, T. (2020). Experiences of Patients With Rheumatic Diseases in the United States During Early Days of the COVID-19 Pandemic. *ACR Open Rheumatol, 2*(6), 335-343. doi:10.1002/acr2.11148

Middlebrook, N., Heneghan, N. R., Evans, D. W., Rushton, A., & Falla, D. (2020). Reliability of temporal summation, thermal and pressure pain thresholds in a healthy cohort and musculoskeletal trauma population. *Plos One, 15*(5), e0233521. doi:10.1371/journal.pone.0233521

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Medicine, 47*(9), 1821-1845. doi:10.1007/s40279-017-0716-0

Mikkelsen, K., Stojanovska, L., & Polenakovic, M. (2017). Exercise and mental health. *Maturitas, 106*, 48-56. doi:10.1016/j.maturitas.2017.09.003

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & Group, P.-P. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1), 1. doi:10.1186/2046-4053-4-1

Moloney, N. A., Hall, T. M., & Doody, C. M. (2012). Reliability of thermal quantitative sensory testing: a systematic review. *J Rehabil Res Dev, 49*(2), 191-207. doi:10.1682/jrrd.2011.03.0044

Montoye, A. H. K., Clevenger, K. A., Pfeiffer, K. A., Nelson, M. B., Bock, J. M., Imboden, M. T., & Kaminsky, L. A. (2020). Development of cut-points for determining activity intensity from a wrist-worn ActiGraph accelerometer in free-living adults. *Journal of Sports Sciences, 38*(22), 2569-2578. doi:10.1080/02640414.2020.1794244

Montoye, A. H. K., Moore, R. W., Bowles, H. R., Korycinski, R., & Pfeiffer, K. A. (2018a). Reporting accelerometer methods in physical activity intervention studies: a systematic review and recommendations for authors. *British Journal of Sports Medicine, 52*(23), 1507-1516. doi:10.1136/bjsports-2015-095947

Montoye, A. H. K., Nelson, M. B., Bock, J. M., Imboden, M. T., Kaminsky, L. A., Mackintosh, K. A., McNarry, M. A., & Pfeiffer, K. A. (2018b). Raw and Count Data Comparability of Hip-Worn ActiGraph GT3X+ and Link Accelerometers. *Medicine & Science in Sports & Exercise, 50*(5), 1103-1112. doi:10.1249/MSS.0000000000001534

Müller, M., Curatolo, M., Limacher, A., Neziri, A. Y., Treichel, F., Battaglia, M., Arendt-Nielsen, L., & Jüni, P. (2019). Predicting transition from acute to chronic low back pain with quantitative

sensory tests—A prospective cohort study in the primary care setting. *European Journal of Pain, 23*(5), 894-907.

National Heart, L., and Blood Institute. (2019). Study Quality Assessment Tools *Study Quality Assessment Tools.* Retrieved from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Nerurkar, L., Siebert, S., McInnes, I. B., & Cavanagh, J. (2019). Rheumatoid arthritis and depression: an inflammatory perspective. *The Lancet Psychiatry, 6*(2), 164-173. doi:10.1016/s2215-0366(18)30255-4

Neuberger, G. B., Aaronson, L. S., Gajewski, B., Embretson, S. E., Cagle, P. E., Loudon, J. K., & Miller, P. A. (2007). Predictors of exercise and effects of exercise on symptoms, function, aerobic fitness, and disease outcomes of rheumatoid arthritis. *Arthritis & Rheumatism, 57*(6), 943-952.

Neuberger, G. B., Press, A. N., Lindsley, H. B., Hinton, R., Cagle, P. E., Carlson, K., Scott, S., Dahl, J., & 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.

Nguyen, P., Le, L. K.-D., Nguyen, D., Gao, L., Dunstan, D. W., & Moodie, M. (2020). The effectiveness of sedentary behaviour interventions on sitting time and screen time in children and adults: an umbrella review of systematic reviews. *International Journal of Behavioral Nutrition and Physical Activity, 17*(1), 117. doi:10.1186/s12966-020-01009-3

Nguyen, S., Häcker, A.-L., Henderson, M., Barnett, T., Mathieu, M.-E., Pagani, L., & Bigras, J.-L. (2016). Physical Activity Programs with Post-Intervention Follow-Up in Children: A Comprehensive Review According to Categories of Intervention. *International Journal of Environmental Research and Public Health, 13*(7), 664. doi:10.3390/ijerph13070664

Nie, H., Arendt-Nielsen, L., Andersen, H., & Graven-Nielsen, T. (2005). Temporal summation of pain evoked by mechanical stimulation in deep and superficial tissue. *Journal of Pain, 6*(6), 348-355. doi:10.1016/j.jpain.2005.01.352

Nikiphorou, E., Santos, E. J. F., Marques, A., Böhm, P., Bijlsma, J., Daien, C. I., Esbensen, B. A., Ferreira, R. J. O., Fragoulis, G. E., Holmes, P., McBain, H., Metsios, G. S., Moe, R. H., Stamm, T. A., de Thurah, A., Zabalan, C., Carmona, L., & Bosworth, A. (2021). 2021 EULAR recommendations for the implementation of self-management strategies in patients with inflammatory arthritis. *Ann Rheum Dis, 80*(10), 1278. doi:10.1136/annrheumdis-2021-220249

Nikolaus, S., Bode, C., Taal, E., & van de Laar, M. A. (2013). Fatigue and factors related to fatigue in rheumatoid arthritis: a systematic review. *Arthritis Care & Research, 65*(7), 1128-1146. doi:10.1002/acr.21949

Nocon, M., Hiemann, T., Müller-Riemenschneider, F., Thalau, F., Roll, S., & Willich, S. N. (2008). Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. *Eur J Cardiovasc Prev Rehabil, 15*(3), 239-246. doi:10.1097/HJR.0b013e3282f55e09

Nordgren, B., Friden, C., Demmelmaier, I., Bergstrom, G., Lundberg, I. E., Dufour, A. B., Opava, C. H., & Group, P. S. (2015). An outsourced health-enhancing physical activity programme for people with rheumatoid arthritis: exploration of adherence and response. *Rheumatology, 54*(6), 1065-1073.

Nussbaum, E. L., & Downes, L. (1998). Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. *Physical Therapy, 78*(2), 160-169.

O'Brien, A. T., Deitos, A., Pego, Y. T., Fregni, F., & Carrillo-de-la-Peña, M. T. (2018a). Defective endogenous pain modulation in fibromyalgia: a meta-analysis of temporal summation and conditioned pain modulation paradigms. *The Journal of Pain, 19*(8), 819-836.

O'Brien, C. M., Duda, J. L., Kitas, G. D., Veldhuijzen Van Zanten, J. J. C. S., Metsios, G. S., & Fenton, S. A. M. (2018b). Correlates of sedentary behaviour and light physical activity in people living

with rheumatoid arthritis: Protocol for a longitudinal study. *Mediterranean Journal of Rheumatology, 29*(2), 106-117.

O'Brien, C. M., Duda, J. L., Kitas, G. D., Veldhuijzen Van Zanten, J. J. C. S., Metsios, G. S., & Fenton, S. A. M. (2020). Measurement of sedentary time and physical activity in rheumatoid arthritis: an ActiGraph and activPAL™ validation study. *Rheumatology International, 40*(9), 1509-1518. doi:10.1007/s00296-020-04608-2

O'Brien, C. M., Ntoumanis, N., Duda, J. L., Kitas, G. D., Veldhuijzen Van Zanten, J. J. C. S., Metsios, G. S., & Fenton, S. A. M. (2021). Pain and fatigue are longitudinally and bi-directionally associated with more sedentary time and less standing time in rheumatoid arthritis. *Rheumatology (Oxford), 60*(10), 4548-4557. doi:10.1093/rheumatology/keab029

O'Leary, H., Larkin, L., Murphy, G. M., & Quinn, K. (2021). Relationship Between Pain and Sedentary Behavior in Rheumatoid Arthritis Patients: A Cross-Sectional Study. *Arthritis Care & Research, 73*(7), 990-997. doi:10.1002/acr.24207

Okely, J. A., Čukić, I., Shaw, R. J., Chastin, S. F., Dall, P. M., Deary, I. J., Der, G., Dontje, M. L., Skelton, D. A., Gale, C. R., Skelton, D. A., Chastin, S., Cox, S., Coulter, E., Čukić, I., Dall, P., Deary, I., Der, G., Dontje, M., Fitzsimons, C., Gale, C., Gill, J., Granat, M., Gray, C., Greig, C., Hindle, E., Laird, K., Mead, G., Mutrie, N., Palmer, V., Radakovic, R., Sattar, N., Shaw, R., Starr, J., Stewart, S., Wyke, S., & on behalf of the Seniors, U. S. P. t. (2019). Positive and negative well-being and objectively measured sedentary behaviour in older adults: evidence from three cohorts. *BMC Geriatrics, 19*(1), 28. doi:10.1186/s12877-019-1026-1

Olesen, S. S., van Goor, H., Bouwense, S. A., Wilder-Smith, O. H., & Drewes, A. M. (2012). Reliability of static and dynamic quantitative sensory testing in patients with painful chronic pancreatitis.

Olsen, C. L., Lie, E., Kvien, T. K., & Zangi, H. A. (2016). Predictors of Fatigue in Rheumatoid Arthritis Patients in Remission or in a Low Disease Activity State. *Arthritis Care & Research, 68*(7), 1043-1048. doi:10.1002/acr.22787

Osborne, N. R., Anastakis, D. J., & Davis, K. D. (2018). Peripheral nerve injuries, pain, and neuroplasticity. *J Hand Ther, 31*(2), 184-194. doi:10.1016/j.jht.2018.01.011

Østergaard, M., Ejbjerg, B., & Szkudlarek, M. (2005). Imaging in early rheumatoid arthritis: roles of magnetic resonance imaging, ultrasonography, conventional radiography and computed tomography. *Best Pract Res Clin Rheumatol, 19*(1), 91-116. doi:10.1016/j.berh.2004.08.006

Owen, N., Healy, G. N., Matthews, C. E., & Dunstan, D. W. (2010). Too much sitting: the population health science of sedentary behavior. *Exercise and sport sciences reviews, 38*(3), 105-113. doi:10.1097/JES.0b013e3181e373a2

Owens, M. A., Bulls, H. W., Trost, Z., Terry, S. C., Gossett, E. W., Wesson-Sides, K. M., & Goodin, B. R. (2015). An examination of pain catastrophizing and endogenous pain modulatory processes in adults with chronic low back pain. *Pain Medicine, 17*(8), 1452-1464.

Park, G., Kim, C. W., Park, S. B., Kim, M. J., & Jang, S. H. (2011). Reliability and usefulness of the pressure pain threshold measurement in patients with myofascial pain. *Annals of rehabilitation medicine, 35*(3), 412-417.

Park, S., Thøgersen-Ntoumani, C., Veldhuijzen Van Zanten, J. J. C. S., & Ntoumanis, N. (2018). The Role of Physical Activity and Sedentary Behavior in Predicting Daily Pain and Fatigue in Older Adults: a Diary Study. *Ann Behav Med, 52*(1), 19-28. doi:10.1007/s12160-017-9921-1

Paul, L., Rafferty, D., Marshall-McKenna, R., Gill, J. M., McInnes, I., Porter, D., & Woodburn, J. (2014). Oxygen cost of walking, physical activity, and sedentary behaviours in rheumatoid arthritis. *Scandinavian Journal of Rheumatology, 43*(1), 28-34.

Paungmali, A., Sitilertpisan, P., Taneyhill, K., Pirunsan, U., & Uthaikhup, S. (2012). Intrarater Reliability of Pain Intensity, Tissue Blood Flow, Thermal Pain Threshold, Pressure Pain Threshold and Lumbo-Pelvic Stability Tests in Subjects with Low Back Pain. *Asian journal of sports medicine, 3*, 8-14. doi:10.5812/asjsm.34718

Pavlakovic, G., & Petzke, F. (2010). The Role of Quantitative Sensory Testing in the Evaluation of Musculoskeletal Pain Conditions. *Curr Rheumatol. Rep., 12*(6), 455-461. doi:10.1007/s11926-010-0131-0

Pedersen, B. K., & Saltin, B. (2006). Evidence for prescribing exercise as therapy in chronic disease. *Scandinavian Journal of Medicine & Science in Sports, 16*(S1), 3-63. doi:10.1111/j.1600-0838.2006.00520.x

Pelfort, X., Torres-Claramunt, R., Sánchez-Soler, J. F., Hinarejos, P., Leal-Blanquet, J., Valverde, D., & Monllau, J. C. (2015). Pressure algometry is a useful tool to quantify pain in the medial part of the knee: an intra- and inter-reliability study in healthy subjects. *Orthop Traumatol Surg Res., 101*(5), 559-563.

Pépin, J. L., Bruno, R. M., Yang, R.-Y., Vercamer, V., Jouhaud, P., Escourrou, P., & Boutouyrie, P. (2020). Wearable Activity Trackers for Monitoring Adherence to Home Confinement During the COVID-19 Pandemic Worldwide: Data Aggregation and Analysis. *J Med Internet Res, 22*(6), e19787. doi:10.2196/19787

Pigg, M., Baad-Hansen, L., Svensson, P., Drangsholt, M., & List, T. (2010). Reliability of intraoral quantitative sensory testing (QST). *Pain, 148*(2), 220-226.

Pinto, A. J., Peçanha, T., Meireles, K., Benatti, F. B., Bonfiglioli, K., De Sá Pinto, A. L., Lima, F. R., Pereira, R. M. R., Irigoyen, M. C. C., Turner, J. E., Kirwan, J. P., Owen, N., Dunstan, D. W., Roschel, H., & Gualano, B. (2020a). A randomized controlled trial to reduce sedentary time in rheumatoid arthritis: Protocol and rationale of the Take a STAND for Health study. *Trials, 21*(1). doi:10.1186/s13063-020-4104-y

Pinto, A. J., Rezende, D., Sieczkowska, S. M., Meireles, K., Bonfiglioli, K., Ribeiro, A. C. M., Bonfa, E., Owen, N., Dunstan, D. W., Roschel, H., & Gualano, B. (2020b). Increased prolonged sitting in rheumatoid arthritis patients during the COVID-19 pandemic: a within-subjects, accelerometer-based study. *medRxiv*. doi:10.1101/2020.09.09.20191395

Plant, M. J., O'Sullivan, M. M., Lewis, P. A., Camilleri, J. P., Coles, E. C., & Jessop, J. D. (2005). What factors influence functional ability in patients with rheumatoid arthritis. Do they alter over time? *Rheumatology, 44*(9), 1181-1185. doi:10.1093/rheumatology/keh707

Plasqui, G. (2008). The role of physical activity in rheumatoid arthritis. *Physiology & Behavior, 94*(2), 270-275. doi:10.1016/j.physbeh.2007.12.012

Pollard, L. C., Choy, E. H., Gonzalez, J., Khoshaba, B., & Scott, D. L. (2006). Fatigue in rheumatoid arthritis reflects pain, not disease activity. *Rheumatology (Oxford), 45*(7), 885-889. doi:10.1093/rheumatology/kel021

Pollard, L. C., Ibrahim, F., Choy, E. H., & Scott, D. L. (2012). Pain thresholds in rheumatoid arthritis: the effect of tender point counts and disease duration. *Journal of Rheumatology, 39*(1), 28-31. doi:10.3899/jrheum.110668

Portney, L. G., & Watkins, M. P. (2009). *Foundations of Clinical Research: Applications to Practice* (3rd Edition ed.). Upper Saddle River, N.J: Pearson/Prentice Hall.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods, 40*(3), 879-891. doi:10.3758/brm.40.3.879

Prevoo, M. L., van 't Hof, M. A., Kuper, H. H., van Leeuwen, M. A., van de Putte, L. B., & van Riel, P. L. (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, 38*(1), 44-48. doi:10.1002/art.1780380107

Prioreschi, A., Hodkinson, B., Avidon, I., Tikly, M., & McVeigh, J. A. (2013). The clinical utility of accelerometry in patients with rheumatoid arthritis. *Rheumatology (Oxford), 52*(9), 1721-1727. doi:10.1093/rheumatology/ket216

Prioreschi, A., Hodkinson, B., Tikly, M., & McVeigh, J. A. (2014). Changes in physical activity measured by accelerometry following initiation of DMARD therapy in rheumatoid arthritis. *Rheumatology, 53*(5), 923-926.

Puccinelli, P. J., da Costa, T. S., Seffrin, A., de Lira, C. A. B., Vancini, R. L., Nikolaidis, P. T., Knechtle, B., Rosemann, T., Hill, L., & Andrade, M. S. (2021). Reduced level of physical activity during COVID-19 pandemic is associated with depression and anxiety levels: an internet-based survey. *Bmc Public Health, 21*(1), 425. doi:10.1186/s12889-021-10470-z

Qiu, X.-j., Zhang, X.-l., Cai, L.-s., Yan, C., Yu, L., Fan, J., Zhang, R.-w., Huang, J.-w., & Duan, X.-w. (2019). Rheumatoid arthritis and risk of anxiety: a meta-analysis of cohort studies. *Clinical Rheumatology, 38*(8), 2053-2061. doi:10.1007/s10067-019-04502-8

Rachele, J. N., McPhail, S. M., Washington, T. L., & Cuddihy, T. F. (2012). Practical physical activity measurement in youth: a review of contemporary approaches. *World J Pediatr, 8*(3), 207-216. doi:10.1007/s12519-012-0359-z

Rausch-Osthoff, A. K., Juhl, C. B., Knittle, K., Dagfinrud, H., Hurkmans, E., Braun, J., Schoones, J., Vliet Vlieland, T. P. M., & Niedermann, K. (2018). Effects of exercise and physical activity promotion: Meta-analysis informing the 2018 EULAR recommendations for physical activity in people with rheumatoid arthritis, spondyloarthritis and hip/knee osteoarthritis. *RMD Open, 4*(2).

Repping-Wuts, H., Uitterhoeve, R., van Riel, P., & van Achterberg, T. (2008). Fatigue as experienced by patients with rheumatoid arthritis (RA): a qualitative study. *International Journal of Nursing Studies, 45*(7), 995-1002. doi:10.1016/j.ijnurstu.2007.06.007

Rhudy, J. L., & France, C. R. (2007). Defining the nociceptive flexion reflex (NFR) threshold in human participants: a comparison of different scoring criteria. *Pain, 128*(3), 244-253.

Rhudy, M. B., Dreisbach, S. B., Moran, M. D., Ruggiero, M. J., & Veerabhadrappa, P. (2020). Cut points of the Actigraph GT9X for moderate and vigorous intensity physical activity at four different wear locations. *Journal of Sports Sciences, 38*(5), 503-510. doi:10.1080/02640414.2019.1707956

Robinson, E., Boyland, E., Chisholm, A., Harrold, J., Maloney, N. G., Marty, L., Mead, B. R., Noonan, R., & Hardman, C. A. (2021). Obesity, eating behavior and physical activity during COVID-19 lockdown: A study of UK adults. *Appetite, 156*, 104853. doi:10.1016/j.appet.2020.104853

Rodriguez-Rey, R., Garrido-Hernansaiz, H., & Collado, S. (2020). Psychological Impact and Associated Factors during the Initial Stage of the Coronavirus (COVID-19) Pandemic among the General Population in Spain. . *Front Psychol*. doi:10.3389/fpsyg.2020.01540

Rolke, R., Baron, R., Maier, C., Toelle, T. R., Treede, D. R., Beyer, A., & Binder, A. (2006a). Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): Standardized protocol and reference values. *Pain, 123*(3), 231-243.

Rolke, R., Magerl, W., Campbell, K. A., Schalber, C., Caspari, S., Birklein, F., & Treede, R. D. (2006b). Quantitative sensory testing: a comprehensive protocol for clinical trials. *European Journal of Pain, 10*(1), 77-77.

Rongen-van Dartel, S., Repping-Wuts, H., & Flendrie, M. (2015). Effect of Aerobic Exercise Training on Fatigue in Rheumatoid Arthritis: A Meta-Analysis. *Arthritis Care & Research, 67*(8), 1054-1062. doi:10.1002/acr.22561

Rosa-Gonçalves, D., Bernardes, M., & Costa, L. (2018). Quality of life and functional capacity in patients with rheumatoid arthritis - Cross-sectional study. *Reumatologia Clinica, 14*(6), 360-366. doi:10.1016/j.reuma.2017.03.002

Rosenberg, D. E., Bellettiere, J., Gardiner, P. A., Villarreal, V. N., Crist, K., & Kerr, J. (2016). Independent Associations Between Sedentary Behaviors and Mental, Cognitive, Physical, and Functional Health Among Older Adults in Retirement Communities. *J Gerontol A Biol Sci Med Sci, 71*(1), 78-83. doi:10.1093/gerona/glv103

Rosenberg, D. E., Norman, G. J., Wagner, N., Patrick, K., Calfas, K. J., & Sallis, J. F. (2010). Reliability and validity of the Sedentary Behavior Questionnaire (SBQ) for adults. *Journal of Physical Activity & Health, 7*(6), 697-705. doi:10.1123/jpah.7.6.697

Rouse, P. C., Veldhuijzen Van Zanten, J. J. C. S., Ntoumanis, N., Metsios, G. S., Yu, C. A., Kitas, G. D., & Duda, J. L. (2015). Measuring the positive psychological well-being of people with rheumatoid arthritis: a cross-sectional validation of the subjective vitality scale. *Arthritis Research & Therapy, 17*, 312. doi:10.1186/s13075-015-0827-7

Rowlands, A. V., Dawkins, N. P., Maylor, B., Edwardson, C. L., Fairclough, S. J., Davies, M. J., Harrington, D. M., Khunti, K., & Yates, T. (2019a). Enhancing the value of accelerometer-assessed physical activity: meaningful visual comparisons of data-driven translational accelerometer metrics. *Sports Medicine, 5*(1), 47. doi:10.1186/s40798-019-0225-9

Rowlands, A. V., Edwardson, C., Davies, M. J., Khunti, K., Harrington, D. M., & Yates, T. (2018). Beyond Cut Points: Accelerometer Metrics that Capture the Physical Activity Profile. *Medicine & Science in Sports & Exercise, 50*(6), 1323-1332. doi:10.1249/mss.0000000000001561

Rowlands, A. V., Fairclough, S. J., Yates, T., Edwardson, C., Davies, M. J., Munir, F., Khunti, K., & Stiles, V. (2019b). Activity intensity, volume & norms: Utility & interpretation of accelerometer metrics. *Medicine & Science in Sports & Exercise, 51*(11), 2410-2422. doi:10.1249/MSS.0000000000002047

Rupp, I., Boshuizen, H. C., Jacobi, C. E., Dinant, H. J., & van den Bos, G. A. (2004). Impact of fatigue on health-related quality of life in rheumatoid arthritis. *Arthritis & Rheumatism, 51*(4), 578-585. doi:10.1002/art.20539

Ryan, R. M., & Deci, E. L. (2001). On happiness and human potentials: a review of research on hedonic and eudaimonic well-being. *Annu Rev Psychol, 52*, 141-166. doi:10.1146/annurev.psych.52.1.141

Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Well-Ness*. New York: Guildford Press.

Ryan, R. M., & Frederick, C. (1997). On energy, personality, and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality, 65*(3), 529-565. doi:10.1111/j.1467-6494.1997.tb00326.x.

Sagelv, E. H., Ekelund, U., Pedersen, S., Brage, S., Hansen, B. H., Johansson, J., Grimsgaard, S., Nordström, A., Horsch, A., Hopstock, L. A., & Morseth, B. (2019). Physical activity levels in adults and elderly from triaxial and uniaxial accelerometry. The Tromsø Study. *Plos One, 14*(12), e0225670. doi:10.1371/journal.pone.0225670

Saidj, M., Jorgensen, T., Jacobsen, R. K., Linneberg, A., & Aadahl, M. (2014). Differential cross-sectional associations of work- and leisure-time sitting, with cardiorespiratory and muscular fitness among working adults. *Scand J Work Environ Health, 40*(5), 531-538. doi:10.5271/sjweh.3443

Samitz, G., Egger, M., & Zwahlen, M. (2011). Domains of physical activity and all-cause mortality: systematic review and dose–response meta-analysis of cohort studies. *International Journal of Epidemiology, 40*(5), 1382-1400. doi:10.1093/ije/dyr112

Santos-Lozano, A., Santín-Medeiros, F., Cardon, G., Torres-Luque, G., Bailón, R., Bergmeir, C., Ruiz, J. R., Lucia, A., & Garatachea, N. (2013). Actigraph GT3X: validation and determination of physical activity intensity cut points. *International Journal of Sports Medicine, 34*(11), 975-982. doi:10.1055/s-0033-1337945

Santos, D. A., Silva, A. M., Baptista, F., Santos, R., Vale, S., Mota, J., & Sardinha, L. B. (2012). Sedentary behavior and physical activity are independently related to functional fitness in older adults. *Exp Gerontol, 47*(12), 908-912. doi:10.1016/j.exger.2012.07.011

Sasaki, J. E., John, D., & Freedson, P. S. (2011). Validation and comparison of ActiGraph activity monitors. *Journal of Science and Medicine in Sport, 14*(5), 411-416. doi:10.1016/j.jsams.2011.04.003

Saunders, T. J., McIsaac, T., Douillette, K., Gaulton, N., Hunter, S., Tremblay, M. S., & Healy, G. N. (2020). Sedentary behaviour and health in adults: an overview of systematic reviews. *Applied Physiology, Nutrition, and Metabolism, 45*(10 (Suppl. 2)), S197-S217. doi:10.1139/apnm-2020-0272 %M 33054341

Schnohr, P., Scharling, H., & Jensen, J. S. (2003). Changes in leisure-time physical activity and risk of death: An observational study of 7,000 men and women. *Am J Epidemiol, 158*(7), 639-644. doi:10.1093/aje/kwg207

Scott, J., & Huskisson, E. (1976). Graphic representation of pain. *Pain, 2*, 175-184. doi:1026900

Sedentary Behaviour Research Network. (2012). Letter to the Editor: Standardized use of the terms "sedentary" and "sedentary behaviours". *Applied Physiology, Nutrition, and Metabolism, 37*(3), 540-542. doi:10.1139/h2012-024 %M 22540258

Sellers, C., Dall, P., Grant, M., & Stansfield, B. (2016). Validity and reliability of the activPAL3 for measuring posture and stepping in adults and young people. *Gait Posture, 43*, 42-47. doi:10.1016/j.gaitpost.2015.10.020

Semanik, P., Song, J., Chang, R. W., Manheim, L., Ainsworth, B., & Dunlop, D. (2010). Assessing physical activity in persons with rheumatoid arthritis using accelerometry. *Medicine and Science in Sports and Exercise, 42*(8), 1493-1501. doi:10.1249/MSS.0b013e3181cfc9da

Semanik, P., Wilbur, J., Sinacore, J., & Chang, R. W. (2004). Physical activity behavior in older women with rheumatoid arthritis. *Arthritis & Rheumatism, 51*(2), 246-252. doi:10.1002/art.20245

Senra, H., Rogers, H., Leibach, G., Altamar, M. L. P., Plaza, S. L. O., Perrin, P., & Durán, M. A. S. (2017). Health-related quality of life and depression in a sample of Latin American adults with rheumatoid arthritis. *International Journal of Rheumatic Diseases, 20*(11), 1684-1693. doi:10.1111/1756-185x.12412

Sherrer, Y. S., Bloch, D. A., Mitchell, D. M., Young, D. Y., & Fries, J. F. (1986). The development of disability in rheumatoid arthritis. *Arthritis & Rheumatism, 29*(4), 494-500. doi:10.1002/art.1780290406.

Singh, N. A., Clements, K. M., & Singh, M. A. F. (2001). The Efficacy of Exercise as a Long-term Antidepressant in Elderly Subjects: A Randomized, Controlled Trial. *The Journals of Gerontology: Series A, 56*(8), M497-M504. doi:10.1093/gerona/56.8.M497

Smets, E. M., Garssen, B., Bonke, B., & Haes, J. (1995). The Multidimensional Fatigue Inventory (MFI) Psychometric Qualities of an Instrument to Assess Fatigue. *J Psychosom Res., 39*(3), 315-325. doi:10.1016/0022-3999(94)00125-o

Smolen, J. S., Aletaha, D., & McInnes, I. B. (2016). Rheumatoid arthritis. *Lancet, 388*(10055), 2023-2038. doi:10.1016/s0140-6736(16)30173-8

Smolen, J. S., Breedveld, F. C., Schiff, M. H., Kalden, J. R., Emery, P., Eberl, G., van Riel, P. L., & Tugwell, P. (2003). A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford), 42*(2), 244-257. doi:10.1093/rheumatology/keg072

Smolen, J. S., Landewé, R., Bijlsma, J., Burmester, G., Chatzidionysiou, K., Dougados, M., Nam, J., Ramiro, S., Voshaar, M., van Vollenhoven, R., Aletaha, D., Aringer, M., Boers, M., Buckley, C. D., Buttgereit, F., Bykerk, V., Cardiel, M., Combe, B., Cutolo, M., van Eijk-Hustings, Y., Emery, P., Finckh, A., Gabay, C., Gomez-Reino, J., Gossec, L., Gottenberg, J. E., Hazes, J. M. W., Huizinga, T., Jani, M., Karateev, D., Kouloumas, M., Kvien, T., Li, Z., Mariette, X., McInnes, I., Mysler, E., Nash, P., Pavelka, K., Poór, G., Richez, C., van Riel, P., Rubbert-Roth, A., Saag, K., da Silva, J., Stamm, T., Takeuchi, T., Westhovens, R., de Wit, M., & van der Heijde, D. (2017). EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis, 76*(6), 960-977. doi:10.1136/annrheumdis-2016-210715

Sokka, T., Häkkinen, A., Kautiainen, H., Maillefert, J. F., Toloza, S., Mørk Hansen, T., Calvo-Alen, J., Oding, R., Liveborn, M., Huisman, M., Alten, R., Pohl, C., Cutolo, M., Immonen, K., Woolf, A., Murphy, E., Sheehy, C., Quirke, E., Celik, S., Yazici, Y., Tlustochowicz, W., Kapolka, D., Skakic,

V., Rojkovich, B., Müller, R., Stropuviene, S., Andersone, D., Drosos, A. A., Lazovskis, J., & Pincus, T. (2008). Physical inactivity in patients with rheumatoid arthritis: data from twenty-one countries in a cross-sectional, international study. *Arthritis & Rheumatism, 59*(1), 42-50. doi:10.1002/art.23255

Sokka, T., Kankainen, A., & Hannonen, P. (2001). Scores for functional disability in patients with rheumatoid arthritis are correlated at higher levels with pain scores than with radiographic scores. *Arthritis & Rheumatism, 43*(2), 386-389. doi:10.1002/1529-0131(200002)43:2<386::AID-ANR19>3.0.CO;2-Z.

St. Clair, E. W., Pisetsky, D. S., & Haynes, B. F. (2004). *Rheumatoid arthritis / editors, E. William St. Clair, David S. Pisetsky, Barton F. Haynes*

with contributing authors: Philadelphia : Lippincott Williams and Wilkins, 2004.

Stamatakis, E., Hillsdon, M., & Primatesta, P. (2007). Domestic Physical Activity in Relationship to Multiple CVD Risk Factors. *Am J Prev Med, 32*(4), 320-327.e323. doi:10.1016/j.amepre.2006.12.020

Stavropoulos-Kalinoglou, A., Metsios, G. S., Veldhuijzen Van Zanten, J. J. C. S., Nightingale, P., Kitas, G. D., & Koutedakis, Y. (2013). Individualised aerobic and resistance exercise training improves cardiorespiratory fitness and reduces cardiovascular risk in patients with rheumatoid arthritis. *Ann Rheum Dis, 72*(11), 1819-1825. doi:10.1136/annrheumdis-2012-202075

Stenstrom, C. H. (1994). Home exercise in rheumatoid arthritis functional class II: goal setting versus pain attention. *Journal of Rheumatology, 21*(4), 627-634.

Stenström, C. H., & Minor, M. A. (2003). Evidence for the benefit of aerobic and strengthening exercise in rheumatoid arthritis. *Arthritis Care & Research, 49*(3), 428-434. doi:10.1002/art.11051

Sterne, J. A. C., Savović, J., Page, M., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., & Higgins, J. P. T. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ (Clinical research ed.), 366*.

Summers, G., Booth, A., Brooke-Wavell, K., Barami, T., & Clemes, S. (2019). Physical activity and sedentary behavior in women with rheumatoid arthritis: A comparison of patients with low and high disease activity and healthy controls. *Open Access Rheumatology: Research and Reviews, 11*, 133-142. doi:10.2147/OARRR.S203511

Suokas, A. K., Walsh, D. A., McWilliams, D. F., Wylde, V., Arendt-Nielsen, L., & Zhang, W. (2012). Quantitative sensory testing in painful osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis and Cartilage, 20*(10), 1075-1085.

Swardh, E., Nordgren, B., Opava, C. H., & Demmelmaier, I. (2020). "A Necessary Investment in Future Health": Perceptions of Physical Activity Maintenance Among People With Rheumatoid Arthritis. *Physical Therapy, 100*(12), 2144-2153. doi:10.1093/ptj/pzaa176

Sylvia, L. G., Bernstein, E. E., Hubbard, J. L., Keating, L., & Anderson, E. J. (2014). Practical guide to measuring physical activity. *Journal of the Academy of Nutrition and Dietetics, 114*(2), 199-208. doi:10.1016/j.jand.2013.09.018

Tack, B. B. (1991). *Dimensions and correlates of fatigue in older adults with rheumatoid arthritis [dissertation].* University of California, San Francisco

Taraldsen, K., Askim, T., Sletvold, O., Einarsen, E. K., Bjåstad, K. G., Indredavik, B., & Helbostad, J. L. (2011). Evaluation of a body-worn sensor system to measure physical activity in older people with impaired function. *Physical Therapy, 91*(2), 277-285. doi:10.2522/ptj.20100159

Taylor, A. H., Cable, N. T., Faulkner, G., Hillsdon, M., Narici, M., & Van Der Bij, A. K. (2004). Physical activity and older adults: a review of health benefits and the effectiveness of interventions. *Journal of Sports Sciences, 22*(8), 703-725. doi:10.1080/02640410410001712421

Testa, D., Calvacchi, S., Petrelli, F., Giannini, D., Bilia, S., Alunno, A., & Puxeddu, I. (2021). One year in review 2021: pathogenesis of rheumatoid arthritis. *Clinical & Experimental Rheumatology, 39*(3), 445-452.

Thomsen, T., Aadahl, M., Beyer, N., Hetland, M. L., Loppenthin, K., Midtgaard, J., Christensen, R., Ostergaard, M., Jennum, P. J., & Esbensen, B. A. (2017). The efficacy of motivational counselling and SMS reminders on daily sitting time in patients with rheumatoid arthritis: a randomised controlled trial. *Ann Rheum Dis, 76*(9), 1603-1606.

Thomsen, T., Aadahl, M., Beyer, N., Hetland, M. L., Loppenthin, K. B., Midtgaard, J., Christensen, R., Nielsen, S. M., Ostergaard, M., Jennum, P., & Esbensen, B. A. (2020). Sustained Long-Term Efficacy of Motivational Counseling and Text Message Reminders on Daily Sitting Time in Patients With Rheumatoid Arthritis: Long-Term Follow-up of a Randomized, Parallel-Group Trial. *Arthritis Care & Research, 72*(11), 1560-1570.

Thomsen, T., Beyer, N., Aadahl, M., Hetland, M. L., Løppenthin, K., Midtgaard, J., & Esbensen, B. A. (2015). Sedentary behaviour in patients with rheumatoid arthritis: A qualitative study. *International Journal of Qualitative Studies on Health and Well-being, 10*(1), 28578. doi:10.3402/qhw.v10.28578

Thomson, J. L., Landry, A. S., Zoellner, J. M., Tudor-Locke, C., Webster, M., Connell, C., & Yadrick, K. (2012). Several steps/day indicators predict changes in anthropometric outcomes: HUB City Steps. *Bmc Public Health, 12*, 983. doi:10.1186/1471-2458-12-983

Tierney, M., Fraser, A., & Kennedy, N. (2012). Physical Activity in Rheumatoid Arthritis: A Systematic Review. *Journal of Physical Activity & Health, 9*(7), 1036-1048. doi:10.1123/jpah.9.7.1036

Treharne, G. J., Hale, E. D., Lyons, A. C., Booth, D. A., Banks, M. J., Erb, N., Douglas, K. M., Mitton, D. L., & Kitas, G. D. (2005). Cardiovascular disease and psychological morbidity among rheumatoid arthritis patients. *Rheumatology (Oxford), 44*(2), 241-246. doi:10.1093/rheumatology/keh441

Treharne, G. J., Lyons, A. C., Booth, D. A., & Kitas, G. D. (2007). Psychological well-being across 1 year with rheumatoid arthritis: Coping resources as buffers of perceived stress. *British Journal of Health Psychology, 12*, 323-345.

Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N., & Owen, N. (2010). Physiological and health implications of a sedentary lifestyle. *Appl Physiol Nutr Metab, 35*(6), 725-740. doi:10.1139/h10-079

Troiano, R. P. (2007). Large-scale applications of accelerometers: new frontiers and new questions. *Medicine & Science in Sports & Exercise, 39*(9), 1501. doi:10.1097/mss.0b013e318150d42e

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine & Science in Sports & Exercise, 40*(1), 181-188. doi:10.1249/mss.0b013e31815a51b3

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British Journal of Sports Medicine, 48*(13), 1019-1023. doi:10.1136/bjsports-2014-093546

Tudor-Locke, C., & Myers, A. M. (2001). Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine, 31*(2), 91-100. doi:10.2165/00007256-200131020-00002

Tudor-Locke, C., Sisson, S. B., Collova, T., Lee, S. M., & Swan, P. (2005). Pedometer-determined step count guidelines for classifying walking intensity in a young ostensibly healthy population. *Canadian Journal of Applied Physiology, 30*(6), 666-676.

Uddin, Z., & MacDermid, J. C. (2016). Quantitative Sensory Testing in Chronic Musculoskeletal Pain. *Pain Medicine, 17*(9), 1694-1703. doi:10.1093/pm/pnv105

Uhlig, T., & Kvien, T. K. (2005). Is rheumatoid arthritis disappearing? *Ann Rheum Dis, 64*(1), 7-10. doi:10.1136/ard.2004.023044

Uhlig, T., Moe, R. H., & Kvien, T. K. (2014). The burden of disease in rheumatoid arthritis. *Pharmacoeconomics, 32*(9), 841-851. doi:10.1007/s40273-014-0174-6

Valencia, C., Fillingim, R. B., Bishop, M., Wu, S. S., Wright, T. W., Moser, M., Farmer, K., & George, S. Z. (2014). Investigation of central pain processing in post-operative shoulder pain and disability. *The Clinical journal of pain, 30*(9), 775.

Van Den Berg, M. H., Ronday, H. K., Peeters, A. J., Le Cessie, S., Van Der Giesen, F. J., Breedveld, F. C., & Vliet Vlieland, T. P. M. (2006). Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. *Arthritis Care & Research, 55*(6), 935-945.

Van Den Ende, C. H. M., Vliet Vlieland, T. P. M., Munneke, M., & Hazes, J. M. W. (1998). Dynamic exercise therapy in rheumatoid arthritis: A systematic review. *British Journal of Rheumatology, 37*(6), 677-687.

Van Der Ploeg, H. P., Merom, D., Chau, J. Y., Bittman, M., Trost, S. G., & Bauman, A. E. (2010). Advances in population surveillance for physical activity and sedentary behavior: reliability and validity of time use surveys. *Am J Epidemiol, 172*(10), 1199-1206. doi:10.1093/aje/kwq265

Van Hees, V. T., Renström, F., Wright, A., Gradmark, A., Catt, M., Chen, K. Y., Löf, M., Bluck, L., Pomeroy, J., Wareham, N. J., Ekelund, U., Brage, S., & Franks, P. W. (2011). Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *Plos One, 6*(7), e22922. doi:10.1371/journal.pone.0022922

Van Lankveld, W., Naring, G., Van't Pad Bosch, P. V. P., & Van de Putte, L. (2000). The negative effect of decreasing the level of activity in coping with pain in rheumatoid arthritis: An increase in psychological distress and disease impact. *Journal of Behavioral Medicine, 23*(4), 377-391. doi:10.1023/A:1005552901946

Van Lankveld, W., Pad Bosch, P., Van de Putte, L., Van der Staak, C., & Naring, G. (1992). Pain in Rheumatoid Arthritis Measured With the Visual Analogue Scale and the Dutch Version of the McGill Pain Questionnaire. *Nederlands Tijdschrift voor Geneeskunde, 136*(24), 1166-1170.

Van Roie, E., Delecluse, C., Opdenacker, J., De Bock, K., Kennis, E., & Boen, F. (2010). Effectiveness of a Lifestyle Physical Activity Versus a Structured Exercise Intervention in Older Adults. *Journal of Aging and Physical Activity, 18*(3), 335-352. doi:10.1123/japa.18.3.335

Van Tuyl, L. H., & Boers, M. (2015). Patient-reported outcomes in core domain sets for rheumatic diseases. *Nat Rev Rheumatol, 11*(12), 705-712. doi:10.1038/nrrheum.2015.116

VanDyke, M. M., Parker, J. C., Smarr, K. L., Hewett, J. E., Johnson, G. E., Slaughter, J. R., & Walker, S. E. (2004). Anxiety in rheumatoid arthritis. *Arthritis Care & Research, 51*(3), 408-412. doi:10.1002/art.20474

Vanhelst, J., Baquet, G., Gottrand, F., & Beghin, L. (2012). Comparative interinstrument reliability of uniaxial and triaxial accelerometers in free-living conditions. *Percept Mot Skills, 114*(2), 584-594. doi:10.2466/03.26.PMS.114.2.584-594

Veldhuijzen Van Zanten, J. J. C. S., Fenton, S. A. M., Brady, S. M., Metsios, G. S., Duda, J. L., & Kitas, G. D. (2020). Mental health and psychological wellbeing in rheumatoid arthritis during COVID-19 – can physical activity help? *Mediterranean Journal of Rheumatology, 31*, 284-287. doi:10.31138/mjr.31.3.284

Veldhuijzen Van Zanten, J. J. C. S., Fenton, S. A. M., Rouse, P. C., Ntoumanis, N., Osailan, A., Yu, C., Metsios, G. S., Kitas, G. D., & Duda, J. L. (2021). Autonomous motivation, cardiorespiratory fitness, and exercise in rheumatoid arthritis: Randomised controlled trial. *Psychology of Sport and Exercise, 55*, 101904. doi:10.1016/j.psychsport.2021.101904

Veldhuijzen Van Zanten, J. J. C. S., Rouse, P. C., Hale, E. D., Ntoumanis, N., Metsios, G. S., Duda, J. L., & Kitas, G. D. (2015). Perceived Barriers, Facilitators and Benefits for Regular Physical Activity and Exercise in Patients with Rheumatoid Arthritis: A Review of the Literature. *Sports Medicine, 45*(10), 1401-1412. doi:10.1007/s40279-015-0363-2

Vladimirova, N., Jespersen, A., Bartels, E. M., Christensen, A. W., Bliddal, H., & Danneskiold-Samsøe, B. (2015). Pain Sensitisation in Women with Active Rheumatoid Arthritis: A Comparative Cross-Sectional Study. *Arthritis, 2015*, 434109. doi:10.1155/2015/434109

Wahid, A., Manek, N., Nichols, M., Kelly, P., Foster, C., & Scarborough, P. (2016). Quantifying the Association Between Physical Activity and Cardiovascular Disease and Diabetes: A Systematic Review and Meta-Analysis. *Journal of the American Heart Association, 5*(9), e002495. doi:doi:10.1161/JAHA.115.002495

Walsh, D. A., & McWilliams, D. F. (2012). Pain in rheumatoid arthritis. *Current pain and headache reports, 16*(6), 509-517.

Walsh, D. A., & McWilliams, D. F. (2014). Mechanisms, impact and management of pain in rheumatoid arthritis. *Nature Reviews Rheumatology, 10*, 581-592. doi:10.1038/nrrheum.2014.64

Walter, S., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine, 17*(1), 101-110.

Walton, D., MacDermid, J., Nielson, W., Teasell, R., Chiasson, M., & Brown, L. (2011). Reliability, standard error, and minimum detectable change of clinical pressure pain threshold testing in people with and without acute neck pain. *Journal of Orthopaedic & Sports Physical Therapy, 41*(9), 644-650.

Wan, S., He, H., Mak, A., Lahiri, M., Luo, N., & Cheung, P. P. (2016). Health related quality of life and its predictors among patients with rheumatoid arthritis. *Applied Nursing Research, 30*, 176-183.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *Cmaj, 174*(6), 801-809. doi:10.1503/cmaj.051351

Ward, M. M., Guthrie, L. C., & Dasgupta, A. (2017). Direct and Indirect Determinants of the Patient Global Assessment in Rheumatoid Arthritis: Differences by Level of Disease Activity. *Arthritis Care & Research, 69*(3), 323-329. doi:10.1002/acr.22953

Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care, 30*(6), 473-483.

Warner-Schmidt, J. L., Vanover, K. E., Chen, E. Y., Marshall, J. J., & Greengard, P. (2011). Antidepressant effects of selective serotonin reuptake inhibitors (SSRIs) are attenuated by antiinflammatory drugs in mice and humans. *Proc Natl Acad Sci USA, 108*(22), 9262-9267. doi:10.1073/pnas.1104836108

Wassenberg, S., Rau, R., Steinfeld, P., & Zeidler, H. (2005). Very low-dose prednisolone in early rheumatoid arthritis retards radiographic progression over two years: a multicenter, double-blind, placebo-controlled trial. *Arthritis & Rheumatism, 52*(11), 3371-3380. doi:10.1002/art.21421

Watson, K. B., Carlson, S. A., Carroll, D. D., & Fulton, J. E. (2014). Comparison of accelerometer cut points to estimate physical activity in US adults. *Journal of Sports Sciences, 32*(7), 660-669. doi:10.1080/02640414.2013.847278

Weinstein, A. A., Drinkard, B. M., Diao, G., Furst, G., Dale, J. K., Straus, S. E., & Gerber, L. H. (2009). Exploratory Analysis of the Relationships between Aerobic Capacity and Self-Reported Fatigue in Patients with Rheumatoid Arthritis, Polymyositis, and Chronic Fatigue Syndrome. *PM&R, 1*(7), 620-628. doi:10.1016/j.pmrj.2009.04.007

Welsing, P. M., van Gestel, A. M., Swinkels, H. L., Kiemeney, L. A., & van Riel, P. L. (2001). The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum, 44*(9), 2009-2017. doi:10.1002/1529-0131(200109)44:9<2009::Aid-art349>3.0.Co;2-l

Westerterp, K. R. (2009). Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol, 105*(6), 823-828. doi:10.1007/s00421-009-1000-2

Whelton, S. P., Chin, A., Xin, X., & He, J. (2002). Effect of Aerobic Exercise on Blood Pressure. *Ann Intern Med, 136*(7), 493-503. doi:10.7326/0003-4819-136-7-200204020-00006

WHO. (1995). The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med, 41*(10), 1403-1409. doi:10.1016/0277-9536(95)00112-k

WHO. (2020a). Coronavirus disease (COVID-2019) situation reports. *Coronavirus disease (COVID-2019) situation reports.* Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

WHO. (2020b). Depression: let's talk. Retrieved from https://www.who.int/health-topics/depression#tab=tab_1

WHO. (2020c). WHO guidelines on physical activity and sedentary behaviour. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/336656/9789240015128-eng.pdf?sequence=1&isAllowed=y

WHO. (2022). WHO Coronavirus (COVID-19) Dashboard. Retrieved from https://covid19.who.int/

Wijndaele, K., I, D. E. B., Godino, J. G., Lynch, B. M., Griffin, S. J., Westgate, K., & Brage, S. (2014). Reliability and validity of a domain-specific last 7-d sedentary time questionnaire. *Medicine & Science in Sports & Exercise, 46*(6), 1248-1260. doi:10.1249/mss.0000000000000214

Williamson, P. R., Altman, D. G., Bagley, H., Barnes, K. L., Blazeby, J. M., Brookes, S. T., Clarke, M., Gargon, E., Gorst, S., Harman, N., Kirkham, J. J., McNair, A., Prinsen, C. A. C., Schmitt, J., Terwee, C. B., & Young, B. (2017). The COMET Handbook: version 1.0. *Trials, 18*(3), 280. doi:10.1186/s13063-017-1978-4

Wolfe, F., Hawley, D. J., & Cathey, M. A. (1991). Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis. *Journal of Rheumatology, 18*(9), 1290-1297.

Wolfe, F., Hawley, D. J., & Wilson, K. (1996). The prevalence and meaning of fatigue in rheumatic disease. *Journal of Rheumatology, 23*(8), 1407-1417.

Wylde, V., Palmer, S., Learmonth, I. D., & Dieppe, P. (2011). Test-retest reliability of Quantitative Sensory Testing in knee osteoarthritis and healthy participants. *Osteoarthritis and Cartilage, 19*(6), 655-658. doi:10.1016/j.joca.2011.02.009

Yarnitsky, D. (2010). Conditioned pain modulation (the diffuse noxious inhibitory control-like effect): its relevance for acute and chronic pain states. *Current Opinion in Anesthesiology, 23*(5), 611-615.

Yarnitsky, D., Bouhassira, D., Drewes, A., Fillingim, R., Granot, M., Hansson, P., Landau, R., Marchand, S., Matre, D., & Nilsen, K. (2015). Recommendations on practice of conditioned pain modulation (CPM) testing. *European Journal of Pain, 19*(6), 805-806.

Yngve, A., Nilsson, A., Sjostrom, M., & Ekelund, U. (2003). Effect of monitor placement and of activity setting on the MTI accelerometer output. *Medicine & Science in Sports & Exercise, 35*(2), 320-326. doi:10.1249/01.Mss.0000048829.75758.A0

Young, D. R., Hivert, M. F., Alhassan, S., Camhi, S. M., Ferguson, J. F., Katzmarzyk, P. T., Lewis, C. E., Owen, N., Perry, C. K., Siddique, J., & Yong, C. M. (2016). Sedentary Behavior and Cardiovascular Morbidity and Mortality: A Science Advisory From the American Heart Association. *Circulation, 134*(13), e262-279. doi:10.1161/cir.0000000000000440

Yu, C., Rouse, P. C., Veldhuijzen Van Zanten, J. J. C. S., Ntoumanis, N., Kitas, G. D., Duda, J. L., & Metsios, G. S. (2015). Subjective and objective levels of physical activity and their association with cardiorespiratory fitness in rheumatoid arthritis patients. *Arthritis Research & Therapy, 17*(1), 59. doi:10.1186/s13075-015-0584-7

Zhang, Y., Zhang, H., Ma, X., & Di, Q. (2020). Mental Health Problems during the COVID-19 Pandemics and the Mitigation Effects of Exercise: A Longitudinal Study of College Students in China. *International Journal of Environmental Research and Public Health, 17*(10). doi:10.3390/ijerph17103722

Ziade, N., El Kibbi, L., Hmamouchi, I., Abdulateef, N., Halabi, H., Hamdi, W., Abutiban, F., El Rakawi, M., Eissa, M., & Masri, B. (2020). Impact of the COVID-19 pandemic on patients with chronic rheumatic diseases: A study in 15 Arab countries. *International Journal of Rheumatic Diseases*. doi:10.1111/1756-185X.13960

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370. doi:10.1111/j.1600-0447.1983.tb09716.x