University of Nottingham
UK | CHINA | MALAYSIA

CENTRAL SENSITISATION AS A PREDICTOR OF SELF-MANAGEMENT IN CHRONIC LOW BACK PAIN

Vasileios Georgopoulos, BSc (PT), BSc (PE), MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

May, 2020

DECLARATIONS

This is to certify that work submitted in this thesis is the result of original research. It has been conducted substantially by myself with assistance outlined below. It has not already been accepted for any degree, diploma or other qualification. All authors and works to which reference has been made are fully acknowledged.

Study design, data collection, analysis and general administration for the Systematic Literature Review study were conducted primarily by myself, under the supervision of Professor David A. Walsh and Dr Paul Hendrick, with data coding advice from Dr Daniel McWilliams and statistical advice from Dr Weiya Zhang. Dr Kehinde Akin-Akinyosoye has contributed to the systematic literature review by performing independent data extraction and quality appraisal of the included titles.

Study design, ethical application, recruitment, data collection, analysis, writing and general administration for (i) the Reliability study and (ii) the Observational study were conducted primarily by myself, under the supervision of Professor David A. Walsh, and Dr Paul Hendrick, with advice from Dr Daniel McWilliams and significant contribution during the recruitment process from Mrs Debbie Wilson, Mrs Mary Martin and Mrs Wendy Furnell.

All data presented in this thesis were analysed by myself, and supervision of the thesis was undertaken by Professor David A. Walsh, and Dr Paul Hendrick.

LIST OF PUBLICATIONS

Papers in Preparation:

Georgopoulos, V., Akin-Akinyosoye, K., McWilliams, D.F., Hendrick, P. and Walsh, D.A., 2020. Central Sensitisation indices as predictors of self-management and self-care outcomes in individuals with Chronic Low Back Pain. **[in preparation]**

Georgopoulos, V., Akin-Akinyosoye, W., McWilliams, D.F., Hendrick, P. and Walsh, D.A., 2020. Reliability and validity of a Quantitative Sensory Testing protocol in healthy participants and patients with Chronic Low Back Pain. **[in preparation]**.

Published Papers:

Georgopoulos, V., Akin-Akinyosoye, K., Zhang, W., McWilliams, D.F., Hendrick, P. and 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), pp.1920-1932.

Conference Presentations:

International Association for the Study of Pain (IASP) World Congress (**Amsterdam, 2021**): Indices of central sensitisation can predict effective self-management in individuals with chronic low back pain (*Accepted for poster presentation*).

Osteoarthritis Research Society International (OARSI) World Congress (**Vienna, 2020**): Pain distribution as an indicator of central mechanisms in individuals with back or knee osteoarthritis pain *(Poster presentation)*.

British Pain Society Annual Scientific Meeting (**London, 2019**): Quantitative Sensory Testing (QST) can predict pain, disability and negative affect in musculoskeletal conditions: a systematic review and meta-analysis (*Oral presentation*).

Conference Abstracts:

Akin-Akinyosoye, K., Georgopoulos, V., Frowd, N., Swaithes, L., Stocks, J., Fernandes, G.S., Valdes, A., McWilliams, D.F., Zhang, W., Doherty, M. and Ferguson, E., 2020. Pain distant from the index site and sensitization in people with knee pain and low back pain. Osteoarthritis and Cartilage, 28, p.S375.

ABSTRACT

Background: Chronic low back pain (CLBP) is one of the most prevalent reasons people seek healthcare assistance worldwide. Guidelines for managing CLBP prioritise the development of self-management strategies. Levels of central sensitisation (CS) may contribute to the relatively poor efficacy of treatments aiming to facilitate self-management. CS might be a dominant factor predicting worse self-management in people with CLBP following interventions aiming to improve such outcomes. Quantitative sensory testing (QST) may provide reliable and valid indices of CS and it may predict musculoskeletal pain and disability. CS might be associated with increasing psychological distress, pain, fatigue and catastrophisation which might also be predictors of ineffective self-management. CS has also been associated in people with knee pain with self-report measures of widespread pain distribution (reported by shading a pain manikin) or a self-report Central Mechanisms Trait score, comprising of items addressing depression, anxiety, neuropathic-like symptoms, pain distribution, catastrophising, sleep, fatigue and cognitive difficulties.

Objectives: [1] to systematically review the literature in order to determine the ability of QST to predict musculoskeletal outcomes; [2] to establish the reliability and validity of distinct QST modalities as classification and measurement tools of CS; [3] to establish a cut off for number of body sites shaded on a self-reported pain manikin that best identifies those with widespread pain and explore whether certain self-reported items taken to indicate central mechanisms involvement contribute to a single latent trait in individuals with CLBP; [4] to determine whether different CS indices are associated specifically with self-management/self-care outcomes at a single time-point; [5] to test whether any cross-sectional associations between baseline CS indices and self-management/self-care outcomes are also present longitudinally, after participants have undertaken an intervention programme that aimed to improve such outcomes.

Methods: A systematic literature review (SLR) was conducted to collate the evidence regarding the ability of QST to predict pain, disability and negative affect using searches of 6 databases up to April 2018. Title screening, data extraction, and methodological quality assessments were performed independently by 2 reviewers. Associations were reported between baseline QST and outcomes using adjusted (β) and unadjusted (r) correlations.

Reliability of Pressure Pain Detection Threshold (PPT), Temporal Summation (TS) and Conditioned Pain Modulation (CPM) conducted at a site distant from the low back were assessed in healthy participants (n=25) and individuals with CLBP (n=25). The QST test site was the dominant forearm and conditioning site the contralateral arm.

Pain distribution was classified according to criteria proposed by the American College of Rheumatology (ACR) and other research groups. Receiver operating characteristics (ROC) analysis established the cut-off point for the optimal number of painful sites needed to classify low PPT (1st quartile). Confirmatory factor analysis (CFA) was used to assess model fit and produce a single Central Mechanisms Trait score based on unique items form 8 distinct self-reported tools.

The ability of baseline indices of CS (PPT, TS, CPM, number of painful sites on a manikin, and Central Mechanisms Trait score) to predict self-management outcomes at 3-months follow-up was assessed in individuals with CLBP (n=97) participating in a cognitive behavioural therapy (CBT)-based group physiotherapy intervention, which aimed to facilitate self-management. Self-management was measured in 8 discrete domains; health-directed behaviour, positive engagement in life, self-monitoring and insight, constructive attitudes and approaches, skill and technique acquisition, social integration and support, health services navigation and emotional distress. Pain (numerical rating scale), depression/anxiety (hospital anxiety-depression scale), fatigue (fatigue severity scale) and catastrophising (pain catastrophising scale) were also measured.

Results: The SLR identified 37 eligible studies (n=3860 participants). Meta-analysis revealed that baseline QST predicted musculoskeletal pain (mean r=0.31, 95%CI: 0.23 to 0.38, n=1057 participants) and disability (mean r=0.30, 95%CI: 0.19 to 0.40, n=290 participants). Baseline modalities quantifying central mechanisms such as TS and CPM were associated with follow-up pain (TS: mean r=0.37, 95%CI: 0.17 to 0.54; CPM: r=0.36, 95%CI: 0.20 to 0.50), and baseline mechanical threshold modalities were predictive of follow-up disability (mean r=0.25, 95%CI: 0.03 to 0.45).

Test-retest and inter-rater reliability were high for PPT and TS in both normal and CLBP populations (ICC=0.76 to 0.92) but low for CPM (ICC=0.43 and 0.46 respectively).

In people with CLBP (n=97), ROC analysis determined that >9/24 painful sites optimally predicted low PPT at the forearm (AUC=0.67, 95%CI: 0.55 to 0.80). The single-factor Central Mechanisms Trait model showed a good fit to the data (CFI=0.92, TLI=0.88; RMSEA=0.09; SRMR=0.07; $x^2(df)=34.19(20)$).

Follow-up questionnaires were completed by 87 people with CLBP (67% female, mean age 65y). Low PPT, inefficient CPM, the ACR and >9/24 classification criteria and the Central Mechanisms trait predicted less positive engagement in life ($r=-0.54$ to 0.31, $p<0.05$), low PPT and inefficient CPM each predicted increased emotional distress (PPT: $r=-0.21$, $p<0.05$; CPM: $r=-0.29$, $p=0.01$), and low PPT predicted worse social integration and support ($r=0.28$, $p<0.01$) at 3 months. Baseline Central Mechanisms trait scores predicted worse performance in health directed behaviour, positive engagement in life, constructive attitudes and approaches, social integration and support and emotional distress at 3-months ($r=-0.56$ to 0.54, $p<0.05$).

In multivariate regression models exploring the relationship between baseline CS indices (QST modalities, widespread pain identification methods, Central Mechanisms trait) and self-management outcomes, adjusted for other variables (age, sex, depression, catastrophisation, pain and fatigue), low PPT, inefficient CPM and Central Mechanisms trait scores, remained significantly associated ($p<0.05$) with social integration and support, positive engagement in life and constructive attitudes and approaches at 3 months respectively.

Conclusion: QST can predict musculoskeletal outcomes across a range of musculoskeletal conditions and discrete pain hypersensitivity indices (PPT and TS) demonstrate high reliability as pain quantification tool. Baseline indices of high CS can predict reduced ability of individuals with CLBP to self-manage their condition 3 months after commencing a CBT-based group physiotherapy intervention. Self-management is a multidimensional concept and its influence by factors other than CS merits further research. Treatments which specifically target CS might help remove barriers to self-management in people with CLBP.

ACKNOWLEDGEMENTS

I am immensely thankful to my parents for offering me the opportunity to pursue my dreams and providing me with endless support throughout the years to persist and achieve anything I set my mind to. I would also like to thank my two supervisors Professor David Walsh and Dr Paul Hendrick for their persistence to excellence, their support and invaluable advice throughout this project. Their constant input has greatly benefited me personally, academically and professionally. Special thanks to Professor David Walsh for introducing me to the field of pain management, which gave me the opportunity to expand my clinical skills, and for inspiring me to seek a career as a clinical/academic.

I would like to share my gratitude to Dr Daniel McWilliams for his kindness, his humour and for being willing to advise me since the first day of my PhD project.

Special thanks go to the clinical and administration staff of the Back Pain Unit, particularly to Will Vogt, Mary Martin, Wendy Furnell and Debbie Wilson for creating an excellent working environment and for assisting me with this project whenever needed.

A lot of appreciation for those numerous individuals who, in their own way, have willingly or unknowingly contributed physically and mentally to the successful completion of this project.

Finally, I would like to thank Antonia Papadopoulou for her endless love, support and understanding throughout this experience. She came into my life when she was mostly needed and helped me to see this project through.

“This body holding me reminds me of my own mortality

Embrace this moment, remember

We are eternal, all this pain is an illusion”

Maynard James Keenan

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ACC	Anterior Cingulate Cortex
ACR	American College of Rheumatology
AUC	Area Under the Curve
BMI	Body Mass Index
BPS	Biopsychosocial
CBT	Cognitive Behavioural Therapy
CCC	Concordance Correlation Coefficient
CFA	Confirmatory factor analysis
CFI	Comparative Fit Index
CI	Confidence Interval
CLBP	Chronic Low Back Pain
CMT	Central Mechanisms Trait
CNS	Central Nervous System
CPCI	Chronic Pain Coping Inventory
CPM	Conditioned Pain Modulation
CPR	Clinical Prediction Rule
CSRI	Client Service Receipt Inventory
CS	Central Sensitisation
DFNS	German Research Network on Neuropathic Pain
DNIC	Diffuse Noxious Inhibitory Control
EQ-5D-5L	Quality of Life Questionnaire
EQVAS	Quality of Life Visual Analogue Scale
FABQ	Fear-avoidance Beliefs Questionnaire
FM	Fibromyalgia
fMRI	functional Magnetic Resonance Imaging
FMSS	Fibromyalgia Pain Severity Scale

FSS	Fatigue Severity Scale
FSVAS	Fatigue Severity Visual Analogue Scale
GP	General Practitioner
HADS	Hospital Anxiety and Depression Scale
HCUQ	Healthcare Utilisation Questionnaire
heiQ	Health Education Impact Questionnaire
heiQ-CAA	Constructive Attitudes and Approaches
heiQ-ED	Emotional Distress
heiQ-HDB	Health-Directed Behaviour
heiQ-HSN	Health Services Navigation
heiQ-PEL	Positive Engagement in Life
heiQ-SIS	Social Integration and Support
heiQ-SMI	Self-monitoring and Insight
heiQ-STA	Skill and Technique Acquisition
IASP	International Association for the Study of Pain
ICC	Intraclass Correlation Coefficient
IQ	Inter-quartile
IQR	Inter-quartile Range
kPa	kiloPascal
LBP	Low Back Pain
LoA	Limits of Agreement
LOC	Locus of Control
MCID	Minimum Clinically Important Difference
MDT	Multidisciplinary
MWU	Mann-Whitney U Test
NHS	National Health Service
NICE	National Institute of Clinical Evidence
NRS	Numerical Rating Scale

NSAIDs	Non-steroidal Anti-inflammatory Drugs
OA	Osteoarthritis
OR	Odds Ratios
PAG	Periaqueductal Grey
PCS	Pain Catastrophising Scale
PD-Q	painDETECT Questionnaire
PDS	Post-traumatic Stress Diagnostic Scale
PFC	Prefrontal Cortex
PhD	Doctorate of Philosophy
PNE	Pain Neuroscience Education
PPT	Pressure Pain Detection Threshold
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSEQ	Pain Self-efficacy Questionnaire
PT	Physiotherapy
QST	Quantitative Sensory Testing
QUIPS	Quality in Prognosis Studies Risk of Bias Tool
RA	Rheumatoid Arthritis
RMSEA	Root Mean Square Error of Approximation
RCT	Randomised Controlled Trial
RMDQ	Roland-Morris Disability Questionnaire
ROC	Receiver Operating Characteristic
SCT	Social Cognitive Theory
SD	Standard Deviation
SDM	Standard Deviations of the Mean
SE	Standard Error
SET	Self-efficacy Theory
SMP	Self-management Programme

SMS	Self-management Strategies
SM/SC	Self-management/Self-care
SRMR	Standardised Root Mean Square Residual
STarT-Back	Subgroups for Targeted Treatment Back Screening Tool
SW	Shapiro-Wilk Normality Test
THR	Total Hip Replacement
TKR	Total Knee Replacement
TLI	Tucker-Lewis Index
TNR	True Negative Rate
TPR	True Positive Rate
TS	Temporal Summation
TSK	Tampa Scale for Kinesiophobia
VAS	Visual Analogue Scale
VIF	Variance Inflation Factor
WAD	Whiplash Associated Disorders
WSRT	Wilcoxon Signed Rank Test
WUD	Wind-up Difference
WUR	Wind-up Ratio
>9/24	24-site body manikin classification method

1. BACKGROUND

1.1. Overview

Chronic low back pain (CLBP) is the most prevalent musculoskeletal condition and constitutes a significant health issue associated with huge public health implications (treatment cost) and significant impact on patients' quality of life (loss of employment, isolation, depression, persistent pain and years lived with disability) (Vos et al., 2016).

Definite causes of CLBP remain elusive although, consistent research findings have implicated the altered function of pain mechanisms in both peripheral and central nervous systems. It has been hypothesised that repetitive acute spinal injuries can cause neuroplastic changes in the central nervous system, which can exacerbate the development of CLBP and the generation of central sensitisation (CS) (a mechanism of centrally augmented pain) (Woolf and Salter, 2000).

The complex clinical presentation of CLBP makes it a notoriously difficult condition to treat effectively despite the large array of available treatments (Liddle et al., 2009). Self-management strategies (SMS), as approaches that aim to provide individuals with skills and abilities for long-term management of the condition without being reliant on others (family members and health-care providers alike), have been considered an essential component of an holistic package for the management of CLBP (NICE, 2016). Self-management interventions have been shown to be effective in increasing quality of life (Lorig et al., 1999) and coping skills such as self-efficacy that help individuals more adequately manage their condition and become more actively involved in the course of their treatment (Damush et al., 2016). However, available SMS for CLBP have demonstrated only moderate long-term benefits in reducing pain and disability levels (Du et al., 2017).

The limited effectiveness of SMS in regards to decreasing levels of pain and disability leads to the development of research exploring the hypothesis of whether pain mechanisms hinder successful self-management in CLBP. CS and self-management are both multidimensional clinical domains with aspects that spread across the entire biopsychosocial spectrum, which constitutes the question of their association logical and of significant scientific as well as clinical value.

This chapter provides the background knowledge that forms the foundation of the project hypothesis by highlighting the complexity of CLBP and how its prevalence can

limit the ability of individuals to live a satisfactory life, the multidimensionality of self-management and how it entails all aspects involved in living with a chronic condition, and the multifaceted role of CS in overshadowing biopsychosocial functions necessary for a successful recovery.

1.2. Chronic Low Back Pain

1.2.1. Description of Low Back Pain

Low back pain (LBP) has been defined as “pain and discomfort, localized below the costal margin and above the inferior gluteal folds, with or without leg pain" by the National Institute for Health and Care Excellence (NICE, 2016) and the European Guidelines for Prevention of Low Back Pain (Burton et al., 2004). All anatomical structures of the lumbar region (spine, muscles, intervertebral discs, and nerves) can be involved in the development of LBP (NIH, 2020) (**Figure 1**). However, in most cases, LBP cannot be attributed to a specific pathology or a recognisable condition, which makes the condition difficult to manage.

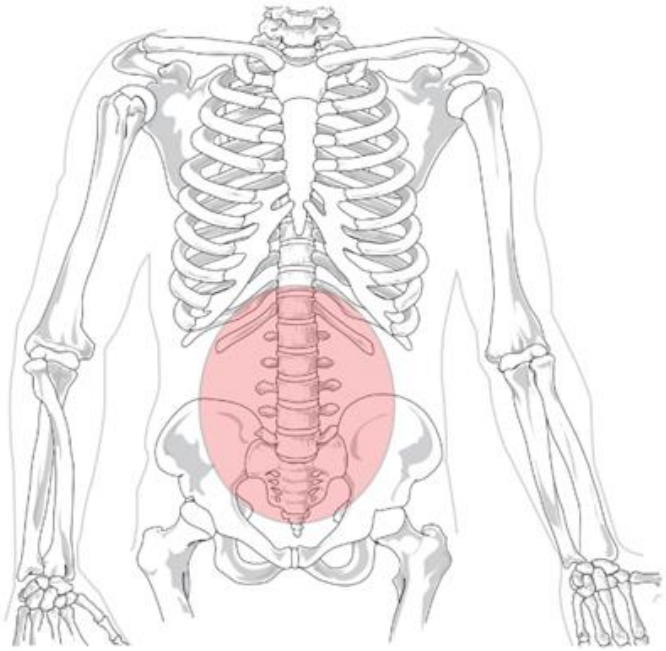

Figure 1. Main area of pain or discomfort manifestation in individuals diagnosed with LBP

Modern definitions of LBP are along the lines of the most widely accepted definition of pain as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage“ (IASP, 1979) and are no longer in agreement with outdated concepts where the experienced back pain was necessarily the result of tissue damage.

Low back pain can be caused by a number of underlying pathologies (intervertebral disc prolapse, nerve root irritation, spondylitis, spondylolisthesis, inflammation, infection, fracture, cancer) with varying levels of severity (Manusov, 2012) and its perpetuation or transition to chronicity is often the result of a complex interaction between multiple co-existing physical and psychological factors such as occupational duties, obesity, sedentary lifestyle, smoking, lack of exercise, depression and anxiety (Ramond et al., 2010, Koes et al., 2006). Neurophysiologically, repetitive acute injuries are considered able to drive neuroplastic changes in the central nervous system, which can promote the alteration of pain processing and the transition from acute to chronic LBP (Woolf and Salter, 2000).

1.2.2. Demographics of Low Back Pain

Despite the common presentation of LBP, its incidence is difficult to calculate. That is common among musculoskeletal conditions due to their highly episodic nature and prevalence of chronicity, which can be barriers to reliable calculations of ‘true’ incidence (the first ever episode) (McBeth and Jones, 2007). LBP frequently occurs in the third decade (20-30) of life and its overall prevalence increases until approximately the sixth or seventh decade (50-70) (Hoy et al., 2010a).

It is anticipated that approximately 80–85% of people will experience an episode of LBP at least once in their lifetime (WHO, 2003). Recent research is along those lines as it is currently expected that 70-85% of the general population will experience at least one acute LBP episode (Buragadda et al., 2018, Adnan et al., 2017). Almost half the amount of people (40-50%) suffering from acute LBP will see their symptoms persist at 3-months with marginal or no improvement and approximately 60-70% of those who will experience improvement will suffer another episode within a year (May, 2010). Ultimately, 62% of individuals experiencing a LBP episode for the first time are considered at risk of developing CLBP (Adnan et al., 2017). Women appear to be more prone in the development of CLBP as, at any point in time worldwide, 16% of females in contrast to 11% of men are suffering from the condition (Hartvigsen et al., 2018, Bridges, 2012). Even though estimations and calculations appear to be

consistent across time periods and regions, the significant diversity in the definitions or classifications of LBP needs to be considered before arriving to conclusions about prevalence, prognosis and impact of the condition. Also, cultural differences among ethnic groups or health-care provision among nations could seriously influence case ascertainment and epidemiologic inference.

Low back pain, acute or chronic, comprises almost half the amount (46.3%) of all prevalent musculoskeletal conditions, with a steady increase in incidence over a 10-year period (2005-2015) and a well-demonstrated socioeconomic burden (Vos et al., 2016). Between 1990 and 2013, LBP was one of the top 25 conditions causing Years Lived with Disability (YLD) globally whereas, specifically in 2013, LBP was the leading cause of disability in developed countries of Europe and America (Vos et al., 2015). Low-income and developing countries appear to be more affected by LBP as 54% more cases were reported over a 25-year period (1990-2015) (Hartvigsen et al., 2018). LBP in its chronic form is the biggest and most prevalent occupational barrier in high-income countries with approximately 2-5% of the working population being temporarily or permanently kept off-work (Hoy et al., 2010b).

The prevalence of LBP leads to a significant worldwide financial burden. Treating LBP is globally more costly than other diseases with significant socioeconomic burden such as cancer, cardiovascular diseases and mental health problems (Dagenais et al., 2008, Maniadakis and Gray, 2000).

In USA, the total cost of LBP-related treatment in 2012 alone was estimated from $84.1 to $624.8 billion (Gore et al., 2012). In the UK specifically, a 2012 evaluation has determined that the annual cost of treatment associated with LBP was £12.3 billion (Bridges, 2012). LBP affects also the clinical working hours devoted to its management as it has been shown that approximately 7% of consultations in general practice (GP) are for LBP resulting in a total loss of 4.1 million days annually (Parsons et al., 2011).

1.2.3. Management of Low Back Pain

Low back pain management constitutes a clinical domain rich in diversity of opinions and intervention strategies. Over the years, different schools of thought have attempted to aid decision making regarding management and numerous treatment approaches have been proposed for the management of the condition (O'Sullivan, 2012, Sahrmann, 2010, Delitto et al., 1995, McKenzie, 1981).

Acute and chronic LBP appear to necessitate distinct treatment pathways for appropriate management. The goals of acute LBP management are pain reduction, return to work, restoration of normal function and return to normal activities following simple exercises within the pain limits (Delitto et al., 2012). Treatment modalities in the acute stage can include mobilisation, spinal manipulation, general exercise, advise, individually tailored exercise, and stabilisation exercises (Liddle et al., 2009). CLBP is considered exceptionally challenging regarding its clinical management and involves multidisciplinary rehabilitation strategies with treatment modalities that include low-graded exercises, pacing, advice, cognitive behavioural therapy (CBT), pain neuroscience education (PNE) and self-tailored pharmacological advice (O'Sullivan et al., 2015, Costa-Black et al., 2010).

Pharmacological approaches also differ according to the stage of the condition (acute or chronic). In acute episodes the aim of medication is to reduce the effect of inflammation and decrease the levels of pain so a quick recovery and return to functionality can be achieved (Miller, 2012). For that reason, the first line of medication is non-steroidal anti-inflammatory drugs (NSAIDs), whereas the use of weak opioids with or without acetaminophen (paracetamol) should be reserved for cases of acute LBP where NSAIDs are contraindicated, cause severe side-effects or have been found to be ineffective (NICE, 2016). The consumption of paracetamol alone has been shown in a systematic review to be no more effective than placebo (Machado et al., 2015). In all cases, dosage has to be self-tailored and according to individual needs or levels of effectiveness and tolerance as strong doses of NSAIDs increase the risk of stomach, liver, kidney and heart problems (Enthoven et al., 2016).

In CLBP pharmacological management should focus on addressing any long-standing nociceptive or neuropathic components that contribute to the overall pain experience (Morlion, 2011). The use of opioids has been proposed if the pain persists or the patient suffers from CLBP (Chaparro et al., 2014) although, the clinical importance of any analgesic effect is questionable (Shaheed et al., 2016) and prolonged consumption of such medication carries the risk of addiction, adverse interaction with other drugs, and various side-effects (Miller, 2012). Certain side-effects such as decreased reaction time, cloudy judgment, and drowsiness can significantly impact on the capacity of individuals to undertake treatment, leading to early discontinuation of pharmacological approaches or treatment modalities that require adequate cognitive function by as many as 35% of the patient population (Rosomoff and Rosomoff, 1999). NICE (2016) have advised against the routine use of opioids in the management of chronic pain. Serotonin reuptake inhibitors,

serotonin–norepinephrine reuptake inhibitors or tricyclic antidepressants, antiseizure (gabapentinoids) and oral steroids have demonstrated some usefulness in managing neuropathic pain (pain caused by a lesion or disease of the somatosensory nervous system) but the evidence supporting their effectiveness in reducing pain in the context of CLBP are considered insubstantial and their use should only be implemented when there are related co-morbidities (Miller, 2012, White et al., 2011). Epidural injections with steroids and local anaesthetic as well as radiofrequency denervation procedures have been also recommended for the management of CLBP (NICE, 2016) however, they either demonstrate questionable effectiveness or only short term pain relief (Juch et al., 2017, Poetscher et al., 2014, Pinto et al., 2012, Chou et al., 2009).

Surgical interventions (excluding disc replacement, which is not recommended) should be only offered when conservative and pharmacological management have not been effective and only when they are supported by radiographic evidence or are part of well-designed randomised controlled trials (RCTs) (NICE, 2016).

The NICE guidelines for the management of LBP (NICE, 2016) reflect the overall paradigm shift towards a multimodal biopsychosocial approach, which is highlighted by a call to prioritise self-management, abandon ineffective practices, promote education and facilitate adaptation or integration of new skills.

1.3. Self-management

1.3.1. The concept of self-management

Self-management is a complex multidimensional concept encapsulating an individual's ability to effectively manage the treatment needs, physical, social and psychological challenges as well as lifestyle modifications that living with a chronic condition can impose (Barlow et al., 2002). Self-management consists of different constructs that range across physical, mental and social health domains (PROMIS, 2020) and encapsulate key elements such as, day-to-day management of the condition, adequate problem-solving, promotion of healthy lifestyle behaviours, adaptation of social, professional and emotional changes, conscious decision-making, appropriate action-taking, optimal resource utilisation and formation of a partner-like relationship with health-care providers (Hartley and Stockley, 2016, Finestone et al., 2015, May, 2010, Lorig and Holman, 2003, Bodenheimer et al., 2002).

Self-management as a clinical domain represents both belief and behaviour and adheres to specific principles that are described best but not exclusively by the social cognitive theory (SCT) and self-efficacy theory (SET), which underpin self-efficacy, a construct recognised as a key element of self-management (Bandura, 2001, Lawrance and McLeroy, 1986, Bandura, 1986). Efficacy beliefs are considered crucial in a patient's goal-setting and decision making ability and, therefore, the theories focus on increasing levels of self-efficacy and social support in order for them to engage in specific behaviours to manage the pain (or the condition in general) more effectively (Bandura, 1986). Self-efficacy has been consistently associated with future health-related behaviours such as exercise, disability, avoidance and quality of life (Yazdi-Ravandi et al., 2013, Denison et al., 2007, Lorig and Holman, 2003, Asghari and Nicholas, 2001, Bandura, 1986).

Even though self-efficacy is allegedly one of the fundamental constructs of self-management, there are other elements that characterise it as a clinical domain. The Locus of Control (LOC) theory encapsulates the degree to which people in general and patients in particular tend to believe that they are in control of outcomes related to their lives or their condition (strong internal LOC), as opposed to external factors that are beyond their sphere of influence (strong external LOC) (Rotter, 1966). Therefore, people with strong external LOC tend to blame external factors for their situation or mishap. In the field of health-care, this type of belief is mostly issued by decreased knowledge or lack of understanding regarding one's condition, and could lead to increased fear or anxiety, factors that have been negatively correlated with LBP treatment outcomes (Gatchel et al., 2007). Increased external LOC has been found to negatively influence perceived disability and quality of life (Sengul et al., 2010, Cheng and Leung, 2000) whereas increased internal LOC has been associated with enhanced physical function and adherence to recommended treatment in individuals with LBP (Keedy et al., 2014, Ono et al., 2008).

Learned helplessness is another construct underpinning self-management and it entails the belief of an individual (e.g. a patient), who has experienced repetitive aversive or directly painful stimuli, that there is no more control over such situations, reinforcing sentiments of helplessness and inevitability (Abramson et al., 1980). This type of behaviour can lead to the development of fear avoidance, catastrophisation, and clinical depression (Miller and Seligman, 1975), factors that have been also found to enhance disability levels (Gatchel et al., 2007), promote pain sensitisation (Overmier, 2002), and negatively impact on the prognosis of musculoskeletal pain (Vlaeyen and Linton, 2012). Helplessness specifically, has been found to be a

predictor of increased pain intensity in patients with chronic pain (Samwel et al., 2006).

Coping and its association with stress that arises from the perceived inability to cope with condition-related circumstances, which can then lead to subsequent behavioural adaptations is another construct that underpins self-management (Lazarus and Folkman, 1984). This construct highlights the process through which the inability to cope with the pain or with barriers imposed by the suffered condition interferes with the ability of individuals to self-appraise, take responsibility for the management of the condition, solve problems, set goals and engage in health-promoting behaviours (Jensen et al., 1991). A reduced coping ability has been shown to be a predictor of increased pain intensity in patients with neck and back pain (Mercado, 2004) and increased disability in individuals suffering from chronic musculoskeletal pain (Samwel et al., 2006).

Congruence between health-care professionals and patients regarding the severity of pain or the condition as a whole is also considered a construct that underpins self-management (Antonovsky, 1979). High congruence between health-care providers and patients can facilitate informed decision-making, increase knowledge about the given condition and enhance adherence to treatment that balances evidence-based guidelines and personal preferences, leading ultimately to enhanced self-management capacity (Suarez-Almazor et al., 2001). Health-care professionals tend to underestimate the severity and impact of pain in others (lack of congruence), which leads to reduced adherence to treatment on behalf of patients, increased health-care utilisation and adaptation of negative behaviours from different stakeholders (patients, patients' families, health-care providers) that do not promote a collaborative relationship (Solomon, 2001).

The constructs described above are inter-correlated and underpin self-management as they are considered personality traits that define belief and behaviour (Bandura, 1986, Lazarus and Folkman, 1984, Antonovsky, 1979). Available research (Vohs and Baumeister, 2016, Lorig et al., 2006, Lorig et al., 2005, Shekelle et al., 2003, Lorig et al., 2001a, Lorig et al., 2001b, Lorig et al., 1999) implicates all those constructs in the ability of individuals to self-manage their condition, as they collectively underpin tasks, such as the confidence to deal with the biological, social, and emotional aspects of a given condition, which must be undertaken daily in order to live with a chronic pathology (Lorig and Holman, 2003). Miles et al, (2011) in a systematic review demonstrated that different psychosocial and behavioural aspects such as self-

efficacy, depression, catastrophising, fear, knowledge levels, skills and overall motivation were determinants of self-management behaviour (ability to perform necessary tasks) and concluded that their improvement should be a fundamental key of all behaviour change techniques.

The accumulation of existing theories, their underpinned constructs and the evidence supporting their relationship with self-management highlight that it is a multidimensional, complex, and dynamic clinical domain that requires continuous development and advancement (Rodgers, 2005). A relatively recent comprehensive model (Ryan and Sawin, 2009) proposes that self-management entails three different dimensions (*context*, *processes*, and *outcomes*) that are further subdivided into separate constructs, which collectively encapsulate self-management and underpin such belief and behaviour:

- Context dimension: risk and protective factors
    - Condition specific factors (condition complexity, condition trajectory, treatment complexity)
    - Physical and social environment factors (transportation and access to care, setting, and provider)
    - Individual and family factors (perspective, literacy, education level, information processing capacity)
- Process dimension: health behaviour change
    - Facilitation of knowledge and beliefs (self-efficacy, outcome expectancy, goal congruence between stakeholders)
    - Enhancement of self-regulation skills and abilities (goal-setting, self-monitoring, reflective thinking, decision-making, planning and action, self-evaluation/self-appraisal, control of negative emotions such as stress/coping and frustration)
    - Social facilitation (influence among peers, support and collaboration between stakeholders)
- Outcomes dimension:
    - Proximal outcomes:
        - Self-management behaviours (condition-specific behaviours, engagement in health-promoting activities, adherence to recommended treatments, patient empowerment)
        - Cost of health-care services (adherence to treatment and engagement in health-promoting activities may or may not reduce resource utilisation and cost)

- Distal outcomes:
    - Health status – disease trajectory (prevention, attenuation, stabilisation, dealing with worsening)
    - Quality of life or wellbeing
    - Cost of health (direct or indirect decrease of health-care utilisation and cost)

According to the above model, constructs within each dimension are intercorrelated and they influence constructs in other dimensions (Ryan and Sawin, 2009). There are demonstrated links between disability and socioeconomic status, access to care, work and educational opportunities (Sawin et al., 2009, Braveman and Gruskin, 2003, Shekelle et al., 2003, Dazinger et al., 2000) as well as between education levels and the ability to manage complex regimens (Sawin et al., 2009, Paasche-Orlow and Wolf, 2007, Riegel et al., 2007, Simons and Blount, 2007, Boldy and Silfo, 2006, Paasche-Orlow et al., 2006, Clark et al., 1991).

Even though knowledge, in and of itself, does not promote change in behaviour (Ryan and Sawin, 2009), acquiring additional knowledge, facilitating social interactions, and enhancing health beliefs are associated with self-regulation, which can further lead to engagement in self-management behaviours and achievement of proximal outcomes (Paasche-Orlow and Wolf, 2007, Warsi et al., 2004, Lorig, 2003, Lorig and Holman, 2003). Equally, contextual and process constructs such as access to healthcare, family relationships, outcome expectancy and collaboration between stakeholders can influence outcomes as the ability to self-manage can facilitate engagement in condition-specific behaviours, which are associated with reduced health-care utilisation and decreased cost (Panagioti et al., 2014, Bodenheimer, 2005).

Despite the strong theoretical foundations of this model, further validation and incorporation of its principles in contemporary qualitative and quantitative research is needed to consolidate those theoretical concepts, identify other underlying dimensions or constructs of self-management and promote service and intervention development.

1.3.2. Self-management in clinical practice

The field of self-management interventions broadly concern the development of strategies and the identification of factors that promote a patient-centric health-care approach (Swendeman et al., 2009, Barlow et al., 2002). According to Du et al.,

(2011) an appropriate self-management programme (SMP), regardless of the condition it aims to manage, must incorporate in its strategy the following essential elements:

i) Increase of self-efficacy
ii) Development of decision-making and problem-solving skills
iii) Self-monitoring of status
iv) Exercise advice and joint protection
v) Enhancement of goal-setting and action-planning abilities
vi) Development of a partner-like behaviour between patients and health professionals
vii) Development of self-tailoring skills
viii) The program must be community-based or taking place at a location close to home.

The effectiveness of interventions that aim to enhance self-management has been the subject of research focus in a variety of non-musculoskeletal pathologies (Nicholas et al., 2016, Damush et al., 2016, Epstein et al., 2015, Richard and Shea, 2011, Morris, 2004). In the musculoskeletal field, the systematic review of Du et al., (2011) revealed that SMPs for chronic musculoskeletal conditions demonstrate small to moderate effectiveness in improving pain and disability in the long term.

Recent research has focused specifically on the effectiveness of SMPs in CLBP and demonstrated that SMPs targeting specifically LBP have moderate effect on pain and small effect on disability in the long term (Du et al., 2017). However, it is unclear whether a moderate statistically significant long-term effect is of equally significant clinical value. Past studies of the same focus and similar findings concluded that self-management does not demonstrate worthwhile long-term effects when compared with minimal intervention (Oliveira et al., 2012).

Such findings might have been influenced by the outcomes that were considered determinants of self-management. Available systematic reviews (Du et al., 2017, Oliveira et al., 2012, Du et al., 2011, May, 2010) have shown that most researchers determined the effectiveness of SMPs based on the improvement of mainly pain and disability and seldom measure of self-efficacy or other behaviours more closely related to the principles of self-management (Lorig and Holman, 2003). Measurements of pain, disability, negative affect (depression, anxiety) and maladaptive beliefs (catastrophising, fear-avoidance) might not adequately reflect

the ability of individuals to self-manage (Nolte et al., 2013). Conversely, qualitative research has shown that patients and other stakeholders (family members and health-care providers) collectively identify increasing applicable knowledge (condition-related or access to information or resources), enhancing independence (physical independence, independence form clinicians, not being a burden to family and/or society), promoting a feeling of 'normality' ('being me', availability of options and choices, maintaining social identity), acquiring condition management skills (managing condition, emotions, stress, empowerment) and optimising emotional (self-efficacy, quality of life), physical (overall health, staying alive, preventing deterioration) as well as social (being useful to family, improve communication skills) health as the most important self-management outcomes that can significantly improve patients' everyday lives (Boger et al., 2015).

Another factor that could have potentially influenced the results of SMP-related systematic reviews is the diversity between what treatment can be classified as self-management intervention. The definition of self-management can vary between research groups and can rely on personal beliefs toward self-management (May, 2010). Future systematic reviews need to set robust eligibility criteria and carefully examine the validity of the self-management programmes under review.

1.3.3. Measurement of self-management

Nolte and Osborne, (2013) have highlighted that self-management programs do not aim to decrease the levels of pain and it is questionable if they can achieve any decrease in the levels of disability. Self-management courses are psychoeducational programs that aim to alter the perception of experienced pain and/or disability and provide patients with the skills to effectively manage their symptoms during episodes (Newman et al., 2004). Boger et al., (2013) suggested that the development of self-management skills should be evaluated via patient-reported outcome measures (e.g. mood, functional status, quality of life, control of symptoms) as this is the only way in which outcomes that are truly important to patients are considered. Therefore, measuring the effectiveness of strategies that aim to increase the self-management capacity of patients with chronic conditions requires the utilisation of outcome measures with a multidimensional focus that simple pain and disability questionnaires might not possess (Nolte and Osborne, 2013).

Measuring the effectiveness of self-management/self-care (SM/SC) interventions is inherently complex as there is significant variability in measurement methods as well

as diversity in what consists self-management (Nolte et al., 2013). Nevertheless, several proxy measures of self-management, such as self-efficacy, coping, health-care utilisation, locus of control, behaviours, goal achievement and health-related knowledge, have been used in research related to chronic pain management (Banerjee et al., 2018).

Self-management interventions primarily aim to address a variety of biopsychosocial and cognitive aspects that can significantly influence the ability of individuals to cope with or recover from their condition. The measurement of such a complex biopsychosocial clinical phenomenon such as self-management by a single measurement tool is particularly challenging, which could be partially attributed to the absence of standardised and widely accepted biomarkers of pain severity (Nolte et al., 2013, Nolte and Osborne, 2013). A recent systematic review (Banerjee et al., 2018) showed that, out of the variety of tools used to capture self-management change in chronic pain research, the Health Education Impact Questionnaire (heiQ) (Osborne et al., 2007) and the Chronic Pain Coping Inventory (CPCI) (Jensen et al., 1995) display psychometric properties that spread across the biopsychosocial spectrum, which is considered essential for the robust evaluation of self-management measures (PROMIS, 2020).

The heiQ has been primarily designed to measure the impact of programs that aim to promote health related knowledge and literacy. It is considered a valid, user friendly and psychometrically sound measurement tool (Banerjee et al., 2018, Schuler et al., 2013) and has been so far translated in more than 20 languages (Elsworth et al., 2015, Epstein et al., 2015). The CPCI has been designed to measure cognitive and behavioural coping, it has also been formally validated in a population with chronic pain (Romano et al., 2003) and has been found to be reliable and valid across cultures (Monticone et al., 2013). Both multidomain tools have displayed high internal consistency (Cronbach's α = 0.70-0.89) with slight superiority of heiQ in regards to measurement of social aspects (social relationships) that the CPCI does not address (Banerjee et al., 2018).

Despite the demonstrated validity of the above tools, both do not aim to measure self-efficacy (Banerjee et al., 2018), even though it is considered an essential component of self-management and self-care (Bandura, 2001, Lawrance and McLeroy, 1986, Bandura, 1986). Based on this limitation, the implementation of the pain self-efficacy questionnaire (PSEQ) (Nicholas, 2007b) as an additional tool for a broader and more clinically comprehensive measurement of self-management could be warranted in

pain-related research. PSEQ has demonstrated good psychometric properties (test-retest: ICC=0.83; internal consistency: Cronbach's α=0.93; construct validity: Cronbach's α≥0.80) (Chiarotto et al., 2016, Asghari and Nicholas, 2009) and cross-cultural adaptation (Lim et al., 2007).

Health-care utilisation has been brought forward as an additional proxy measure of effectiveness of self-managements approaches with conflicting findings (Griffiths et al., 2007, Salisbury et al., 2002, Lorig et al., 1999). Despite the marginal benefits of self-management in pain and disability associated with chronic pain states, a recent systematic review (Panagioti et al., 2014) demonstrated that there is robust evidence on the cost effectiveness of those programs but the review highlighted that cost-related benefits appear to be stronger in programmes for respiratory and cardiovascular conditions than those focusing on musculoskeletal disorders. Health-care utilisation measurements are warranted in self-management research although, they must be read with caution as they mostly rely on primary care records as well as patient self-report of utilisation, which have been both found to be prone to recall bias (Jordan et al., 2006).

1.3.4. Predictors of self-management

Different self-management approaches have demonstrated small to moderate short and long-term effectiveness on pain and disability in individuals with CLBP (Du et al., 2017). What factors influence or predict long-term self-management outcomes have not been effectively and/or robustly studied. Past qualitative CLBP-related research has investigated patients' perspectives on self-management and indicated that pain, among other factors, appears to be a significant obstacle in adhering with treatment such as exercise (Cooper et al., 2009, Morris, 2004). Such findings have been supported also by quantitative research as Escolar-Reina et al., (2009) showed that pain intensity is associated with adherence to exercise regimes or more broadly to rehabilitation programs. However, adherence to exercise and adaptation of health-promoting behaviours might be only one of the dimensions that encapsulate self-management and it is unclear whether pain predicts other psychosocial outcomes relevant to self-management.

Disability poses as another construct that is potentially linked with poor self-management outcomes (Nolte and Osborne, 2013, Oliveira et al., 2012, Du et al., 2011). As with pain, disability is a broad term that can be related with different factors such as pain, pain beliefs, and several other psychophysical drivers such as self-

efficacy (Ferrari et al., 2019). Disability has demonstrated longitudinal associations with self-efficacy (Denison et al., 2007) as well as with functional non-adjustment, an arbitrary umbrella term that encapsulates several biopsychosocial factors including health-care utilisation (Koleck et al., 2006).

Depression and maladaptive beliefs appear to demonstrate stronger longitudinal links with self-management outcomes than reduced physical function (Miles et al., 2011). Baseline levels of depression predict worse self-efficacy to perform work-related duties (Haugli et al., 2003) as well as severe psychological and social limitations (van der Hulst et al., 2008). Generally, individuals displaying the characteristics of depression and/or anxiety are less likely to benefit from self-management strategies as such factors interfere with the ability of people to perform a desired behaviour (May, 2010).

Maladaptive beliefs such as catastrophising and fear-avoidance have also been found to predict poor self-management outcomes (Nicholas et al., 2016, Nicholas et al., 2013). Patients suffering from CLBP with established maladaptive pain beliefs are less likely ($r = -0.29$) to be positively affected by treatment that aims to improve self-management outcomes such as physical functioning (Walsh and Radcliffe, 2002).

Apart from psychological factors (self-efficacy and depression levels, motivation, beliefs etc.), several social (family and social support) as well as environmental (work context, exercise opportunities) factors could act as mediators or moderators in regards to the outcome of rehabilitation programs aiming to improve self-management (Snelgrove and Liossi, 2013). Age, sex, the number of children, job satisfaction and levels of education have been found to longitudinally influence outcomes directly (self-efficacy, quality of life) or indirectly (pain and disability) associated with self-management (Ferrari et al., 2019, Roberts and Allen, 2016, Koleck et al., 2006). Medically-induced cognitive impairment manifested as decreased reaction time, cloudy judgment, and drowsiness due to excessive or prolonged use of strong analgesics could also reduce the ability of individuals to self-manage (Rhee et al., 2009).

Cultural background and cross-cultural differences might also need to be considered as predictors of self-management outcomes. Bandura, (2001) claims that different cultures display different characteristics regarding their efficacy beliefs or management approaches. Culture is considered a very strong individual or collective characteristic that shapes decision-making about treatment, beliefs about pain or self-management as well as behaviour, actions and treatment expectations (Walker et al.,

1995). Culture can also be a determinant of how a person feels as part of a group intervention programme. A culture's individualistic or collectivistic nature can define the outcome of self-management programs as patients from an individualistic culture could feel more efficacious within a program with an individualised and personal orientation whereas patients from a collectivistic culture may be more productive in a group-oriented regime (Bandura, 2001). Self-efficacy levels of healthy individuals has been previously found to be associated with the country of origin of each individual, indicating that culture is implicated in how able people consider themselves to undertake a task or an activity (Earley et al., 1999). Nevertheless, despite the demonstrated links between culture and pain (Rahim-Williams et al., 2012) as well as pain sensitivity (Al-Harthy et al., 2016), existing primary research has not explored any potential links between culture and self-management outcomes in the context of musculoskeletal conditions in general and CLBP in particular.

Despite any existing links in literature with regards to the degree various psychosocial factors predict long-term self-management outcomes, no comprehensive research has been conducted to explore whether different pain mechanisms (nociceptive, neuropathic, central sensitisation) are predictors of self-management and degree of whether certain socioeconomic factors mediate, moderate or confound such a relationship. Pain is the commonest reported symptom of all musculoskeletal conditions (Goldenberg, 2010) and, as a complex clinical phenomenon, it entails cognitive-evaluative, emotional and sensory dimensions (Melzack and Casey, 1968). Self-management is also a complex clinical construct encapsulating physical, psychological, cognitive and social dimensions (Barlow et al., 2002) that might be distinctively or collectively influenced by abnormal pain processing and an augmented pain experience.

1.4. Pain and pain mechanisms

Pain is an unpleasant multidimensional experience sharing both sensory and emotional characteristics that cannot be easily expressed or measured by simple conventional means (Arendt-Nielsen and Yarnitsky, 2009). The International Association for the Study of Pain (IASP, 1979) defines pain as "*an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage*". This definition, despite its potential flaws is the most comprehensive and widely accepted as it incorporates the biological ("actual or

potential damage") and psychological ("unpleasant emotional experience") elements of pain.

1.4.1. Theories of Pain

1.4.1.1. Biological theories

The clinical phenomenon of pain has been the subject of clinical and philosophical debate since classical antiquity. However, theories that tried to underpin or explain pain started to be formulated in the Renaissance and continued in the modern era (19th and 20th century). Initial theories were the 1664's Specificity Theory by Descartes (pain is a specific sensation independent of touch and other senses following a one-way pathway from the periphery/site of injury to the brain) (Antoine-Mahut and Gaukroger, 2017) and the 1874's Intensive Theory by Erb (pain can be generated by any intense enough sensory stimulus or emotion) (Moayedi and Davis, 2013).

These early theories either failed to take into account structures that are activated by nociceptive stimuli (nociceptors) (Intensive Theory) or structures within the central nervous system that can be activated by both nociceptive and non-nociceptive stimuli (Specificity Theory). Both theories epitomised a biological outlook on pain, forming the Biomedical Model of pain, which attempts to explain pain within a dualistic paradigm of separate and independent function between mind and body.

Erb's theory expanded Descarte's outlook by adding the emotional aspect in the concept of pain, which laid the foundation for the proposition of the 1965's Gate Control Theory (Melzack and Wall, 1965). The theory suggested that specific cells at the dorsal horn of the spinal cord (substantia gelatinosa) act as a gate that controls the transmission of impulses from peripheral nociceptors to the brain where they activate central mechanisms responsible for the perception of pain and any subsequent response to it. Even though the theory provided a framework that entailed the combined function of peripheral excitation to external or internal triggers and central processing, it could not adequately explain abnormal pain experiences such as phantom limb pain, which are indicative of processing at a cortical level within the brain.

1.4.1.2. Psychosocial theories

Contemporary pain theories adapt an anthropocentric approach where human psychology is a key contributor to the human pain experience and advocate that every individual needs to receive care according to their own personal needs (Linton and Shaw, 2011). One theory abiding by those tenets is the 1983's Fear-avoidance Theory (Lethem et al., 1983), which considers that avoidance of activities or actions can lead to prolonged and amplified pain-related fear that can subsequently lead to increased pain and disability. Another theory is the 2002's Diathesis-Stress Theory (Turk, 2002), which takes into account predisposing factors of individuals that could facilitate the development of prolonged and amplified pain after an instigating event. Evidence have supported these two theories and have showed the trajectory people with fear-avoidance can follow after an acute pain incident (Leeuw et al., 2007) as well as how pre-existing factors associated with anxiety or stress can facilitate the development of pain catastrophising (Andersen, 2012). However, extensive research yet remains to be conducted to confirm the proposed causal links between avoidance and pain amplification or to elucidate the mechanisms through which anxiety can lead to exaggerated pain perception (Wideman et al., 2013, Sullivan et al., 2011).

A collection of theories labelled together as Social Cognitive Theories suggest that interpersonal social factors need to be considered when attempting to treat individuals in pain as they might be important drivers of abnormal pain responses (Craig, 2009, Norman and Conner, 1996). Ethnic differences and the influence of culture have been also proposed as determinants of pain experience (Campbell and Edwards, 2012, Peacock and Patel, 2008) however, more evidence is needed to support those theories as research demonstrating concrete links between social factors and pain is only scarce (Miró et al., 2019).

1.4.1.3. Biopsychosocial theory

Developments in neuroscience and acquired insight into pain mechanisms has promoted a paradigm shift in the management of musculoskeletal conditions and an adaptation of a different model for the explanation and treatment of pain. In past decades, the biomedical model described above was the paradigm under which most interventions were delivered having, as is widely believed, a significant negative impact on perceived disability as well as on the development of chronicity (Deyo et al., 2009, Dagenais et al., 2008). The biomedical model might have imposed a

contextual barrier on clinicians, preventing the adjustment of their approach according to the biopsychosocial diversity displayed by patients and the adaptation of an evidence-based biopsychosocial framework (Borkan et al., 2002).

The biopsychosocial (BPS) model of health was first suggested by George L. Engel and Jon Romano in 1977 as an alternative to the predominant, biomedical approach (Borrell-Carrió et al., 2004) and proposes an holistic outlook of pain, under which the development of illness is the result of a complex interaction between biological (genetic, biochemical), psychological (mood, personality, behaviour) and social factors (cultural, familial, socioeconomic, medical) that affect, drive or mediate the experience of pain and, subsequently, influence treatment outcomes (Engel, 1977) (**Figure 2**).

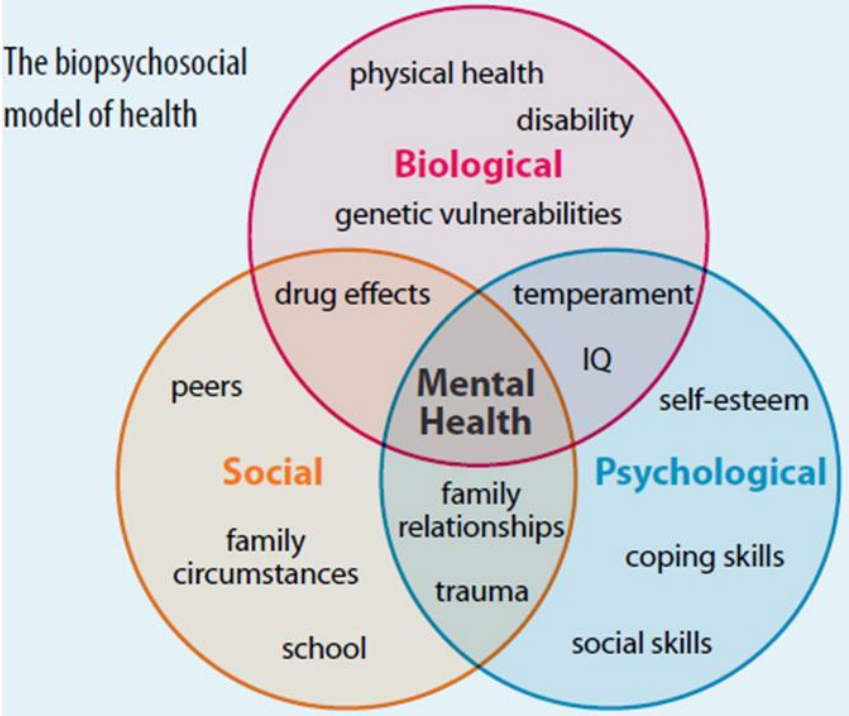

Figure 2. Biopsychosocial Model of Health (https://www.physio-pedia.com/images/c/cb/Biopsychosocial-model-of-health.PNG)

The BPS model operates as a way of understanding the patient's subjective experience of pain and contributes to improved clinical practice by facilitating accurate diagnosis, enhancing health outcomes, and promoting the delivery of humane care (Borrell-Carrió et al., 2004). Application of the BPS model requires the consideration of different factors as well as the cognitive-evaluative and emotional aspects of pain

in order to achieve a more holistic understanding and identify the predominant driver of the pain experience (Wijma et al., 2016). The BPS model has been developed as a platform to express a new philosophy of clinical care as well as a practical clinical guide and is widely accepted and adapted by clinicians and researchers alike (Borrell-Carrió et al., 2004). Nevertheless, even though the model facilitates the identification of the predominant mechanism of pain, it does not provide a mechanistic 'blueprint' of how those biopsychosocial elements are perceived and processed centrally (supraspinal regions), collectively resulting to the generation of the unpleasant experience of pain.

1.4.1.4. Neuromatrix pain theory

The supraspinal regions have been the subject of scientific focus since 1968 with the Neuromatrix Theory (Melzack and Casey, 1968), which suggested that pain is a multidimensional experience (a cognitive-evaluative dimension influencing affective and sensory dimensions of pain) and the result of neurosignature patterns of brain processing that is generated by a complex neural network widely distributed across the cerebral cortex (**Figure 3**). The Neuromatrix Theory highlighted the multidimensional nature of pain and how processing within supraspinal regions plays a key role in pain experience (Keefe et al., 1996) and involves complex cerebral functions other than pain such as cognition, emotion, motivation, and localisation (Ossipov et al., 2010).

The experience of pain alongside other biopsychosocial components could generate increased levels of stress and/or anxiety (Andersen et al., 2016). Therefore, pain can have a leading role in the way feelings or emotions are experienced (cognitive appraisal, fear, happiness and sadness) and, in turn, different emotions (negative or not) can influence the levels of experienced pain (Apkarian et al., 2005). Evidence from functional magnetic resonance imaging (fMRI) studies have demonstrated similar alterations of grey and white matter on brain regions of patients suffering from chronic pain and severe forms of depression or anxiety, suggesting overlapping of pain and emotion processing pathways in the brain (Edwards et al., 2016b).

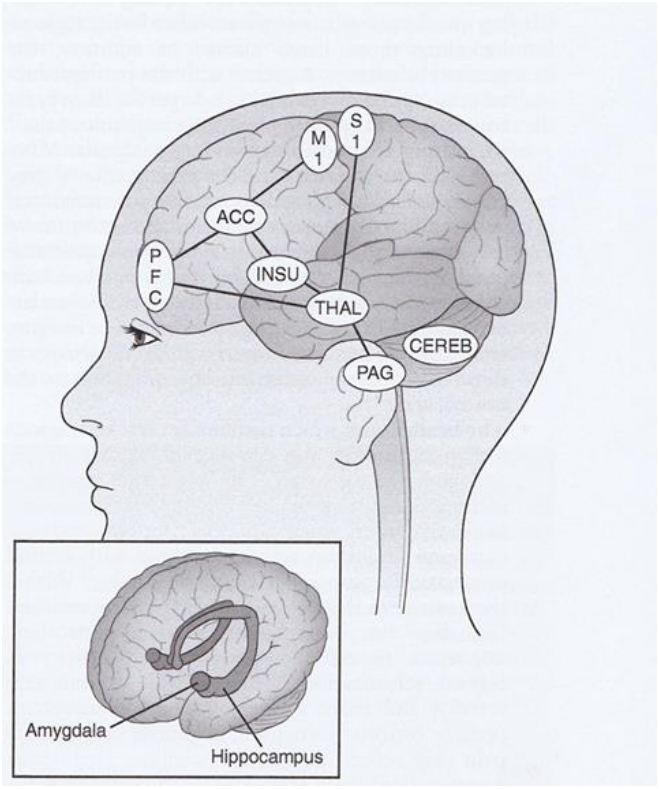

Figure 3. The pain neuromatrix in the brain (Nijs et al., 2015).

ACC: Anterior Cingulate Cortex, **CEREB:** Cerebellum, **INSU:** Insula, **M1:** Primary Motor Cortex, **PAG:** Periaqueductal Grey, **PFC:** Prefrontal Cortex, **S1:** Primary Somatosensory Cortex, **THAL:** Thalamus

Modern pain neuroscience attempts to explain that the perception of pain does not necessarily happen as a response to actual tissue trauma, but it can be the output of active processing in a sophisticated network of neurons known as the brain neuromatrix (Gracely et al., 1994). According to the neuromatrix pain theory, pain perception lies entirely within the brain and it consists also a cognitive instead of just a physical process. Passive registration in the brain of tissue trauma is not enough to produce pain but rather, active generation of subjective experiences (qualia) through a network of neurons (neuromatrix) must occur for pain to be perceived (Chapman, 1996). The patterns of the neuromatrix activity and which brain region will be involved in the final pain output are largely affected by sensory, affective, and cognitive factors that determine the multidimensionality of pain and its related behavioural responses (Melzack, 2001) (**Figure 4**).

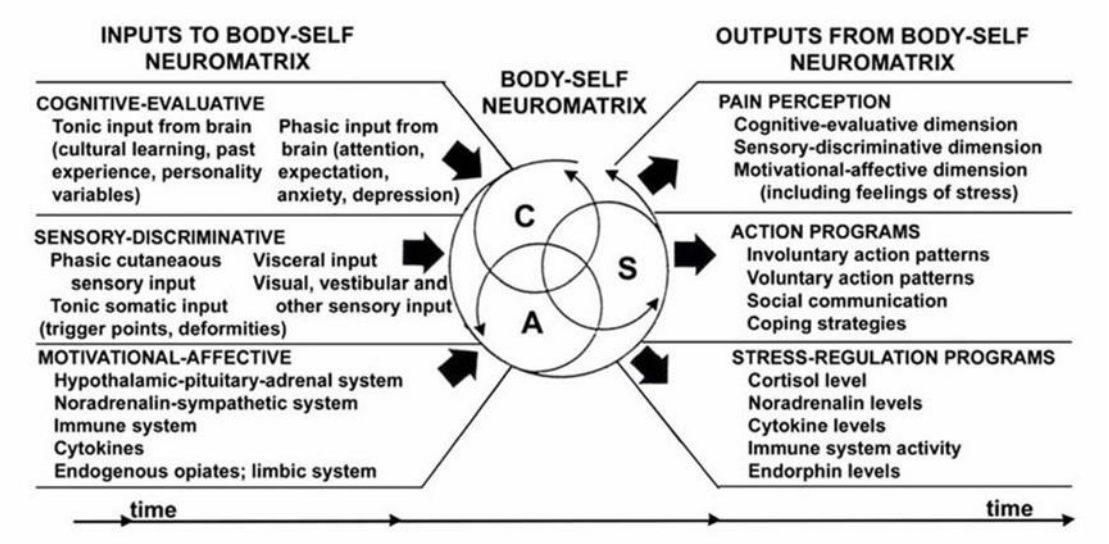

Figure 4. Pain processing within the neuromatrix in the brain (Melzack, 2001).

Several key regions within the brain neuromatrix, due to their neurophysiological function, are primarily or secondarily involved in the generation, modulation, and perpetuation of pain. The amygdala which is responsible for monitoring pain and formulating negative feelings or pain-related memories (Simons et al., 2014, Kattoor et al., 2013, Li et al., 2013), the anterior cingulate cortex which is involved in the affective-motivational aspects of the brain (Peyron et al., 2000), the prefrontal cortex which is responsible for the cognitive-evaluative dimension of pain (Atlas and Wager, 2012, Taylor et al., 2012) and the insula which is responsible for processing the emotional dimension of pain experience (Peyron et al., 2000) have been found to be critical parts of the neuromatrix (Fuchs et al., 2014).

Cerebral function is a key component in the experience of pain and in the production of cognitive responses as a result of that experience. What the Neuromatrix Theory of pain illustrates is that emotional or cognitive malfunction can drive, generate, amplify, or even perpetuate pain in the absence of tissue pathology, a process that probably influences the capacity of individuals to develop or adapt health-related behaviours.

All these theories were developed with the sole purpose of conceptualisation of pain as well as to offer a meaningful explanation about why pain persists beyond expected healing timeframes. These theories were the theoretical background upon which modern-day neuroscience was able to decrypt the pathways that lead to a conscious experience of pain.

1.4.2. Pain processing pathways

A simple model of pain processing for LBP is that afferent sensory nerve fibres located in the peripheral tissues of the lower back transmit sensory information to the dorsal horn of the spinal cord, activating the ascending spinothalamic tract, which then brings the information to the brain where it is processed in the cervical cortex across brain regions (primary and secondary somatosensory cortex, anterior and mid-cingulate cortex, and insula) (Tracey, 2005, Treede et al., 1999, Devinsky et al., 1995, Kenshalo Jr and Isensee, 1983). A more detailed description of pain processing pathways entails 4-stages and can be seen in **Figure 5**.

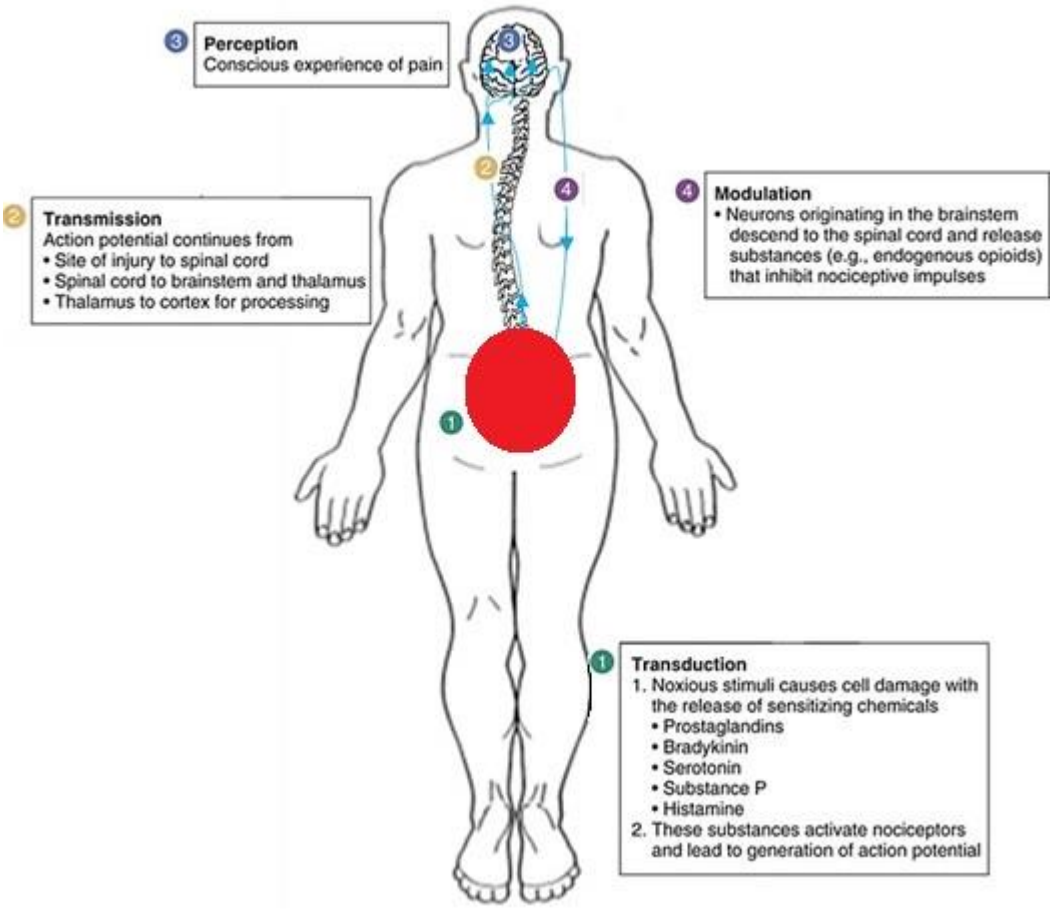

Figure 5. Physiology of pain processing in low back pain (adapted from: https://nursekey.com/20-drugs-used-for-pain-management/)

1.4.3. Temporal classification of pain

The temporal characteristic of pain determines its clinical classification into *'acute'* or *'chronic',* which is entirely based on the length of time it is experienced. Nevertheless, each classification demonstrates its own unique features and characteristics.

1.4.3.1. Acute pain

Acute pain has been further defined as the *'awareness of noxious signalling (nociception) from recently damaged tissue, complicated by sensitization in the periphery and within the central nervous system'* (IASP, 2010). Its experience is generated from the activation of neurophysiological pathways that are usually linked to a specific disease or injury and its intensity changes with the extend of inflammatory processes, tissue healing, and movement (Lavand'homme, 2011). For example, tissue injury in the lumbar region can trigger an inflammatory reaction and facilitate the production of catecholamines (autonomic response), that will sensitise nociceptors locally as well as pathways processing noxious stimuli centrally (sensory response), followed by an instantaneous unpleasant and stressful experience such as pain at the lower back.

In principle, acute pain serves an important biological purpose as it warns individuals for potential or actual tissue damage, therefore, minimising the risk of injury and promoting protective behaviour to avoid reinjury or facilitate tissue healing (Schug, 2011). The duration of acute pain (theoretically 3-6 months) is arbitrarily used to distinguish it from chronic pain which demonstrates distinct characteristics and typically lasts beyond expected tissue healing timeframes (Chapman et al., 2011).

1.4.3.2. Chronic pain

Chronic pain has been defined as *'pain that extends beyond the expected period of healing'* (Merskey, 1986). Nevertheless, the 'period of healing' is not standardised in research and it can significantly vary across conditions. In contrast to acute pain, chronic pain is an unpleasant experience that goes beyond physiological response to tissue damage as it also involves psychological and behavioural mechanisms that can significantly shape its experience and impact (Hasenbring et al., 2001).

Chronic musculoskeletal pain is considered to be the result of a combination of increased tonic impulses from peripheral tissues (prolonged or repetitive nociceptive input to the dorsal horn receptors) and abnormal function of descending facilitation mechanisms in the brain, although the degree of responsibility for each underlying mechanism is inconclusive (Sarzi-Puttini et al., 2011). Contemporary literature suggests that both responsible mechanisms for the generation of pain (central and peripheral) function in parallel rather than in sequence and it is hypothesised that there is a certain degree of overlapping between them (Bourke et al., 2015).

1.4.4. Pathophysiology of chronic pain

The overall pathophysiology of chronic pain implicates peripheral, central, and cognitive mechanisms. Pain is a developing phenomenon that can occur with or without nociceptive input (Bourke et al., 2015, Kidd and Urban, 2001) and consists of a conscious experience that depends upon and cannot happen without cortical activity (Treede et al., 1999). This multidimensional complexity implies that there can be nociceptive signals that may not be perceived as pain and, equally, not all sensations of pain are the result of nociception (Nijs et al., 2015).

Chronic pain can be the final output of progressively induced pathological changes and malfunctioning of pain modulation mechanisms such as descending nociceptive inhibition (Daenen et al., 2013b, Staud et al., 2005, Banic et al., 2004, Price et al., 2002) and descending nociceptive facilitation (Nijs et al., 2015) (**Figure 6**).

High intensity nociceptive input and neuroplastic changes in the central nervous system (brain and dorsal horn in the spine) are implicated in the development of chronic widespread pain (central sensitisation) and have been found to be influenced by genetic, physical, and environmental factors (Phillips and Clauw, 2011), as well as adverse cognitive presentations known as maladaptive beliefs (catastrophising, fear of movement, expectations of treatment outcomes, etc.) (Zusman, 2002, Ursin and Eriksen, 2001).

There are mixed evidence about the true causes of LBP but, over the years, different pathologies that can occur in the muscles and ligaments surrounding the lumbar spine, the nerve roots or ganglia (radicular pain), the facet joints (facet joint syndrome), the sacroiliac joints (sacroiliac joints syndrome), the intervertebral discs (discogenic pain) or in the spinal canal (lumbar spinal stenosis) have been implicated in the development of the condition (Allegri et al., 2016). However, in the absence of a clear underlying pathology, CLBP could be also a state of abnormal pain processing where harmless stimuli are perceived as potentially threatening triggering an unsubstantiated and augmented pain response. From a neurophysiological point of view, CLBP implicates mechanisms in both peripheral and central nervous systems with an equally significant role in the final pain output.

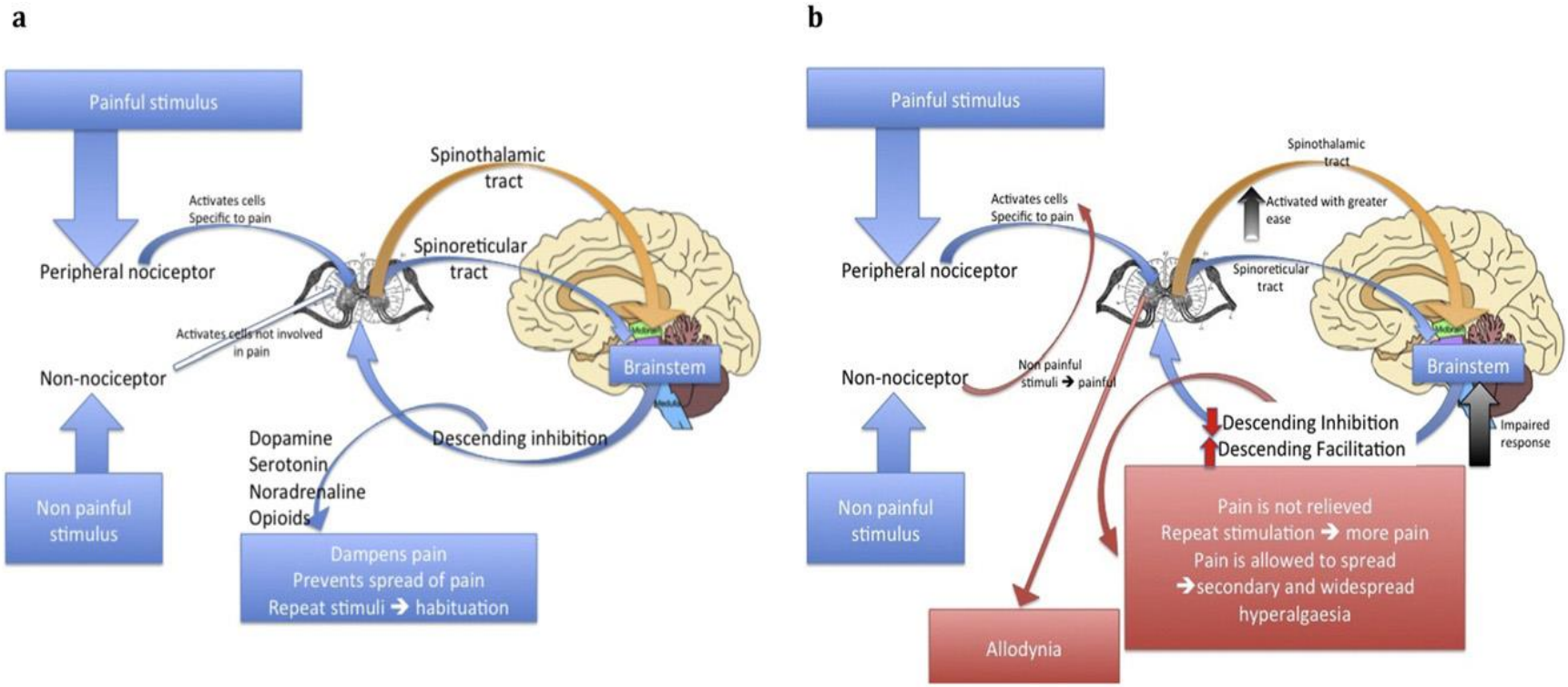

Figure 6. Ascending and descending mechanisms involved in pain experience and control. a: Function of mechanisms in a non-sensitised state. b: Function of mechanisms in a sensitised state (Bourke et al., 2015).

1.4.5. Peripheral Sensitisation

Continued stimulation either because of sustained tissue injury or inflammatory processes can increase the sensitivity of peripheral nociceptive fibres resulting to a reduced threshold of activation and subsequently to a steady influx of ascending action potentials to the central nervous system. Such a phenomenon within peripheral tissues is called 'peripheral sensitisation'. Peripheral sensitisation is defined as the "increased responsiveness and reduced threshold of nociceptive neurons in the periphery to the stimulation of their receptive fields" (IASP, 2017). Nerve lesions anywhere along the nerve can also facilitate the development of peripheral sensitization; as such injuries trigger the production and concentration of inflammatory mediators such as catecholamine, prostaglandins, histamine, serotonin, Tumour Necrosis Factor (TNF), cytokines and Adenosine Triphosphate, that have been found to cause peripheral sensitisation (O'Neill and Felson, 2018, Lechner and Lewin, 2009, Campbell and Meyer, 2006).

1.4.6. Central Sensitisation

Peripheral sensitisation can instigate a constant influx of nociceptive input to the spinal cord (Schaible, 2006). Repetitive acute injuries in the peripheral or the central nervous system can cause neuroplastic changes in the central nervous system, which can facilitate the development of central sensitisation (CS) (a marker of centrally augmented pain) (Woolf and Salter, 2000).

Central sensitisation refers to "increased responsiveness of nociceptive neurons in the central nervous system to their normal or subthreshold afferent input" due to dysfunction of endogenous pain control systems (normal function of peripheral neurons but altered function of central neurons only) (IASP, 2017). The phenomenon of CS may follow the sustained activation of peripheral nociceptive pathways due to an injury (trauma or inflammation) affecting processes such as the transmission and modulation of pain (Fornasari, 2012).

Preliminary evidence of CS was initially presented in animal experiments leading to the development of the first mechanistic models (Woolf, 2011). The first clear evidence of central sensitisation in human volunteers came from a psychophysical study that demonstrated that the secondary mechanical hyperalgesia was dependant on sensory inflow to the central nervous system (CNS) (LaMotte et al., 1991).

Subsequent research on human volunteers provided further evidence about the nature, development and perpetuation of CS. Specifically, the role of different nociceptive fibres (A and C fibres) on the development of CS was established by demonstrating that secondary hyperalgesia to mechanical stimuli is induced by C-fibre nociceptor discharge and is mediated by A-fibre nociceptor excitation (Treede and Magerl, 2000, Ziegler et al., 1999). Klede et al. (2003), highlighted the role of CNS on the mediation of secondary mechanical hyperalgesia by showing that the development of allodynia and punctate hyperalgesia in human skin is centrally mediated, whereas the axon reflex vasodilation is of peripheral origin. Further research on human participants ascertained the specific role of the brainstem on the development of CS as it was shown that only cortical activity, mainly in the primary somatosensory area, was significantly correlated with intensity of pain (Lee et al., 2008). Contemporary research demonstrated the involvement of descending pathways in the development of CS. Rempe et al. (2014) showed that increased spinal activity and decreased activity in supraspinal centres involved in pain modulation during secondary mechanical hyperalgesia suggesting facilitation of nociception via decreased endogenous inhibition.

On a biochemical level, nerve activation (whether by inflammation or not) leads to the release of neuromodulators such as substance-P and pro-inflammatory mediators (bradykinin, prostaglandin) (Samad et al., 2001) as well as to the activation of spinal glial cells associated with neuroimmune function (microglia and astrocytes) (Chacur et al., 2009). Activated glia cells lead to increased production of proinflammatory substances (cytokines), chemokines (Interleukin-8) and various neuro/glial-excitatory substances (p38 MAPK, NF-κB), that contribute to the initiation and maintenance of central sensitisation by further enhancing glial activation and neurotransmitter release at a presynaptic level (microglia), as well as facilitating excitation at a postsynaptic level (astrocytes) (Cao and Zhang, 2008). Ultimately, spinal central sensitisation results (in part) from action of factors such as brain-derived neurotrophic factor (BDNF) released from central terminals of primary afferents to increase responsiveness of second order neurones (Woolf, 2011). Also, increased release of neurotransmitters such as glutamate may be triggered as a direct result of sustained activation of nociceptive mechanisms which, in turn, promotes long-term changes at the level of the dorsal horn (Bourke et al., 2015) contributing to the development of the phenomenon known as CS (Fornasari, 2012). Glutamate displays a mediation of fast excitatory transmission capacity by binding and interacting with both *the alpha-amino-3-hydroxy-5-methyl-4-isoxazoleproprionic acid* (AMPA) and *N-methyl-D-*

aspartate (NMDA) receptors which are largely responsible for the transmission of noxious stimuli from the dorsal horn to the brain (**Figure 7**) (Fornasari, 2012).

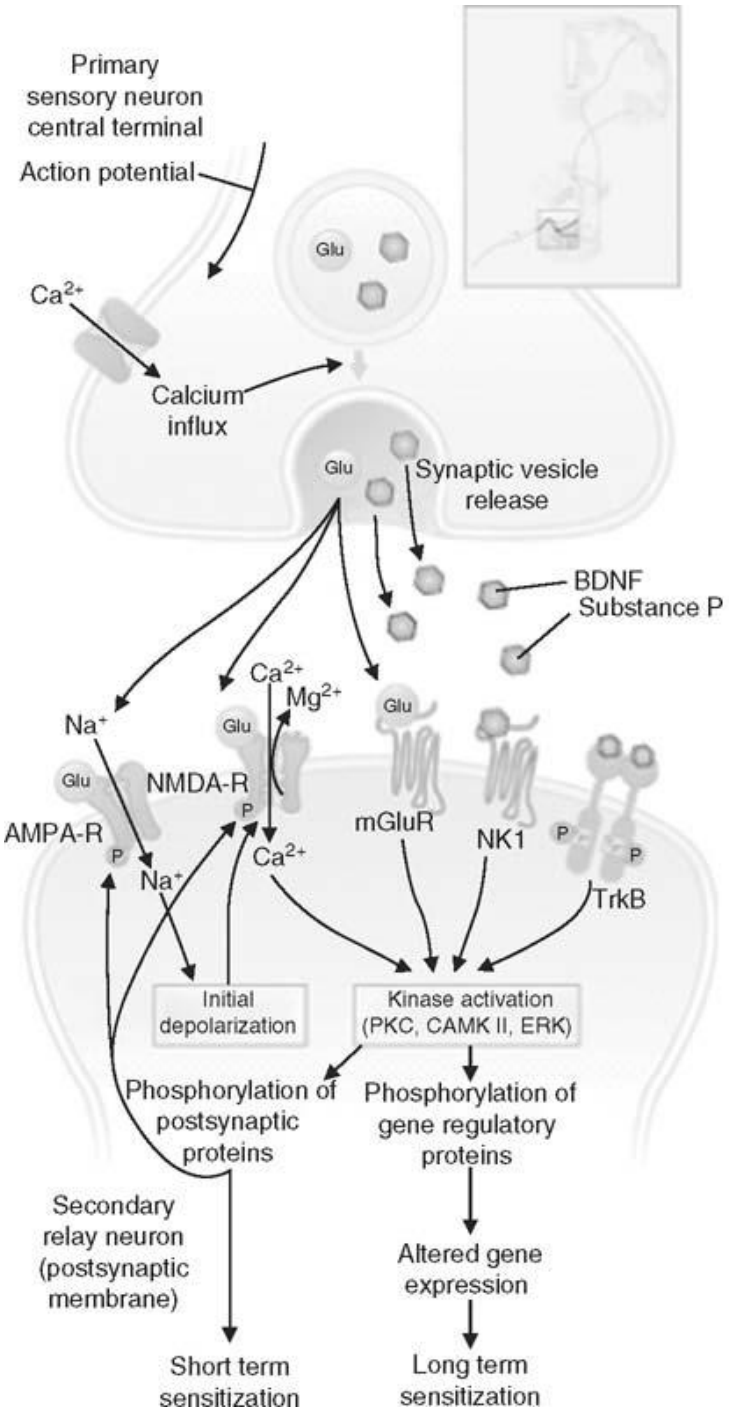

Figure 7. Central sensitization (Fornasari, 2012).

AMPA-R: alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor**, BDNF:** brain-derived neurotrophic factor, **CAMK II:** calmodulin-dependent protein kinase II, **ERK:** extracellular signal-related kinase, **NMDA-R:** N-methyl-D-aspartate receptor, **PKC:** protein kinase C

The sensitisation occurring at a spinal level (dorsal horn) drives the continuous flow of stimuli up to the brain, regardless of them being noxious or not, by maintaining the excitability of spinal pain pathways (Bourke et al., 2015). However, the processes responsible for the development and the continuation of CS are not limited to the dorsal horn or to the amplification of noxious afferent stimuli. Several brain areas such as the insula, anterior cingulate cortex, prefrontal cortex, have demonstrated increased response to standard input in patients with CS as have brain areas that are not involved in acute pain sensations (brain stem nuclei, dorsolateral frontal cortex, parietal associated cortex) (Seifert and Maihöfner, 2009). The brain's capacity for sensory processing demonstrates alterations as well, which can ultimately lead to dysregulation of pain-inhibitory mechanisms (descending inhibition) (Nijs et al., 2015) that are responsible for the reduction of pain sensation (Millan, 2002).

A combination of altered function in ascending (increased activity) and descending (decreased activity) pathways clinically manifests as hypersensitivity to painful (hyperalgesia) and non-painful (allodynia) stimuli (Graven-Nielsen and Arendt-Nielsen, 2002) (**Figure 8**). Additionally, such a combination is responsible for the spread of hypersensitivity in neighbouring dermatomes (secondary hyperalgesia) and at remote sites (widespread hyperalgesia) as well as the development of enhanced temporal summation (stimuli with similar or identical intensity becoming increasingly more painful) (Arendt-Nielsen et al., 2011, Woolf, 2011).

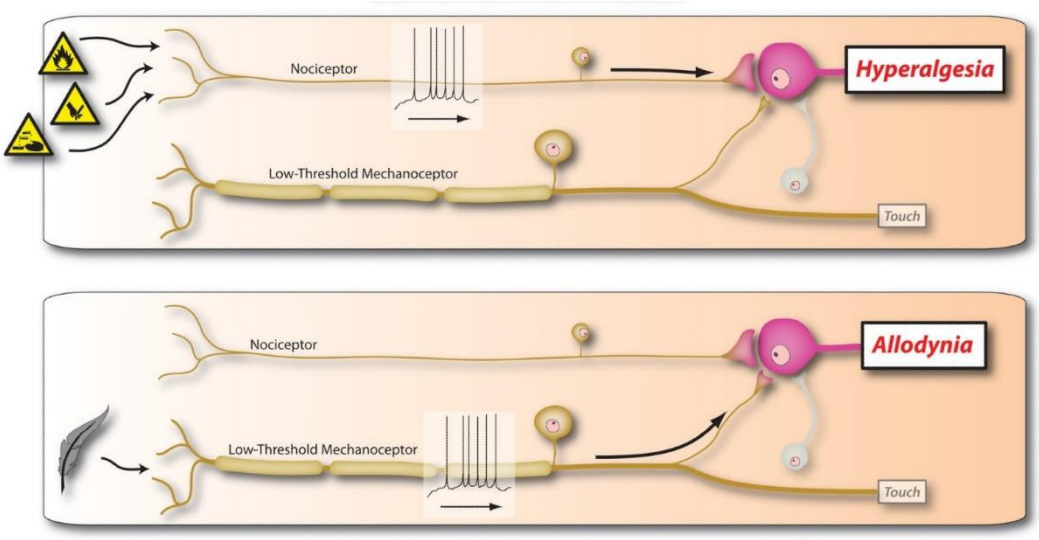

Figure 8. Central sensitization in somatosensory pathways with increased synaptic efficacy and reduced inhibition leads to central amplification which either enhances the pain response to noxious stimuli in amplitude, duration and spatial extent (hyperalgesia), or can strengthen otherwise normally ineffective synapses that can now activate the pain circuit (allodynia) (Woolf, 2011).

Temporal Summation (TS) or wind-up-like pain is an abnormal response to the wind-up effect, which is an electrophysiological phenomenon that occurs over a short time scale (milliseconds) and is considered a feature of CS in terms that it requires central rather than peripheral pain mechanism activity in order to occur (Meeus and Nijs, 2007). Wind-up is a spinal phenomenon and represents increased neuronal response during an array of identical noxious stimuli that returns to normal levels after the end of stimulation. TS is an abnormal response that possibly involves supraspinal processes and is distinguished from wind-up because the elicited painful sensation is sustained autonomously for seconds or minutes after the noxious or innocuous stimulus has come to pass (Woolf, 2011, Petersen-Felix et al., 1996). The presence of TS indicates malfunction of ascending and descending pathways and it is used in research and clinical practice as a measure of 'central gain of pain' (Arendt-Nielsen et al., 2018).

Another neurophysiological expression of altered endogenous descending pain pathway function is the loss or reduction of the 'pain inhibits pain' phenomenon (Yarnitsky et al., 2015, Millan, 2002). Humans with impaired function of descending pathways demonstrate an exaggerated pain response to additional painful stimuli, which might mean either reduced descending inhibition or enhanced descending facilitation (Arendt-Nielsen et al., 2018). Whichever the case, abnormal function of descending mechanisms is associated with widespread hyperalgesia to sites distant to the primary site of pathology (e.g. lower back) (Schliessbach et al., 2013).

Evidence of CS are consistently found in individuals with low back pain, manifested as local and widespread hyperalgesia (Corrêa et al., 2015, Imamura et al., 2013, Giesbrecht and Battié, 2005, Giesecke et al., 2004), elevated TS (Tesarz et al., 2016, Yousef and Al-deeb, 2013), and impaired descending pain modulation (Owens et al., 2015, Mlekusch et al., 2013, Peters et al., 1992). The presence of CS makes the clinical picture of chronic musculoskeletal conditions much more complex (Nijs et al., 2009) and the possibility to achieve positive outcomes with treatment is significantly reduced (Jull et al., 2007). Nevertheless, CS remains a theoretical concept for many clinicians as there is no valid and generally accepted way to successfully diagnose it for patients suffering from a chronic musculoskeletal condition (Girbés et al., 2013).

The pathophysiology of different functional somatic syndromes is not quite well understood and it cannot be conclusively established whether the experienced pain is due to CS (Woolf, 2011). Centrally driven pain hypersensitivity is not present within all patients with chronic pain (Woolf, 2011) rendering identification of those patients

and decision-making for the right management approach even harder (Murphy et al., 2011). Approximately 25% of a population with chronic pain is currently anticipated to demonstrate CS, a percentage that has been shown to be consistent across chronic musculoskeletal pain states (Lluch et al., 2014a, Mlekusch et al., 2013, Sterling et al., 2003a).

Several methods theoretically expressing centrally driven pain hypersensitivity or widespread hyperalgesia beyond the primary area of pain have been proposed. However, identification of CS is a quite challenging clinical approach as there is no gold standard method for that purpose. Nevertheless, the following methods are frequently used in research related to a variety of musculoskeletal disorders such as osteoarthritis (OA), LBP, shoulder pain, neck pain and whiplash (O'Leary et al., 2017).

1.4.7. Quantitative Sensory Testing (QST)

1.4.7.1. Description of QST

Sensory examination is an umbrella term for the assessment, via the utilisation of simple tools (cotton wool, tuning fork, test tubes, coins, tooth picks), of nerve fibres (large myelinated Aβ, small thinly myelinated Aδ and small unmyelinated C) responsible for the perception of touch, vibration, proprioception, pinprick/blunt pressure sensitivity and sensitivity to cold and heat stimuli (Cruz-Almeida and Fillingim, 2014). Assessing some or all of these CNS modalities can provide important information for the differential diagnosis between different pain phenotypes (Arendt-Nielsen and Yarnitsky, 2009). However, bedside neurologic sensory testing demonstrates limitations in regards to the calibration of stimuli intensity and the lack of standardisation of its testing procedures (Backonja et al., 2013).

Quantitative Sensory Testing is an umbrella term for a battery of different modalities (methods of specific tissue stimulation under specific application protocols) that provide important information about different types of pain processing (Courtney et al., 2010, Arendt-Nielsen and Yarnitsky, 2009). QST is widely used to assess how subjects perceive touch, vibration, proprioception, and sharp/blunt pressure sensitivity or sensitivity to cold/heat stimuli (Cruz-Almeida and Fillingim, 2014). The methods can also explore mechanisms responsible for the modulation of pain locally and globally in individuals with musculoskeletal disorders (Courtney et al., 2010, Pavlaković and Petzke, 2010) or quantify sensory alterations in healthy individuals

and patients alike (Rolke et al., 2006a). A comprehensive list of sensory stimuli and QST testing methods is presented in **Table 1**.

The different modalities that collectively constitute the QST approach have been developed based on stimulus properties (stimulus application, stimulus intensity, stimulus summation), response quality of the activated sensation, and the ability to quantify the response intensity (Hansson et al., 2007). Detection threshold modalities are most commonly used for the quantification of sensory alterations (gain/loss) and can be delivered via different types of stimuli (pressure, vibration, electrical, thermal) (Uddin et al., 2016). Pain detection (first point of painful sensation) or pain tolerance (last point of tolerable pain) threshold modalities like pressure pain detection threshold (PPT) are considered "static" and therefore somewhat limited in the assessment of more complex pain-processing pathways as they evoke a single sensation (Arendt-Nielsen and Yarnitsky, 2009).

Threshold modalities are considered an index of abnormal nerve function of the peripheral nervous system whereas other, more "dynamic" modalities, such as conditioned pain modulation (CPM) and TS are operating within a different pain quantification paradigm where different pain-processing pathways are activated (Uddin et al., 2016). Dynamic modalities are considered more appropriate for the measurement of central pain processing and indicative of centrally-driven pain sensitisation as they can test central integration (temporal/spatial summation) and descending control (descending pain inhibition) (Arendt-Nielsen and Yarnitsky, 2009).

Quantitative Sensory Testing has been promoted as an optimal method to explore the involvement of central mechanisms responsible for the development or maintenance of local or widespread pain in various musculoskeletal disorders (Arendt-Nielsen et al., 2015, Uddin et al., 2014, Courtney et al., 2010, Fernández-Carnero et al., 2010, Pavlaković and Petzke, 2010, Sterling et al., 2003a, Staud, 2002). However, central nervous system excitation is not the only factor that can affect QST measurements. Peripheral joint pathology, visceral or superficial tissue damage as well as certain physical (age, gender), social (occupation, family status) and psychological (negative affect) characteristics have the potential to influence responses to different stimuli, constituting such factors important for the interpretation of QST findings (Arendt-Nielsen et al., 2018, Arendt-Nielsen and Yarnitsky, 2009).

Table 1. Modalities, Receptors and Testing Methods (Krumova et al., 2012)

QST Parameter	Laboratory Tests	Clinical Tests	Principal Receptors and Axon Type	Postulated Mechanisms of Hyperalgesia/ Allodynia	Clinical Relevance
Mechanical					
Vibration threshold	Graded tuning fork or vibrometer	Tuning fork	Pacinian, Aβ	Unknown	Lemniscal
Mechanical detection thresholds	Calibrated von Frey filaments	Cotton wool	Aβ		Lemniscal
Punctate pain thresholds	Pin or calibrated sharp metal pin pricks	Toothpick	Unencapsulated, Aδ and C		Spinothalamic
Pressure pain threshold	Algometer	Analogue Algometer, thumb	Intramuscular afferents, iii and iv Aδ and C	Unknown	Spinothalamic
Dynamic mechanical	Brush, cotton wool, Q-tip	Brush, cotton wool, Q-tip	Meissner's Pacinian, hair follicle Aβ and C	Central sensitisation	Lemniscal
Wind-up	Pin prick	Toothpick	Aδ	Central sensitisation, reduced inhibition	Spinothalamic
Thermal					
Cold detection threshold	Computer-controlled thermotester	Thermoroller, test tubes, coins	Unencapsulated, Aδ		Spinothalamic
Warm detection threshold		Thermoroller, test tubes, coins	Unencapsulated, C		Spinothalamic
Cold pain threshold		Thermoroller, test tubes, ice cube, cold pressor test	Unencapsulated, Aδ and C	Central and peripheral sensitisation, reduced inhibition	Spinothalamic
Heat pain threshold		Thermoroller, test tubes	Unencapsulated, Aδ and C	Peripheral sensitisation	Spinothalamic

Quantitative Sensory Testing has been used to ascertain and study pain processing differences between cases and controls, to explore the behaviour of pain mechanisms and to predict or monitor the effectiveness of treatment across a range of musculoskeletal pathologies (Arendt-Nielsen et al., 2018, Pavlaković and Petzke, 2010). QST has been also used to evaluate the effectiveness of pharmacological interventions or to ascertain the involvement of specific pain mechanisms in pain inhibition as different medication have been found to influence different pain mechanisms (Arendt-Nielsen and Yarnitsky, 2009). Acetaminophen (Paracetamol) acts centrally reinforcing descending inhibitory pathways and it may exert an inhibitory action on the enzyme cyclooxygenase in the CNS (Nijs et al., 2011). Opioids such as Tramadol induce analgesic effects through a variety of different targets on the noradrenergic system, serotoninergic system and opioid receptors system (Hitchings et al., 2018). Serotonin reuptake inhibitors (Citalopram, Fluoxetin, Sertraline), serotonin–norepinephrine reuptake inhibitors (Duloxetine) and tricyclic antidepressants (Amitriptyline, Nortriptyline) are thought to modulate pain through action on the central and peripheral nervous systems by primarily altering the neurotransmission of noradrenaline (norepinephrine) and serotonin (Häuser et al., 2012). Anticonvulsant drugs-Gabapentinoids (Gabapentin and Pregabalin) inhibit the function of certain voltage-dependent calcium channels ($\alpha_2\delta$ subunit) by reducing the release of certain neurotransmitters (noradrenaline, serotonin, dopa-mine and substance P) in hyperexcited neurons (Tzellos et al., 2010). Specifically for LBP, QST has been utilised to ascertain its ability to predict the transition of the condition from the acute to chronic stage (Müller et al., 2019); to longitudinally investigate the temporal changes of pain sensitivity (Marcuzzi et al., 2018); to discriminate patients from healthy controls (Tesarz et al., 2016, Owens et al., 2015, Neziri et al., 2012); to confirm the contribution of central mechanisms to the pain distribution (Corrêa et al., 2015, Imamura et al., 2013) as well as to the overall pain experience (Puta et al., 2012) and to identify within-population somatosensory differences (Rabey et al., 2015).

1.4.7.2. Reliability of QST

Calculation and presentation of the reliability of clinical tests when pain tests are used to follow up patients or to investigate the effect of any particular treatment facilitates the translation of QST techniques from the laboratory to a clinical setting (Manresa et al., 2011). The term reliability is used in statistics or psychometric science to express the level of similarity between measurements undertaken under consistent conditions

(Trochim, 2006). When measurements of the same test are taken in a successive manner and under same conditions, the agreement between the results is referred as repeatability or test–retest reliability (BIPM et al., 2008). Inter-rater and intra-rater reliability are two other classes of reliability estimates that respectively measure the degree of agreement among raters and the degree of the consistency between ratings given by the same rater across multiple instances (Hogan et al., 2000). Inter-rater reliability is also known as inter-rater agreement or concordance and measures the variation in measurements when taken by different persons but with the same method or instrument (Hogan et al., 2000). Conversely, intra-rater reliability is a type of test-retest reliability and refers to the self-consistency of a single rater or instrument in the scoring of potentially different subjects or when looking at the same data on multiple occasions (BIPM et al., 2008). The different classes of reliability, when measured appropriately, and bias is minimised, can be used also for validation purposes.

The use of QST is not common practice in clinical settings because its test-retest reliability has not been very well established in patient populations (Wylde et al., 2011b). QST profiles for PPT, TS and CPM are extensively reported on healthy participants regarding their test-retest (Balaguier et al., 2016, Kong et al., 2013, Cathcart et al., 2009, Cathcart and Pritchard, 2006, Nussbaum and Downes, 1998) and inter-rater reliability (Chesterton et al., 2007, Nussbaum and Downes, 1998). In patient populations, reliability is available only for population cohorts of various peripheral or central pathologies (Geber et al., 2011), knee OA (Wylde et al., 2011b) and myofascial pain (Park et al., 2011) whereas for LBP specifically, test-retest reliability has been examined for thermal and pressure pain detection threshold modalities (Paungmali et al., 2012) as well as for the nociceptive withdrawal reflex (NWR) test (Manresa et al., 2011).

In terms of modalities, PPT is widely used in clinical research as a means to assess increased sensitivity in the muscles through applied and gradually increasing pressure (Park et al., 2011). In recent research, TS and CPM are frequently integrated into QST protocols as more relevant studies are conducted and more understanding is gained on pain processing pathways (Arendt-Nielsen et al., 2018). The use of dynamic modalities and the adaptation of mechanistically diverse protocols have highlighted discrepancies in the way modality measurements are calculated and interpreted. Sensory profiling appears to be a well-established and standardised approach for sensation of pain detection threshold modalities (mean value between measurements) whereas TS can be measured either as a ratio (Geber et al., 2011, Rolke et al., 2006a) or as a subtraction between two stimuli (Petersen et al., 2016).

Similarly, different CPM paradigms are adopted among researchers, particularly in regards to conditioning stimuli, which can be unpleasant to participants and therefore affect its reliability by altering how individuals respond to stimuli (Kennedy et al., 2016).

Overall, PPT has been consistently found to be a reliable approach (ICC = 0.75 – 0.94) for sensory testing in healthy and patient populations (Park et al., 2011, Geber et al., 2011) including individuals with LBP (Paungmali et al., 2012). Conversely, TS has been shown to be a reliable method when applied on healthy individuals (ICC = 0.87) (Graven-Nielsen et al., 2015) but its reliability on patient populations has been found to be moderate (ICC = 0.43) (Geber et al., 2011). CPM also displays moderate reliability in healthy participants and patients alike (ICC = 0.57 – 0.59) (Martel et al., 2013, Cathcart et al., 2009). The moderate reliability demonstrated in 'dynamic' modalities (TS, CPM), as opposed to the consistently shown good reliability of 'static' modalities such as PPT, could be the result of heterogeneity in application and calculation methods, which highlights the need for further exploration of QST reliability as well as standardised approaches across populations and protocols (Hall et al., 2015). The development and adaptation of reliable QST protocols is critical to pain-related research and relies heavily on adhering to standardised and widely recognised testing procedures (Arendt-Nielsen and Yarnitsky, 2009).

1.4.7.3. Validity of QST

In the context of scientific testing, the term validation is better referred as test validity, which essentially describes the process of ascertaining whether a test accurately measures what it is supposed to measure or if it can be used as proposed by its developers and/or its users (American Educational Research Association et al., 1999). Nevertheless, validity is considered important in science because it can provide inference regarding appropriate application of tests, can help develop ethical and cost-effective methods, ensures that a method is accurate in measuring the theory or construct in question within similar or across other situations, people, stimuli, and times (Pearl and Bareinboim, 2014).

Available studies do not adequately address QST validity as some of its fundamental aspects are not considered when interpreting findings from sensory measurements. For example, certain QST modalities such as PPT and TS are considered to be tests with good face validity as, conceptually, they are easy to understand as pain measuring approaches by researchers and participants (Uddin and MacDermid,

2016). PPT demonstrates poor correlation with self-reported measurements of pain (Hübscher et al., 2013) indicating low content validity whereas PPT, TS and CPM were found to demonstrate good predictive validity in terms of pain, disability and negative affect (Georgopoulos et al., 2019). However, other aspects of validity (discriminative validity, construct validity) are not so well established or are under-reported. As stated above, QST is a method used for the quantification of sensory alterations and the identification of widespread pain in musculoskeletal disorders. Nevertheless, inferences regarding statistical associations between modalities (internal validity) as well as associations between modalities and alternative methods of hypersensitivity measurements (external validity) are often under-reported.

1.4.8. Pain distribution

Unpleasant sensory experiences in increased duration and spatial extent beyond a clearly defined peripheral driver is suggestive of centrally-driven pain amplification as a result of increased tissue excitation or reduced inhibition (Woolf, 2011). Decreased pain thresholds, allodynia, hyperalgesia, and pain after the end of a noxious stimulus at sites distant to the site of pathology are considered features of CS in many musculoskeletal disorders where central hyperexcitability can be prevalent to approximately 25-30% of the population (Lluch et al., 2014b, Hidalgo-Lozano et al., 2010, Schliessbach et al., 2010, Fernández-Carnero et al., 2009, Freeman et al., 2009). Changes to descending pain control mechanisms have been associated with a wider distribution of pain (widespread pain) in patients with fibromyalgia (Bosma et al., 2016). The American College of Rheumatology (ACR) suggests a combination of number of painful sites (≥7 painful sites across 4 different body regions) and self-reported prevalence of symptoms associated with such type of pain (fatigue, unrefreshed sleep, cognitive difficulties, depression, headaches and abdominal pain/cramps) for the identification of widespread pain (Wolfe et al., 2019).

The number of painful sites (Wylde et al., 2011a, Croft et al., 1996), the use of patient-reporting tools (Mayer et al., 2012), localised pain sensitisation of increased duration (Woolf, 2011) as well as qualitative collection of information about pain characteristics during the examination of patients (Smart et al., 2012b) have been proposed as markers of widespread pain sensitisation, although such classification methods have not been fully validated against QST. Nevertheless, evidence of associations between different methods of widespread pain identification have been demonstrated in individuals with OA, where low pain thresholds at a distant site have

been associated with the number of painful sites, as determined by the pain distribution on a body manikin (Akin-Akinyosoye et al., 2018).

None of the above-mentioned classification methods have been validated in a population suffering from CLBP. Given the presentation of the condition and the frequent prevalence of radicular pain patterns, it is unclear whether such methods could be implemented for classification purposes in relation to such pathology.

1.4.9. Central Mechanisms trait

Pain intensity and markers of central sensitisation have been previously associated with traits of negative affect (anxiety, depression) (Akin-Akinyosoye et al., 2018, Blackburn et al., 2012), maladaptive beliefs (catastrophising) (Akin-Akinyosoye et al., 2018, Campbell et al., 2015), neuropathic-like pain (Akin-Akinyosoye et al., 2018, Hochman et al., 2013), fatigue (Akin-Akinyosoye et al., 2018, Snijders et al., 2011), sleep disturbance (Akin-Akinyosoye et al., 2018, Campbell et al., 2015), distribution of pain (Akin-Akinyosoye et al., 2018), and cognitive impact (Akin-Akinyosoye et al., 2018) in patients with knee pain measured via self-reported outcome measures. High scores in those measures have distinctively predicted poor outcomes after treatments targeted to painful joints (Lluch et al., 2018, Brown et al., 2016, Lluch Girbés et al., 2016, Moss et al., 2016, Campbell et al., 2015, Moreton et al., 2015, Riddle et al., 2010, Sullivan et al., 2009, Harden et al., 2003), which could be indicative that treatments addressing central components of pain processing might be more successful in alleviating symptoms linked to a pathology in peripheral joints.

Contemporary research has found that 8 discrete self-reported items measuring each of the above characteristics, contributed to a single latent ‘Central Mechanisms’ trait which was associated with lower pain thresholds at a distal site cross-sectionally (Akin-Akinyosoye et al., 2018) and with persistent knee pain longitudinally (Akin-Akinyosoye et al., 2020). Given the possible differences in pain processing between musculoskeletal pathologies and the lack of pain sensitivity prevalence across individuals suffering with chronic pain (Woolf, 2011), it is unclear whether a similar trait could be used to predict persistent pain in individuals with CLBP or whether such a trait could demonstrate associations with outcomes other than pain intensity.

1.5. Scope of the project

Low back pain demonstrates a significant impact on society, health-care, and patient quality of life. In its chronic stage (CLBP) it can permanently alter a patients' standards of living and require long-term management approaches which rely heavily on resources and human expertise. CLBP demonstrates significant clinical diversity and patient heterogeneity which renders the establishment of a standardised approach in regards to its management very difficult. Perhaps this is one of the reasons why self-management for LBP demonstrate moderate evidence of effectiveness (Du et al., 2017).

Self-management is a multidimensional clinical concept that aims to alter patients' cognitive function towards their condition, their self and their environment (family and/or clinicians) (Barlow et al., 2002). Judging the effectiveness of self-management based on improvements in levels of pain and disability does not necessarily reflect the ability of individuals to self-manage as such outcomes cannot reflect the degree that patients have adapted their behaviour or how much their knowledge and overall understanding about their condition has changed since the onset of treatment.

Modern neuroscience has shown that centrally-driven pain hypersensitivity, referred also as Central Sensitisation, is also a multidimensional clinical phenomenon with a significant impact on individuals' physical, psychological and cognitive function. CS and self-management are complex but similar concepts as both entail sensory, emotional and cognitive-evaluative processing in the cerebral cortex. CS can affect individuals' emotional state, beliefs, behaviour, expectations and levels of physical activity, which are the same aspects that self-management interventions aim to improve.

To this day, no studies have investigated the potential link between CS and self-management outcomes to establish whether alterations in specific pain mechanisms constitute barriers to effective self-management. Findings from such research have the capacity to have significant impact on differential diagnosis of chronic pain, facilitate patient subgrouping based on the presence of CS and aid the development of more effective self-management approaches in CLBP and potentially in other chronic pain states.

1.6. Hypothesis

The hypothesis underlying this project is that CS, as measured by QST and other indices, can predict long-term self-management outcomes in individuals with CLBP.

1.7. Aims and Objectives

1.7.1. Aims

The main aim of this thesis was to ascertain whether CS indices are predictive of self-management outcomes in a population with CLBP prior to participation in an intervention programme that aims to improve such outcomes.

Secondary aims were to establish whether QST can predict other musculoskeletal outcomes as well as whether it is a reliable and valid tool to identify and measure CS in people with CLBP.

1.7.2. Objectives

- To systematically review the literature in order to determine the ability of QST to predict musculoskeletal outcomes.
- To establish the reliability and validity of distinct QST modalities as classification and measurement tools of CS.
- To establish a cut off for number of body sites shaded on a self-reported pain manikin that best identifies those with widespread pain and explore whether certain self-reported items taken to indicate central mechanisms involvement contribute to a single latent trait in individuals with CLBP.
- To determine whether different CS indices i.e. PPT, TS, CPM, self-reported pain distribution on a body manikin and the Central Mechanisms trait are associated specifically with self-management/self-care outcomes at a single time-point.
- To test whether any cross-sectional associations between baseline CS indices and self-management/self-care outcomes are also present longitudinally, after participants have undertaken an intervention programme that aimed to improve such outcomes.
- To determine whether different baseline CS indices associated equally with different SM/SC measures at 3 months follow-up.

- To explore associations between different self-reported and objective CS indices.
- To identify and explore the relative importance of CS markers compared to other predictors of self-management.
- To explore changes in SM/SC outcomes following participation in a physiotherapy-based group intervention programme and to identify predictors of poor SM/SC other than CS.

2. METHODS

2.1. Summary

Prognosis is an umbrella, single-word, term that encapsulates the scientific act of foreseeing, predicting, or estimating the probability or risk of future outcomes (Moons et al., 2009). In a clinical context, prognosis concerns the probability of an individual (healthy or not) developing a particular condition or achieving a desirable or undesirable outcome over a specific period of time, based on a collection of clinical and non-clinical factors.

The overall aim of this project is to examine whether the presence of centrally-driven pain hypersensitivity in individuals with CLBP can predict self-management outcomes after participation in a group-based pain management programme.

This chapter describes all the methodological steps taken to achieve the project aim. Methods used in each chapter are also briefly summarised in the respective chapter. The study designs incorporated in this thesis that allowed the collection of necessary data are:

1. A systematic literature review exploring the ability of QST (an index of CS) to predict musculoskeletal outcomes in individuals suffering from various musculoskeletal conditions.
2. A reliability study exploring the test-retest and inter-rater reliability as well as the internal and external validity of a QST protocol in healthy participants and individuals suffering from CLBP.
3. An observational study exploring whether CS in volunteers with CLBP undertaking a pain management group programme is associated with self-management/self-care at baseline, prior to participation in a CBT-based intervention aiming to improve self-management outcomes.
4. Analysis of longitudinally collected data to explore whether baseline levels of CS can predict SM/SC outcomes at 3-months follow-up.

Details about each study that contributed data in order to inform the overall results and facilitate the completion of this project are given in **Table 2**. The chapter also describes in detail the analytical procedures undertaken that lead to the synthesis of the results presented in subsequent chapters.

Table 2. Details of studies contributing to the overall results of this project

Results Chapter	Study Sample	Aims	Time points
Systematic Literature Review: capacity of QST to predict musculoskeletal outcomes	International studies featuring adults suffering from any musculoskeletal pathology	To establish the ability of QST to predict long-term musculoskeletal outcomes	Longitudinal musculoskeletal outcomes
Reliability Study: reliability and validity of a QST protocol	Healthy participants and individuals with CLBP	To establish the test-retest reliability, inter-rater reliability, internal validity and external validity of a QST protocol	Baseline – up to 2 weeks follow-up
Observational Study: baseline associations of CS and SM/SC outcomes	Individuals with CLBP undertaking a group-based rehabilitation programme	To establish the cross-sectional associations between CS and SM/SC outcomes	Baseline (start of treatment)
Observational Study: longitudinal associations of baseline CS and follow-up SM/SC outcomes	Individuals with CLBP completing a group-based rehabilitation programme	To establish the prospective associations between CS and SM/SC outcomes	Baseline – 3-months follow-up (end of treatment)

CLBP: Chronic Low Back Pain, **CS:** Central Sensitisation, **QST:** Quantitative Sensory Testing, **SM/SC:** Self-management/Self-care

2.2. Studies outline

2.2.1. QST and predicting outcomes for musculoskeletal pain, disability, and negative affect: a systematic review and meta-analysis

In the musculoskeletal context, QST has been used to ascertain and study differences between healthy subjects and patients, to explore the behaviour of pain mechanisms and to predict or monitor the effectiveness of treatment (Suokas et al., 2012). Musculoskeletal conditions can have a significant impact on society, health-care, and patient quality of life. Most musculoskeletal conditions in the acute stage are successfully managed and patients' health status is restored within an adequate time frame. Conversely, chronic pain, once it is established, may permanently alter individual standards of living and require long-term management approaches which rely heavily on resources and individual expertise of health-care providers.

Latest national guidelines for the management of CLBP (NICE, 2016) have called for the development of novel interventions that promote self-management and facilitate the development of enhanced SM/SC skills. Current literature is lacking the necessary evidence that can longitudinally link QST with musculoskeletal or SM/SC outcomes. Therefore, a systematic literature review is needed to explore the capacity of QST to predict such outcomes and address the literature gap.

The systematic literature review adheres to an a-priori but not publicly registered protocol and was composed under the guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Liberati et al., 2009).

2.2.2. Reliability and validity of a QST protocol in healthy participants and patients with CLBP

Quantitative Sensory Testing has been used to ascertain and study differences between healthy subjects and patients, to explore the behaviour of pain mechanisms and to predict or monitor the effectiveness of treatment (Suokas et al., 2012). Even though signs and symptoms, as clinical observations, might not be relevant or necessarily correlate with mechanism-based findings, it is useful to evaluate and quantify clinical phenomena derived from pain mechanism changes such as

hyperalgesia, allodynia, wind-up, referred pain and tenderness (Uddin and MacDermid, 2016).

Calculation and presentation of the demonstrated reliability and validity of a pain assessment tool is essential when pain tests are used to assess current pain levels, follow up patients in time and investigate the effect of any particular treatment. Determination of the reliability and validity of QST-related approaches is considered a significant step for the successful integration of QST approaches into every-day clinical practice (Manresa et al., 2011).

The reliability study adheres to an a-priori but not publicly registered protocol. Extensive pre-study and intra-departmental training was undertaken (September 2017 to January 2018) to acquire proficiency with QST.

2.2.3. The ability of CS indices to predict SM/SC in individuals with CLBP: an observational cohort study

It is not known, currently, what factors predict effective self-management. Evidence of CS varies between individuals with chronic pain, and may contribute to the relatively poor efficacy of self-management programs. The degree of association between baseline levels of CS and SM/SC outcomes merits exploration as there is a hypothesised link between those two clinical domains.

Recruited participants from the Sherwood Forest Hospitals NHS, Foundation Trust, Primary Integrated Community Services and City Care NHS Trust provided data which were first analysed cross-sectionally for the purposes of addressing objectives within the current thesis. Cross-sectional exploration of available data could allow hypotheses to be tested at the same time-point by bivariate and multivariable approaches and provide inference that could facilitate the progression to prospective exploration of the same hypotheses.

A cross-sectional study can be limited in providing insight regarding the nature of associations between variables as it lacks the temporal element needed for such inference (Shadish et al., 2002). An ideal approach to successfully achieve prognostic inference is via designing longitudinal prognostic studies that implement a multivariable approach in order to determine the most important predictors (or combinations of those) of the outcome under examination (Moons et al., 2009).

A multivariable approach can allow the development of prognostic models, which can be ultimately used in health-care to identify predictors of an outcome or calculate the probability for the prevalence of that outcome (Reilly and Evans, 2006). Multivariable analyses can be also used for designing appropriate models that can lead to the identification of confounders (variables that are associated both with the independent as well as the dependent variable) (Katz, 2011), mediators (often unobserved variables that can better explain an existing relationship as they can be influenced by the predictor variable and then themselves influence the outcome) (MacKinnon, 2012), and moderators (variables that once present can interact with the predictor and the outcome by increasing or reducing the strength of their association) (Cohen et al., 2013).

There appears to be a certain degree of similarity between prognostic and aetiologic studies although, identifying predictors of an outcome does not necessarily imply identification of the cause of that outcome (Brotman et al., 2005). Multivariable approaches can be used in order to ascertain whether a particular outcome is the direct result of a specific (causal) confounder while, at the same time, taking into account other confounders (Moons et al., 2009).

Although a prognostic model may infer causality this should not be the aim of such modelling neither a requirement (Moons et al., 2009). Ascertaining causality is a complex process that needs to abide by rigorously set standards, to which the *strength of association* is only the first of a long list of necessary steps (Schünemann et al., 2011). Therefore, causation, mediation and moderation analyses are beyond the scope of this study.

The methods described in relation to the observational study have been decided upon the sole purpose of establishing whether CS is a predictor of SM/SC outcomes and identifying other potential independent predictors or confounders.

The observational study adheres to an a-priori registered protocol (ClinicalTrials.gov: NCT03972332).

2.3. Study design

- Systematic literature review and meta-analysis: Extensive literature search was conducted to identify studies that can collectively indicate whether CS, when measured via QST, can predict musculoskeletal outcomes.

- Reliability study: Method accuracy and generalisability study. Data collected at two time-points as well as on a single time point by 2 independent raters were used to assess the reliability of the method and its generalisability in another population.
- Prospective observational study: Cross-sectional analyses of baseline data to assess the association of CS indices with SM/SC constructs in individuals with CLBP at the same time-point, followed by analysis of prospectively collected baseline and follow-up data to assess the association of baseline CS indices with follow-up (3-months) SM/SC outcomes in the same population.

2.4. Ethics

No ethical approval was required for conducting the systematic literature review. In regards to the reliability study, the part featuring the healthy participants required ethics application and approval by the Faculty of Medicine & Health Sciences Research Ethics Committee of the University of Nottingham, United Kingdom. The observational study as well as the reliability study part featuring individuals with CLBP required ethics application and approval by the East Midlands - Nottingham 1 Research Ethics Committee of the Health Research Authority, United Kingdom.

2.5. Sample size considerations

Sample size calculations for the test-retest and inter-rater reliability study were identical for both featured study groups (healthy participants and individuals with CLBP). The sample size has been defined according to the minimally acceptable level of reliability ($\rho 0=0.5$), the targeted reliability score ($\rho 1=0.8$) (intraclass correlation coefficient), and the number of repeated measurements (n), based on the assumption that type I and type II errors are fixed within the range of 0.05 to 0.20 (Walter et al., 1998). Considering that the study comprised two different sessions (n=2), featuring one measurement for each modality within each session, and with minimally accepted reliability of ρ0=0.5 and expected reliability of ρ=0.8 (Micalos et al., 2009, Rhudy and France, 2007), the minimum sample size was calculated to be 22 subjects. To account for the potential presence of high variability (spread of QST data), data were collected from 25 individuals for each study group (Manresa et al., 2011).

The sample size calculation for the observational study has been designed specifically for prediction studies using multiple linear regression approaches

(Knofczynski and Mundfrom, 2008). Based on this method, the sample size determination is based on the estimated squared multiple correlation-coefficient (R^2). Past CLBP-related literature indicates that combinations of similar biopsychosocial variables explain 38% to 49% of the variance of SM/SC outcomes (Koleck et al., 2006). It was therefore estimated that at least 40% (R^2=0.40) of the variation of SM/SC outcomes in the study would be explained by a combination of independent variables such as, baseline QST, age, sex, pain, depression, catastrophising and fatigue. Based on these estimates (40%) and the number of independent variables aimed to be included in the model (5 to 7), a sample size of 90 to 120 participants was considered sufficient to power the study and allow reliable regression modelling (Knofczynski and Mundfrom, 2008).

2.6. Materials and tools

2.6.1. Self-management and Self-care self-reported tools

The Health Education Impact Questionnaire (heiQ) (Osborne et al., 2007) is a compilation of 8 self-reported scales (4-point scales ranging from 1-strongly disagree to 4-strongly agree) designed to measure self-management in 8 discrete domains; health-directed behaviour (heiQ-HDB), positive engagement in life (heiQ-PEL), self-monitoring and insight (heiQ-SMI), constructive attitudes and approaches (heiQ-CAA), skill and technique acquisition (heiQ-STA), social integration and support (heiQ-SIS), health services navigation (heiQ-HSN) and emotional distress (heiQ-ED). The heiQ is a reliable tool (test-retest ICC=0.80–0.94) (Schuler et al., 2014) that demonstrates satisfactory discriminant validity (Cronbach's α≥0.80) (Elsworth et al., 2015) as well as high concurrent validity (Cronbach's α=0.88) and internal consistency (Cronbach's α≥0.70) (Morita et al., 2013) with psychometric properties that spread across the physical, psychological and social constructs that define self-management (Banerjee et al., 2018).

The Pain Self-efficacy Questionnaire (PSEQ) (Nicholas, 1994) is a collection of 10 separate 6-point scales (0-not at all confident to 6-completely confident) that has been developed for people suffering from chronic pain. It requires respondents to consider their pain when rating their self-efficacy beliefs on different aspects of daily life. PSEQ is a tool with strong reliability (ICC=0.83) (Asghari and Nicholas, 2009), high internal consistency (Cronbach’s α=0.93) (Lim et al., 2007), responsiveness (Area Under Curve=0.75), unidimensionality and construct validity (Cronbach's α≥0.80) (Chiarotto

et al., 2016), which can be also used as a proxy measure of self-care (Nicholas, 2007b).

Optimal healthcare utilisation is considered an indicator of good self-care capacity (Panagioti et al., 2014). For the purposes of this study, the Client Service Receipt Inventory (CSRI) (Beecham and Knapp, 2001), specifically its version developed for arthritis (Patel et al., 2009), was modified into the Healthcare Utilisation Questionnaire (HCUQ), which aimed to extract information regarding number of condition-specific visits, consultations and hospitalisations (simple sum) as well as type of medication within the last 3 months. The CSRI has shown good concurrent validity (ρ_c=0.63) with other health-care utilisation collection methods (GP records) (Byford et al., 2007) and significant adaptability across conditions and cultures (PSSRU, 2020). Data collection regarding number of visits can provide important inference for conducting economic evaluation of health-care services (Ellard et al., 2017).

2.6.2. Central Sensitisation Indices

A collection of different CS indices proposed by literature have been implemented for the purposes of this thesis featuring distinct QST modalities, pain distribution classification methods and a Central Mechanisms trait comprised by discrete self-reported items.

2.6.2.1. Quantitative Sensory Testing

The QST protocol implemented for the purposes of this thesis features a combination of “static” (pain pressure detection threshold) and “dynamic” (temporal summation, conditioned pain modulation) modalities and was developed according to the recommendations and validated work of international research groups (Arendt-Nielsen et al., 2018, Yarnitsky et al., 2015, Rolke et al., 2006a).

2.6.2.1.1. Pressure Pain Detection Threshold (PPT)

For the purposes of PPT, an electronic data collection unit was used featuring a handheld algometer (Medoc AlgoMed – Computerised Pressure Algometer) connected with a laptop where the amount of applied pressure was displayed on the screen. Increasing pressure with an 1cm-wide rubber probe of the handheld algometer was applied to the participants’ non-dominant forearm (dot on

radiobrachialis muscle) at a rate of 50kPa/sec (Rolke et al., 2006a). Each participant was asked to press a button at a handheld device as soon as the sensation of pressure started to become painful (pain pressure detection threshold). The push-button mechanism was also connected with the laptop and when pressed automatically stored the pressure value (kPa) in the system. The pressing of the button produced a distinct audible sound that alerted the examiner to stop the testing. The push-button device was held by the participant's dominant hand in order to achieve the best reaction time possible. The procedure was initially applied for familiarisation purposes on the dominant arm (forearm) that was holding the button (training site). The process was then repeated a few minutes later on the forearm of the non-dominant arm (testing site) 3 times (Rolke et al., 2006a).

2.6.2.1.2. Temporal Summation (TS)

The TS of pain was assessed on repeated application of a pinprick stimulus (256mN pinprick) delivered via the retractable blunt needle of a specially manufactured pen. The tip was disinfected between individuals with the use of medical disinfectant wipes (2% Chlorhexidine in 70% Alcohol). For familiarisation purposes, TS testing was initially applied on the non-dominant arm (forearm) while participants sat comfortably on the examination plinth. For the actual test (Test A), participants were asked to maintain their relaxed position and a single stimulus with the blunt needle was applied on the participant's dominant forearm (dot on radiobrachialis muscle) followed by ten repetitive stimuli at a rate of 1/sec (Arendt-Nielsen and Yarnitsky, 2009). Immediately after the single stimulus, each participant was asked to rate the experienced intensity of pain or sharpness on a paper copy of a 10cm Visual Analogue Scale (VAS) where the left and right edges of the line were signifying no pain/sharpness and worst pain/sharpness respectively. After the ten repetitive stimuli, they were asked to rate the average intensity of pain or sharpness on the same piece of paper. The process was repeated a few minutes later (Test B), after participants reported that their skin at the test site felt normal to them.

2.6.2.1.3. Conditioned Pain Modulation (CPM)

For the purposes of CPM testing, participants' unconditioned PPT was assessed on their non-dominant arm while holding the push-button mechanism with their dominant arm in an identical way as the one described in the PPT testing. The participants' conditioned PPT was assessed while their dominant arm was compressed by a 15cm

wide pressure cuff similar to those used to measure blood pressure. The pressure cuff was inflated above systolic pressure (to a maximum of 270mm/Hg) in order to occlude arterial blood flow to the arm and thereby generate ischaemic pain (Yarnitsky et al., 2015). Once maximum pressure was achieved, participants were asked to repeatedly squeeze a foam ball of a tennis ball size in order to develop ischaemic pain or discomfort (conditioning stimulus). Upon every five squeezes of the ball, participants were asked to rate their pain or discomfort in the ball-holding hand on a 0-10 Numerical Rating Scale (NRS) with 0 being “no pain” and 10 “the worst pain imaginable”. The conditioned PPT was assessed in their non-dominant arm when participants indicated that their ischaemic pain or discomfort in their dominant arm reached 4 out of 10 on the NRS. Immediately after the conditioned PPT, the pressure cuff was released signifying the end of all QST applications.

2.6.2.2. Pain Distribution on a Body Manikin

A paper copy of a body manikin, featured alongside other self-reported outcome measures in the questionnaire booklet, was used to record the participants’ self-reported pain distribution. The manikin was coded topographically, according to the shading of pain in 7 (head (1), left arm (2), right arm (3), left leg (4), right leg (5), abdomen (6) and spinal axis (7)) (Wylde et al., 2011a) and 24 (right arm (1, 5, 9, 11), left arm (2, 6, 10, 12), right leg (3, 7, 13, 17), left leg (4, 8, 14, 18), head/neck (15, 16), abdomen (19), upper back/shoulders (20, 21), chest (22, 23) and spinal axis (24)) (Croft et al., 1996) regions. The manikin was also coded according to the ACR’s widespread pain classification criteria (WP2019) (Wolfe et al., 2019), based on pain shading over at least 4 of the (1) left and (2) right upper limb (shoulder girdle, upper arm, lower arm), (3) left and (4) right lower limb (hip-buttock-trochanter, upper leg, lower leg) and (5) axis (neck, upper back, lower back) regions and experienced across ≥7 out of the 15 possible sites (neck (1), upper back (2), lower back (3), left arm (4, 5, 6), right arm (7, 8, 9), left leg (10, 11, 12) and right leg (13, 14, 15)). A visual representation of the distinct classification methods is given in **Figure 9**.

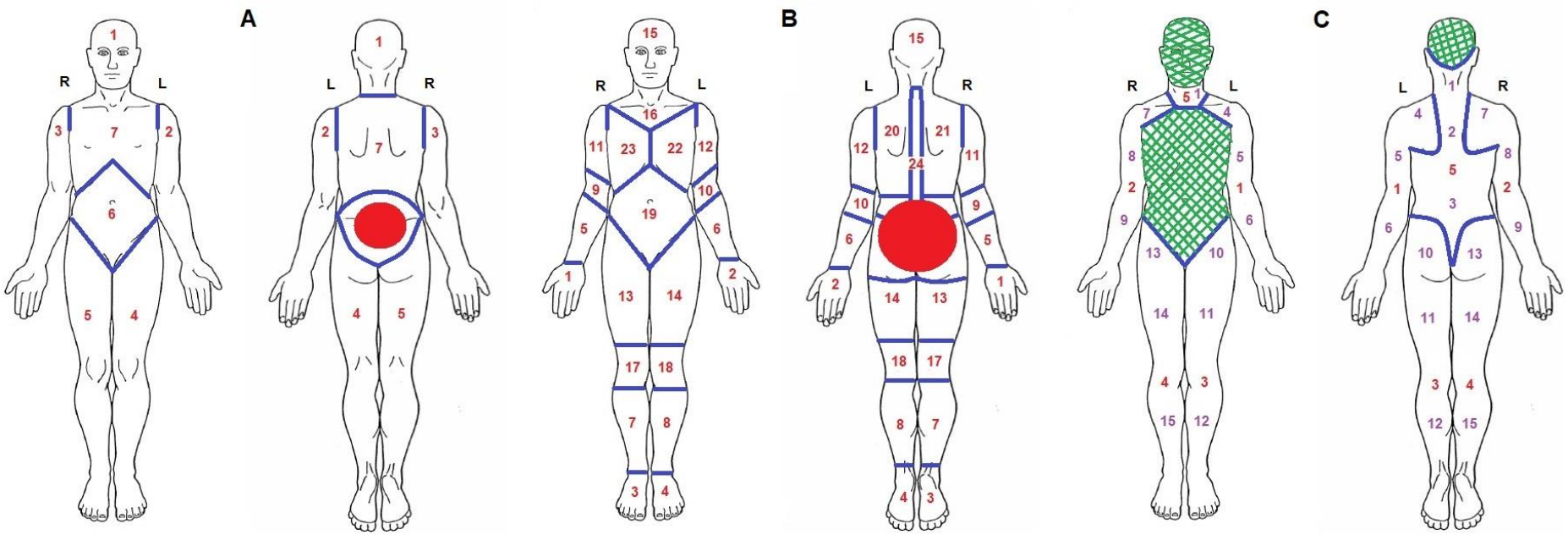

Figure 9. Depiction of discrete diagrammatic manikin scoring based on the distribution of pain on A: 7 anatomical sites, B: 24 anatomical sites and C: the ACR classification criteria

Classifications are made based on the number of painful sites the pain is distributed other than the main area of pain (lower back). The ACR classification criteria divide the manikin into 5 regions (red numbers) and 15 sites (purple numbers).

2.6.2.3. Central Mechanisms Trait

A single Central Mechanisms trait score was calculated from 8 items (**Table 3**) measuring anxiety, catastrophising, cognitive impairment, depression, fatigue, neuropathic-like pain, pain distribution and sleep that have been found in people with knee pain to contribute significantly to one factor, interpreted as “central pain mechanisms” and have demonstrated good internal consistency (Akin-Akinyosoye et al., 2018). Pain distribution classified as “pain in more than 9 out of 24 sites additional to low back pain” was captured using areas shaded by the participant on a body manikin. Item scores from each of the 8 self-report items were used to derive a Central Mechanisms trait score for each participant.

Table 3. Items comprising the Central Mechanisms Trait

Traits	Originating questionnaire	Original format of item
1. Neuropathic-like pain	painDETECT Questionnaire	Is cold or heat (bath water) in this area occasionally painful?
2. Anxiety	Hospital Anxiety and Depression Scale- Anxiety Subscale	I get sudden feelings of panic.
3. Depression	Hospital Anxiety and Depression Scale- Depression Subscale	I still enjoy the things I used to enjoy.
4. Cognitive impact	Fibromyalgia Severity Scale	Please could you indicate your level of concentration problems (forgetfulness and problem solving) severity score over the past week?
5. Catastrophising	Pain Catastrophising Scale	I keep thinking about how much it hurts.
6. Sleep	Roland-Morris Disability Questionnaire	I sleep less well because of my back
7. Pain distribution	-	This question is about **recent pain** you may have had in **any part of your body**. Please shade in the diagram below to indicate where you have suffered **any pain for most days in the previous month.** By pain we also mean aching, discomfort and/or stiffness. Please **do not include** pain due to feverish illness such as flu.
8. Fatigue	Fatigue Severity Scale	Total score

2.6.3. Secondary self-reported tools

2.6.3.1. Pain

Back pain intensity was assessed with an 11-point NRS (Williamson and Hoggart, 2005). Patients were asked to rate their current pain on a scale from 0 to 10 with 0 indicating no pain and 10 the worst pain imaginable. The NRS demonstrates good acceptability by patients with chronic pain (Williams et al., 2000) and shows excellent test-retest reliability (ICC=0.96) and construct validity (Spearman's ρ=0.91) (Ferraz et al., 1990).

Neuropathic components of CLBP were assessed with the painDETECT Questionnaire (PD-Q) (Freynhagen et al., 2006). Participant responses regarding the course, radiation and quality of their pain contributed to a total score (min. 0, max. 38). The higher the score, the higher the likelihood of a neuropathic component. A "neuropathic pain component" is considered unlikely (<15%) with total PD-Q scores of ≤12. A "neuropathic pain component" is likely (>90%) with total PD-Q scores of ≥19. A "neuropathic pain component" is uncertain with total scores of 13 to 18. PainDETECT has demonstrated strong psychometric evidence for the identification and measurement of neuropathic-like pain with high test-retest reliability (ICC≥0.79) (Tampin et al., 2017), high internal consistency (Cronbach's α=0.80) (Cappelleri et al., 2014) and high construct validity (Pearson's r=0.80) (Alkan et al., 2013).

2.6.3.2. Negative affect and maladaptive beliefs

Levels of anxiety and depression were assessed with the Hospital Anxiety and Depression Scale (HADS) (Zigmond and Snaith, 1983). Participants were asked to indicate their levels of emotion frequency on a 4-point scale (0-never, 3-all the time) in 7 statements. The tool is divided into an anxiety and a depression subscale. A sum of all responses for each subscale contributed to a total score for anxiety and depression respectively (min. 0, max. 21). Scores of 0-7 in respective subscales are considered normal, with 8-10 borderline and ≥11 indicating clinical 'caseness'. HADS has demonstrated high test-retest reliability (ICC=0.94), high internal consistency (Cronbach's α=0.88) and high concurrent validity with Beck Depression Index (BDI) (Pearson's r=0.72) and State-Trait Anxiety Inventory (STAI) (Pearson's r=0.75) (Michopoulos et al., 2008).

Catastrophisation was assessed with the Pain Catastrophization Scale (PCS) (Sullivan et al., 1995). Participants were asked to indicate their levels of belief frequency on a 5-point scale (0-not at all, 4-all the time) in 13 statements. A sum of all responses was calculated (min. 0, max. 52). The higher the value, the higher the levels of catastrophising. PCS demonstrates high test-retest reliability (ICC=0.81) (Meyer et al., 2008) as well as high internal consistency (Cronbach's α=0.93) and adequate concurrent validity (Pearson's r=0.42) (Osman et al., 1997).

Kinesiophobia was assessed with the Tampa Scale for Kinesiophobia (TSK) (Miller et al., 1991). Participants were asked to indicate their levels of agreement with 17 questions on a 4-point scale (1-not at all, 3-all the time). A sum of all responses was calculated (min. 17, max. 68). The higher the value, the higher the levels of kinesiophobia. In relation to a population with low back pain, TSK is considered a reliable tool for the measurement of kinesiophobia (test-retest ICC=0.80) and demonstrated high internal consistency (Cronbach's α=0.70) and moderate concurrent validity (Pearson's r=0.59) with the Fear-avoidance Beliefs Questionnaire (FABQ) (Swinkels-Meewisse et al., 2003).

2.6.3.3. Disability and other limiting factors

Levels of disability were assessed with the Roland-Morris Disability Questionnaire (RMDQ) (Roland and Morris, 1983). Participant agreement with 24 statements regarding their ability to perform certain activities (dressing, housework, walking) or functions (sleep) contributed to a total score (min. 0, max. 24). The higher the value, the higher the disability level. RMDQ demonstrates high test-retest reliability (ICC=0.86) as well as construct validity (Pearson's r=0.81) and responsiveness (ROC=0.74 to 0.78) (Davies and Nitz, 2009).

Levels of fatigue were assessed with the Fatigue Severity Scale (FSS) (Krupp et al., 1989). Participants were asked to indicate their degree of agreement with 9 statements on an 8-point scale (0-strong disagreement, 7-strong agreement). A sum of all responses was calculated (min. 0, max. 63). Higher values indicated higher levels of fatigue. Patients were also asked to rate their global fatigue on an 11-point scale with 0 indicating worst and 10 normal fatigue. FSS has demonstrated acceptable test-retest reliability (ICC=0.75) (Learmonth et al., 2013) as well as excellent internal consistency (Cronbach's α=0.94) and high construct validity (Pearson's r=-0.76).

Fibromyalgia severity was assessed with the Fibromyalgia Pain Severity Scale (FMSS) (Wolfe et al., 2016). Participant responses regarding pain location on body manikin, symptom severity at 3 questions about tiredness, sleep and forgetfulness on a 4-point scale (0-no problem, 3-severe) and whether they experienced headaches, depression and abdominal pain amongst 37 other symptoms contributed to a total score (min: 0, max: 31). The higher the overall value the higher the severity of fibromyalgia pain. A simple count of painful sites has demonstrated very good test-retest reliability (ICC=0.71) (Tunks et al., 1995) and the entire collection of criteria has shown good discriminant validity (sensitivity: 64%, specificity:96 %, positive predictive value: 97%, negative predictive value: 56%, and positive likelihood ratio: 16.3) (Usui et al., 2013).

2.6.3.4. Quality of life

Quality of Life was assessed with the EQ-5D-5L (Herdman et al., 2011). EQ-5D-5L consists of two components: health state description and evaluation. Health status is measured in terms of 5 dimensions (5D); mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Participants were asked to rate their level of severity for each dimension using a 5-level (5L) scale (0-no problem, 5-extremely severe). A combined 5-digit value according to responses is composed (11111-excellent quality of life, 55555-severely affected quality of life) and an index value is calculated for statistical analyses with the use of an 'Index Value Calculator' (Van Hout et al., 2012). EQ-5D-5L has demonstrated good construct validity (Pearson's r=0.47 to 0.72) (Nolan et al., 2016), moderate convergent validity (Spearman's ρ=0.34 to 0.57) and moderate discriminant validity (Spearman's ρ=0.24 to 0.45) (Janssen et al., 2013).

2.6.3.5. STarT-Back

Severity was determined with the Subgroups for Targeted Treatment Back Screening Tool (STarT-Back) (Hill et al., 2008). Participant agreement with 9 statements assessing physical (leg pain, co-morbid pain, and disability) and psychological (bothersomeness, catastrophising, fear, anxiety, and depression) predicting factors of poor prognosis contributed to a total score (min. 0, max. 9). Scores of ≤3 indicated a low risk of poor prognosis. Scores of ≥4 indicated a high risk of poor prognosis whereas a subclassification of medium risk could be indicated from score of ≤3 out of the last 5 statements. STarT-Back has demonstrated excellent test-retest reliability

(ICC=0.90), good internal consistency (Cronbach's α=0.73), good predictive validity (AUC: 0.81, 95%CI: 0.78 to 0.84) (Morsø et al., 2013) and high construct validity (Spearman's ρ=0.74) (Bruyère et al., 2014).

2.7. Study participants and recruitment

2.7.1. Reliability study

Healthy individuals affiliated with the University of Nottingham (students, lecturers, senior academics) were invited to participate in the study. They were recruited through posters and personal requests from the University of Nottingham Departments of Academic Rheumatology, Health Sciences, Epidemiology, Psychiatry and Department of Mathematics. Patients were recruited from a list of participants who had already participated in an observational study that aimed to explore the capacity of different CS indices to predict self-management outcomes and had already signed a consent form indicating their desire to be contacted about future studies. At the moment of participation in this reliability study, all participants of the patient group were actively involved or had already completed a group intervention programme aiming to facilitate SM/SC. Written informed consent was obtained from all individuals, healthy and patients, before the beginning of the study.

Inclusion criteria for the healthy group of the study were: adults (≥18 years old) and being currently healthy (defined as people who had no acute or chronic pain at any part of their body). Exclusion criteria were: diagnosed with acute or chronic pain, participation in an NHS rehabilitation program, pregnancy and insufficient understanding of the English language. To be included in the patient group, participants should be adults (≥18 years old), have the ability to give informed consent, be diagnosed with CLBP, have agreed to participate in a group intervention programme at a hospital or community setting (cognitive behavioural therapy (CBT)-based physiotherapy (PT) or multidisciplinary (MDT) group intervention programme featuring a combination of neuroscience education, psychological support, relaxation techniques, pacing, exercise, medication and goal-setting) and be able to speak and understand English.

Patients were excluded from this reliability study if they displayed inability to give informed consent due to cognitive impairment or otherwise, were having history of additional co-morbidities such as cancer, diabetic neuropathies, fractures and/or other conditions causing greater disability than their back pain, were pregnant or if

they had not agreed to be contacted about future studies when consenting for participation in the larger observational study.

2.7.2. Observational study

Individuals with CLBP were enrolled in the study on day 1 of their participation in a group intervention programme, which aimed to facilitate SM/SC. Participants were recruited via hospital pain management units and primary care practices in the Nottingham metropolitan area from May 2018 to August 2019. All recruited individuals were participants in a cognitive behavioural therapy (CBT)-based physiotherapy (PT) or multidisciplinary (MDT) group intervention programme featuring a combination of neuroscience education, psychological support, relaxation techniques, pacing, exercise, medication and goal-setting. Both programmes were delivered in an interactive seminar format where their participation with questions, past experiences and real-life examples was constantly encouraged. Participants, where possible or necessary, were prompted through open ended questions to solve simple day to day problems associated with their condition. A breakdown of events and seminar topics for both programmes is given in **Table 4**. The biopsychosocial aspects of their pain were a-priori measured and evaluated by a group of clinicians unrelated to the present study using self-reported tools and their clinical decision-making skills. Patients were allocated to the MDT programme if they had long-standing CLBP (>1 year), elevated levels of pain, disability and emotional distress and increased maladaptive beliefs. Patients with relatively recent onset of CLBP (>3-12 months), moderate to low levels of pain, disability and emotional distress were allocated to the PT group intervention programme. All participants were allowed to continue their usual care or pursue other management strategies during their participation in the programme.

Inclusion/exclusion criteria for participation in the programmes and in the study were identical, therefore all programme participants were eligible for participation in the study. Specifically, individuals were eligible if they were adults (>18y), had the ability to give informed consent, were diagnosed with CLBP, were enlisted for participation in a pain management programme and were able to speak and understand English. Individuals were excluded if they were pregnant, unable to give informed consent or understand key aspects of the study due to cognitive impairment or otherwise and gave history of additional co-morbidities such as cancer, diabetic neuropathies, fractures or other conditions causing greater disability than their back pain.

Table 4. Set-up, session topics and schedule of events for group intervention programmes

CBT-based Group Intervention Programmes		
	Back in Control	**Back to Fitness**
Programmes Set-up	**Multidisciplinary Intervention Programme**	**Physiotherapy Intervention Programme**
	Psychologists; Occupational Therapists; Physiotherapists; Specialist Nurses; Technical instructors	**Physiotherapists**
	10 days (Full-day)	**5 days (Half-day)**
Sessions	**Seminar Topic and Session Events**	
Day 1	Benefits of Exercise; Tailored Exercises; Values; Pacing	Benefits of Exercise; Tailored Exercises; Pacing
Day 2	Anatomy; Tailored Exercises; Goal-setting; Posture Education	Anatomy; Tailored Exercises; Goal-setting; Relaxation and Pain
Day 3	Explain Pain; Tailored Exercises; Workshop; Pain and Relaxation	Explain Pain; Tailored Exercises; Workshop; Avoidance
Day 4	Avoidance; Tailored Exercises; Values; Gradual Exposure	Meditation; Tailored Exercises; Problem-solving; Sleep Hygiene
Day 5	Sleep Hygiene; Tailored Exercises; Goal-setting; Relaxation	Posture Education; Tailored Exercises; Coping with Flare-ups; Relaxation
Day 6	Meditation; Tailored Exercises; Problem-solving; Medication Advice	-
Day 7	Pain and Emotions; Tailored Exercises; CBT; Relaxation	-
Day 8	Work and Pain; Tailored Exercises; Workshop; Pacing	-
Day 9	Relaxation and Pain; Tailored Exercises; Values; Pacing	-
Day 10	Coping with Flare-ups; Tailored Exercises; Goal-setting; Relaxation	-

2.8. Study procedures

2.8.1. Systematic review

2.8.1.1. Databases search

A systematic online search was conducted in the following databases: CENTRAL, MEDLINE, EMBASE, AMED, CINAHL and PubMed databases from 1948 until April 2018. In the absence of a previously standardised search strategy for QST and musculoskeletal conditions, a unique strategy (**Appendix 1**) was based on previous systematic reviews. The QST elements of the search strategy were adapted from a systematic review on the utilisation of QST in painful OA (Suokas et al., 2012) and the musculoskeletal components were adapted from a systematic review on musculoskeletal intervention and imaging (French et al., 2010). The search strategy was not limited to a specific study design in order to maximise the potential to retrieve relevant studies and because statistical association analysis can be frequently found in randomised controlled trials (RCTs) as well as in prospective cohort studies. A list of the search terms and their combinations that were used in the aforementioned databases is demonstrated in **Appendix 1**. Citation tracking from identified studies as well as from relevant reviews was also employed to maximise the efficiency of the search strategy. No contact of authors to retrieve missing data was attempted.

2.8.1.2. Inclusion/exclusion criteria

Quantitative sensory testing was operationally defined as a method that attempts to measure, in a quantifiable way, responses to sensory stimuli applied on the skin with the aim to be used as an indicator of altered pain sensitivity. Studies that featured QST in their methodology were considered for inclusion in the systematic literature review only if they satisfied the criteria summarised in **Table 5**. All identified studies were downloaded and imported to EndNote X8 (Thomson Reuters) where the duplicates were removed. Two reviewers independently undertook the two-phase screening process for all the identified titles. Phase one (VG and DAW), was the evaluation of the titles and abstracts of the identified studies while phase two (VG and KA) consisted of full text retrieval of all studies deemed eligible for inclusion at the end of phase one.

Table 5. Study eligibility criteria

Inclusion Criteria
1. Prospective studies that had recruited adult participants with any musculoskeletal condition and had used QST to predict a longitudinal outcome. 2. QST modalities used one or more of chemical, electrical, mechanical and/or thermal stimuli applied to skin, muscle or joint 3. Univariate, bivariate or multivariate statistical relationship between QST and outcomes reported, or report data which allow such calculation. 4. QST protocol describes stimulus modality, anatomical site and intensity. 5. Published in English language as an original research article in a peer reviewed journal.
Exclusion Criteria
1. Studies reporting only cross-sectional data. 2. Duplicate publication of data (follow-up analysis of already published data). 3. Books or book chapters, PhD theses or other dissertations, abstracts of conference presentations.

PhD: Doctorate of Philosophy, **QST:** Quantitative Sensory Testing

2.8.1.3. Data extraction from eligible studies

To increase reliability, two independent reviewers (VG and KA) extracted the data from all included studies using a bespoke spreadsheet. Data were extracted on study design, setting, sample selection, length of follow-up, musculoskeletal condition, affected joint or body part, diagnostic criteria, demographic data (mean age, sex and number of participants), pain severity at baseline and at follow-up, stimulus protocol, QST modalities and outcome measures, and the anatomical site of QST. Correlation coefficient (r), regression coefficient (β), odds ratios (OR), area under the curve (AUC) and Chi^2 (x^2) values were collected along with their p-value, standard deviation (SD), standard error (SE) and 95% confidence interval (CI). When data were extracted from a regression model, the prognostic factors that the derived value was adjusted for were also extracted (Hayden et al., 2009). In all cases of disagreement on the extracted data or their interpretation, consensus was achieved through discussion, whenever necessary including all members of the research team.

2.8.1.4. Quality and content assessment of eligible studies

The quality of included studies was appraised by the Quality In Prognosis Studies (QUIPS) Tool (Hayden et al., 2006) for observational cohort studies as well as RCTs. The tool evaluates risk of bias for each study in 6 distinct research domains; (1) Study participation, (2) Study attrition, (3) Prognostic factor measurement, (4) Outcome

measurement, (5) Study confounding and (6) Statistical analyses and reporting, with each domain being qualitatively classified as demonstrating high, medium or low risk of bias (Hayden et al., 2013).

2.8.2. Reliability and Observational studies

2.8.2.1. Clinical assessment and applications for healthy and patient participants

Information regarding age, sex, height and weight were recorded from participants of both the reliability and observational studies as they are considered personal variables that can influence pain sensitisation levels and subsequently QST measurements. Extraneous variables regarding the environment of the examinations (study setting), the room of the examinations (university room, laboratory room), the temperature of the room and the timeframe (time passed between time points) were also recorded.

For the purposes of pain testing in both the reliability and observational studies, before any assessment was conducted, all participants were examined on the forearm of both arms to see if there were any skin or elbow injuries that could preclude the application of equipment to the standard test site. They were asked to adopt a sitting position on an examination plinth with a back support and a pillow on their lap to support their arm. Both arms rested on the pillow with an approximate elbow flexion of 90° and, depending on QST modality, one of the arms held a handheld push-button mechanism essential for the detection threshold modalities.

For the purposes of the reliability study, examinations in both the healthy and patient group were performed by the same researcher (Rater 1). In both groups, each individual was invited to participate in two sessions (baseline/follow-up) separated by at least a week and no more than 15 days. One more session of measurements (n=3) and a second rater (Rater 2) were added at the baseline session of the healthy group to calculate the inter-rater reliability. In that session Rater 1 was the first rater with all participants. For both groups, each session took place in a temperature-controlled room (18-20 °C) and, whenever possible, follow-up measurements were scheduled at a time as close as possible to the baseline measurement time, in order to minimise the possibility of effects caused by circadian variations (Sandrini et al., 1986). All participants were advised and encouraged to maintain their routines and daily activities throughout the testing period. For the healthy group, any regular or urgent

painkiller consumption within 24 hours before the examination was recorded and, in such occasions, appointments were rescheduled to avoid the influence of analgesics on pain testing. As patients were under daily dosage of various analgesic medication, a significant effect on their overall pain sensitivity was not anticipated.

For the purposes of the observational study, a questionnaire booklet featuring all self-reported outcome measures **(2.6.1**, **2.6.2.2**, **2.6.3**) and pain sensitivity assessment methods (**2.6.2.1**) were delivered to participants by the author of this thesis (VG) twice; (1) on the first day, before their commencement of the group intervention programme, and (2) on the last programme day, after they completed their intervention. For the reliability study purposes and particularly for determining the external validity of the different CS indices, data already collected via the questionnaire booklet were extracted from the group of individuals with CLBP only in relation to the intensity (NRS, painDETECT) and distribution of pain (manikin).

Participants who enrolled in the observational study on day 1 of their participation in a group intervention programme (baseline) were asked to undertake an identical clinical examination on the last day (approximately 3 months after day 1) of their intervention (follow-up), featuring QST assessment and completion of the questionnaire booklet. At follow-up, QST assessment was offered to all participants and, to minimise loss to follow-up for primary outcome measures of SM/SC, participants who did not complete the programme or attend for QST had the option to return the questionnaire booklet via mail (Sprague et al., 2003). All participants were eligible for a follow-up QST examination and no minimum amount of attended intervention sessions was required.

2.9. Data synthesis and management

2.9.1. Systematic review

2.9.1.1. Data coding

Coding of the extracted data was conducted by one reviewer (VG) and was validated by one co-investigator (DM). Data from included studies were primarily categorised according to association values featured in the studies; r-correlation coefficients (unadjusted correlation) and β-coefficients (adjusted correlation). All extracted OR values were log-transformed to β-coefficients with the use of RevMan 5 (Review Manager. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane

Collaboration, 2014) and were therefore fitted in the adjusted correlation data cohort. Data were further subdivided according to the musculoskeletal outcome (pain, disability, negative affect) measured in the study and, given sufficient data, separate meta-analysis was conducted for each outcome. When a study used a single outcome measure to observe more than one outcome (e.g. pain and disability), then the data were categorised according to which outcome the tool was prioritising. If there was an equal weight on both outcomes then the data on both subsets were included.

In secondary, exploratory analyses, data were further subgrouped according to study design (RCT or prospective cohort) and QST modalities based on the type of stimulation such as mechanical (pressure/punctate detection or tolerance threshold with algometers, von Frey monofilaments or pinpricks), thermal (cold-hot detection or pain thresholds), movement (pain provocation testing) and electrical (detection or pain thresholds). For the purposes of this review, we further sub-classified QST modalities as 'static' and 'dynamic', with static modalities including pain detection and tolerance thresholds, and dynamic modalities investigating changes in certain mechanisms of pain processing with specialised stimulation (descending pain modulation, temporal and spatial summation) (Arendt-Nielsen and Yarnitsky, 2009). All QST and outcome measurements were extracted at baseline and subsequent follow-ups. If different QST application sites were reported, data were extracted for all and grouped into local, distal and remote to the affected joint sites. As local site was defined the primary area of clinical pain (knee, neck, low back, shoulder, elbow, hip). Distal sites (e.g. tibialis anterior in knee OA) were in the same limb as the musculoskeletal pathology, and distant sites were elsewhere in the body.

2.9.2. Reliability and Observational study

2.9.2.1. Data normality exploration

Data normality was determined via the use of Shapiro-Wilk normality testing in both the reliability and observational studies. In cases of data distribution different to normal ($p<0.05$), logarithmic transformation was performed to achieve, where possible, distribution closer to normal (Keene, 1995). In cases of zero values, the smallest measured value of the dataset (0.1) was added as a small constant to allow logarithmic transformation (Bartlett, 1947).

For the purposes of regression modelling, normality testing of the residuals of distinct multivariable regression models was conducted with the use of the Shapiro-Wilk

normality test where the dependent variable was found to display distribution significantly different to normal after transformation. When model residuals display normal distribution, linear regression modelling should be conducted even if the continuous dependent variable displays distribution significantly different to normal (Katz, 2011).

2.9.2.2. Descriptives

Throughout this thesis, normally distributed data were presented as means ± standard deviations of the mean (SDM) and those found significantly different to normal as medians with their interquartile range (Rolke et al., 2006a). Descriptives regarding patient demographics and baseline as well as follow-up mean/median values of each variable were presented. For the purposes of the reliability study, descriptives were presented separately for each group and subsequently for the combined population. However, for the purposes of the observational study, descriptives were primarily presented for the total population and subsequently in groups according to programme participation (PT, MDT).

In the reliability study, to assess significant differences in QST between the two time points, paired t-tests (for normally distributed data) and Wilcoxon signed-rank tests (for not normally distributed data) were conducted (Rolke et al., 2006a). Additionally, unpaired t-tests were conducted to examine whether there were significant differences in the findings between the 2 groups of participants (healthy participants and individuals with CLBP). In the absence of significant differences, the populations were combined (n=50) and all reliability analyses were repeated for the combined population to explore whether ICCs and correlation-coefficients could be affected by the sample size of each population group (n=25).

In the observational study, significant differences were assessed between the two time points (baseline – follow-up) as well as between participants who undertook QST assessment and those who did not through paired and unpaired Wilcoxon signed-rank tests respectively (Rolke et al., 2006a). Group differences on each outcome were also established in the observational study with the use of independent 2-group Mann-Whitney U Tests (Rolke et al., 2006a).

The Minimum Clinically Important Difference (MCID) was established for the overall population as well as for each intervention group separately based on the baseline SDM for each primary and secondary outcome (Copay et al., 2007). As MCID was

defined a follow-up increase or decrease for each primary and secondary outcome of at least one half (±0.5) of the baseline SDM of that outcome, which corresponds with the limit of the human mental discriminative capacity and consistently appears across most patient self-reported outcomes (Norman et al., 2003).

2.9.2.3. Data management and interpretation

For the purposes of both the reliability and observational studies, the QST application protocol consisted of a battery of three different assessment approaches, as well as calculation and interpretation methods for each modality. These were implemented identically on both study groups (healthy and patients) of the reliability study and on both time-points (baseline – follow-up) of the observational study for the entire participant cohort.

PPT was measured in kiloPascals (kPa) signifying the amount of pressure needed for a participant to report pressure as painful. Participants' PPT was taken as the arithmetic mean of 3 replicate measurements. A low PPT value indicated more pain sensitivity. Similarly, CPM measured the difference (in kPa) between and unconditioned stimulus, which was taken to be the arithmetic mean of the unconditioned replicate PPT measurements (PPT^{Mean}), and a single PPT measurement taken under the application of a conditioning stimulus (PPT^{Con}). Participant CPM ($CPM^{PPT\text{-}mean}$) was determined as the difference between the conditioned and unconditioned PPT values (PPT^{Con} - PPT^{Mean}) (Yarnitsky et al., 2015, Yarnitsky, 2010). Increased sensitivity during the conditioning stimulus (values located in the 1st quartile) was considered as inefficient CPM while reduced sensitivity during the conditioning stimulus (values located in the 2nd to 4th quartile) was considered efficient (Marcuzzi et al., 2018). TS pain was calculated as a difference (subtraction between the score of the single stimulus and the average pain experienced during the ten subsequent stimuli). Each value was used as the index of TS, expressed as a wind-up difference (TS^{WUD}). The average of the two TS values (Test A and Test B) was considered the mean overall TS of the individual. A larger positive value of TS indicated increased sensitisation.

For the purposes of the reliability study, before establishing the above definitive QST calculation methods, alternate means of calculation were explored for CPM and TS. An interim PPT stimulus was used as the unconditioned stimulus ($PPT^{Unc)}$ before the application of the conditioned stimulus (PPT^{Con}). Participant CPM was again determined as the difference between the conditioned and unconditioned PPT values

(PPT^{Con} - PPT^{Unc}) although, this time, CPM was calculated with the use of PPT^{Unc} as unconditioned stimuli (CPM^{Unc}) instead of the PPT^{Mean}, which is the standardised method given above ($CPM^{PPT\text{-}mean}$). To ascertain whether there is a meaningful difference between the different PPT measurements (PPT^{Mean}, PPT^{Unc} and PPT^{Con}) and explore whether implementation of an interim unconditioned stimulus is redundant, the Kruskal–Wallis one-way analysis of variance and a paired Wilcoxon signed-rank test were employed.

An alternate methods of TS calculation was also employed. Temporal summation pain was calculated as a difference (subtraction between the score of the single stimulus and the average pain experienced during the ten subsequent stimuli), which is the standardised method given above, and as a ratio (the average pain experienced during the ten stimuli divided by the rating of the single stimulus) (Rolke et al., 2006a). The ratio value was also used as the index of TS, expressed as a wind-up ratio (TS^{WUR}). As per the standardised method, the average of the two ratios (from Test A and Test B) was considered the mean overall TS^{WUR} of each participant. A larger positive value of TS^{WUR} indicates increased sensitisation.

For the purposes of internal and external validity analyses, the TS^{WUD} and $CPM^{PPT\text{-}mean}$ were used as the standard calculation method. A summary of the different calculation methods for each QST modality is presented in **Table 6**.

2.9.2.4. Missing data

Missing data, whenever present, were dealt with multiple imputation (MacDonald, 2002) to achieve valid statistical inference (Schafer, 1997).

2.9.2.5. Multicollinearity assessment

Independent variables of all combinations featured in regression models were tested for multicollinearity with variance inflation factor (VIF) testing. For interpretation purposes, a VIF value of 1 indicated no correlation between the independent variable, a value between 1 and 5 indicated moderate but no severe correlation, and a value greater than 5 indicated critical levels of multicollinearity, which would then require for that variable to be removed (Katz, 2011, Glantz et al., 1990).

Table 6. Standardised and alternative calculation methods for QST modalities used in quantitative synthesis

QST Calculation Methods				
	QST Modalities	**Measurement Unit**	**Stimuli**	**Method**
Standardised Method	**PPT^{Mean}**	kPa	PPT	Arithmetic mean of 3 replicate measurements
	TS^{WUD}	0-10cm	TS-Single, TS-10 repeated	Wind-up calculated as a difference: Subtraction between the score of the single stimulus and the average pain experienced during the ten subsequent stimuli
	$CPM^{PPT\text{-}mean}$	kPa	PPT^{Con} - PPT^{Mean}	Difference between the conditioned (PPT^{Con}) and unconditioned PPT values (PPT^{Mean})
Alternative Method	**TS^{WUR}**	0-10cm	TS-Single, TS-10 repeated	Wind-up calculated as a ratio: Average pain experienced during the ten stimuli divided by the rating of the single stimulus
	CPM^{Unc}	kPa	PPT^{Con} - PPT^{Unc}	Difference between the conditioned (PPT^{Con}) and an interim unconditioned PPT stimulus (PPT^{Unc})

CPM: Conditioned Pain Modulation, **kPa:** kiloPascals, **PPT**: Pain Pressure Detection Threshold, **TS:** Temporal Summation, **WUD:** Wind-up Difference, **WUR:** Wind-up Ratio

2.10. Data analyses and analytical procedures

2.10.1. Systematic review

2.10.1.1. Meta-analysis

Forest and funnel plots were developed from pooled data of comparable studies by using a random effects model of analysis in R (meta package, R Core Team 2017, Austria). To increase the sample size of the model and allow for more rigorous analysis, data were pooled for meta-analysis only where there were at least 3 eligible studies. When a single study reported both correlation and regression coefficients, both values were used in separate meta-analysis models. In cases where a single study reported more than one result from the same analytical approach (unadjusted correlation or adjusted correlation), the stronger association value was incorporated into the model (Hübscher et al., 2013). If associations were of similar strength, the statistically significant value ($p<0.05$) was preferred for analysis, usually indicating the larger numbers of participants. In situations where unadjusted and adjusted β-coefficient values were reported only the adjusted values were incorporated in the models (Altman, 2001). When multiple associations from the same study were statistically significant, the one related to the most clinically relevant aspect of the outcome (worst pain or pain with movement) was included. Single studies that examined the relationship of QST and outcome in 2 different musculoskeletal conditions where subdivided into two separate studies for the purposes of statistical analysis and were included in the same model.

Pearson's r or Spearman's ρ were included in the same models and were z-transformed during the analysis to normalise the sampling distribution of unadjusted correlation (r) and decrease the bias of the average correlation (Corey et al., 1998).

2.10.1.1.1. Heterogeneity testing

Given the variability in study design, QST modalities, and follow-up outcome measures, statistics to test the null hypothesis of statistical validity (Cochran's Q test) and to quantify the percentage of variance attributable to study heterogeneity rather than chance (I^2 statistic) were calculated and reported for each forest plot (Higgins et al., 2003). Where statistically significant, heterogeneity was determined by an I^2 with an associated p-value of <0.1. I^2 values of 25% were considered as low heterogeneity,

of 50% as moderate, and of 75% as high (Higgins et al., 2003). As per study protocol, a cut-off I^2 value of 50% to perform subgroup analysis was a-priori considered, with subgroups defined based on methodological quality, QST application site (local versus distal or distant), musculoskeletal condition, QST protocol and QST modality. Based on levels of heterogeneity and where there were sufficient data, further post-hoc exploratory analyses were permitted.

The post-hoc analyses reported here explore relationships between baseline QST and follow-up pain according to different QST stimulus within specific modalities, site of clinical pain (axial or peripheral), study design and studies that in their regression models had adjusted for baseline pain. Data were converted when necessary to ensure that higher numerical QST values reflected greater sensitivity.

2.10.1.1.2. Publication bias assessment

To assess for publication bias funnel plots were developed and to assess funnel plot asymmetry Egger's test was performed (Egger et al., 1997). Judging overall risk of bias for each study is recommended where judgments can be made within a specific context such as developing clinical practice guidelines (Higgins et al., 2011) or for undertaking sensitivity analyses (Hayden et al., 2013). The overall risk of bias for both study designs was determined to allow combined subgroup analyses according to levels of bias (high, moderate, low). Study confounding and appropriate statistical analysis were a-priori set as the most important domains for QUIPS (Hayden et al., 2006). The likely magnitude and direction of bias was considered for an overall judgement whenever there was a different measurement of bias between domains of the same tool (Higgins et al., 2011). All discrepancies were discussed between the two reviewers (VG and KA) and the overall risk of bias was determined via consensus.

For interpretation purposes, the strength of any unadjusted association was considered little or zero, fair, moderate to good and good to excellent when r values were between 0.00 to 0.25, 0.25 to 0.50, 0.50 to 0.75 and >0.75 respectively (Portney and Watkins, 2009).

The criteria for exclusion from the meta-analysis were the absence of unadjusted or adjusted correlation data.

2.10.2. Reliability study

2.10.2.1. Test-retest and inter-rater reliability testing

The test-retest reliability of the PPT, TS and CPM was established using a variety of methods that focused on the measurement of reliability (Manresa et al., 2014). For each separate modality, a two-way random effects absolute agreement model was used to measure the inter-rater reliability (Rater 1 and Rater 2) for the healthy group as well as the test-retest reliability for a single rater (Rater 1) for both study groups. A single measures intraclass correlation coefficient (ICC) was reported to express each reliability. The ICC values can range from 0 to 1, which indicate no correlation and perfect correlation respectively. For interpretation purposes, reliability values of 0.00 to 0.25 were taken to suggest no to little correlation, 0.26 to 0.49 low correlation, 0.50 to 0.69 moderate correlation, 0.70 to 0.89 high correlation and 0.90 to 1.00 very high correlation (Portney and Watkins, 2009). ICC calculations were based on the assumption that data were normally distributed. For cases where normality of the data was not achieved through available transformation procedures, and to further inform the reliability of measurements, the concordance correlation coefficient (CCC) was calculated (Liao and Lewis, 2000, Lawrence and Lin, 1992). Similar to ICC interpretations, the strength of CCCs was considered little or zero, fair, moderate to good and good to excellent when CCC values were between 0.00 to 0.25, 0.25 to 0.50, 0.50 to 0.75 and >0.75 respectively (Portney and Watkins, 2009). ICCs were also calculated to express the degrees of correlation between measurements of operationally similar modalities (e.g. comparison of overall PPT value with unconditioned CPM stimulus).

Bland-Altman analysis was conducted to establish the 95% limits of agreement (LoA). Bland-Altman plots were produced to identify potential outliers and to visually evaluate the agreement between measurements of a single rater for each modality across populations as well as between both raters in the healthy population. The Bland-Altman analysis is based on the analysis of the average versus the difference of the means between two separate measurements or different time-points. The LoA represent the average difference ± 1.96 times the standard deviation of the differences and demonstrate the range within which 95% of the differences between the means in two single time points is expected to lie (Manresa et al., 2011). When comparing methods or assessing repeatability of small sample sizes, it is important

to calculate confidence intervals for 95% LoA to increase the accuracy of the estimate (Bland and Altman, 1999).

2.10.2.2. Validity testing

2.10.2.2.1. Internal validity

As the theoretical concept of all three protocol modalities (PPT, TS and CPM) is the quantification of pain sensitivity, correlation-coefficient data (Pearson's r) were presented to report the internal (association between QST modalities, association between measurements of operationally similar stimuli conducted for different modalities) and external (association between QST modalities and self-reported pain outcomes such as NRS and painDETECT) validity of the QST protocol. Similar to ICC and CCC interpretations, the strength of correlation-coefficients was considered little or zero, fair, moderate to good and good to excellent when r values were between 0.00 to 0.25, 0.25 to 0.50, 0.50 to 0.75 and >0.75 respectively (Portney and Watkins, 2009).

2.10.2.2.2. External validity

To further assess the external validity of the QST protocol, the association of each modality with the participants' pain distribution expressed in the total number of painful sites on a body manikin was explored. The Manikin was coded topographically as per the criteria proposed by different study groups (Wolfe et al., 2019, Wylde et al., 2011a, Croft et al., 1996).

2.10.3. Observational study

2.10.3.1. Pain sensitisation classifications

From a total population suffering from CLBP, approximately 25% is anticipated to demonstrate central sensitisation and be classified as such (Smart et al., 2012b), which is consistent across chronic musculoskeletal pain states (Lluch et al., 2014a, Mlekusch et al., 2013, Sterling et al., 2003a). Decreased pain thresholds at sites anatomically unrelated to the primary area of pain is considered a feature of central sensitisation in many musculoskeletal disorders (Smart et al., 2012b, Hidalgo-Lozano et al., 2010, Schliessbach et al., 2010, Fernández-Carnero et al., 2009, Freeman et

al., 2009). Since the application site for all three QST modalities was a site distant (forearm) to the primary areal of pain (lower back), a cut-off of 25% was used (Neziri et al., 2011) and therefore, the 25% of the population with the lowest values (below 1st quartile) for PPT-CPM and the 25% with the highest values (above 3rd quartile) for TS was considered as displaying features of central sensitisation. Given the chronicity of back pain as well as the potentially existing co-morbidities in the study sample that could be contributing to the overall pain experience, z-transformation of QST data based on age and sex-matched controls (Rolke et al., 2006a) could bias the results into classifying the entire patient population as demonstrating centrally-driven pain hypersensitivity. Classifying individuals based on quartiles/percentiles seems more suitable and of greater clinical value than z-transformation for this particular study as it allows flexibility in choosing the percentage that appears to be more clinically meaningful and adheres to consistent evidence across the musculoskeletal spectrum (Neziri et al., 2011). Nevertheless, this approach needs to be further validated by subsequent research.

For non-QST classification purposes, participants were operationally defined as 'sensitised' in a binary manner according to their pain distribution expressed as the total number of painful sites on a body manikin. Classifications were also undertaken according to different criteria, proposed by different study groups (Wolfe et al., 2019, Wylde et al., 2011a, Croft et al., 1996) (**Figure 9**). Receiver operating characteristics (ROC) analysis was conducted to establish the cut-off point for the optimal number of painful sites needed to classify low PPT (1st quartile), high TS (4th quartile) and inefficient CPM (1st quartile). To visually assess the performance of each a-priori classification, AUC-ROC curves were developed (Akin-Akinyosoye et al., 2018). Classification models that maximized sensitivity while maintaining a minimum specificity of 75% were considered acceptable for reliably predicting QST gain-of-function (Neblett et al., 2013). For interpretation purposes, an AUC value of ≥0.90 was indicative of high accuracy, 0.70-0.90 of moderate accuracy, 0.50-0.70 of low accuracy, and 0.50 was considered a 'toss-up' (chance) (Swets, 1988).

2.10.3.2. Central Mechanisms trait

Confirmatory factor analysis (CFA) was used to assess the fit of a single Central Mechanisms model composed by 8 distinct self-reported items shown to be associated with central sensitisation (Akin-Akinyosoye et al., 2018). CFA also allowed for the derivation of a single Central Mechanisms Trait score that was used for

analyses alongside other CS indices. An item was considered to display satisfactory loading only if it demonstrated an estimate value of ≥0.40 (Elsworth et al., 2015). Due to the relatively small sample size for reliable CFA inference, assessment of model fit was also based on Root Mean Square Error of Approximation (RMSEA) fit index (Fabrigar et al., 1999), which is the discrepancy between the model and the data per degree of freedom for the model. RMSEA values of <0.05 constitute good fit, 0.05 to 0.08 acceptable fit, 0.08 to 0.10 marginal fit and >0.10 poor fit (Browne and Cudeck, 1992). Additional values indicative of model fit are Comparative Fit Index (CFI) ≥0.95, Tucker-Lewis Index (TLI) ≥0.95, Chi^2 p-value ≤0.05, Standardised Root Mean Square Residual (SRMR) ≥0.07 (Yu, 2002).

2.10.3.3. Unadjusted correlations

2.10.3.3.1. Cross-sectional and prospective unadjusted associations between baseline demographic variables, CS indices and SM/SC outcomes at baseline and at 3-months follow-up

Baseline unadjusted associations between central sensitisation indices (QST, pain distribution, central mechanisms trait) and self-management/self-care outcomes were established by the Pearson's product-moment correlation test or Spearman's rank-order correlation test wherever data normality was not achieved through transformation and were expressed in correlation-coefficients (Pearson's r or Spearman's ρ).

As the theoretical concept of all central sensitisation indices is measurement and classification of pain sensitivity, correlation-coefficient data were presented to report the consistency between the indices (association between each of the CS indices). Bivariate analyses were also conducted between CS indices and pain, disability and neuropathic pain to further explore the external validity of the indices. The strength of association between CS indices, self-management outcomes and anthropometric and other variables (age, sex, BMI) was also explored through bivariate correlations.

In exploratory analyses, associations between secondary outcomes and self-management/self-care as well as CS indices and secondary outcomes were explored and expressed in the same way. Also, the bivariate associations between each different baseline variable (age, sex, pain, depression, fatigue, quality of life) aimed to be used as independent variables in multivariable regression models were explored.

Correlation analyses were undertaken to explore prospective associations between baseline demographic data and follow-up SM/SC outcomes as well as between CS indices at baseline and SM/SC outcomes at follow-up. Also, the bivariate associations between baseline and follow-up SM/SC constructs (HEIQ domains, PSEQ, HCUQ) as well as between each different baseline variable (age, sex, pain, depression, fatigue, quality of life) aimed to be used as independent variables in multivariable regression models were explored.

The strength of all correlation-coefficients (Pearson's and Spearman's), after adjusted for multiple comparisons (Benjamini and Hochberg corrected) (Jafari and Ansari-Pour, 2019), was considered little or zero, fair, moderate to good and good to excellent when r values were between 0.00 to 0.25, 0.25 to 0.50, 0.50 to 0.75 and >0.75 respectively (Portney and Watkins, 2009).

2.10.3.4. Adjusted correlation and regression modelling

2.10.3.4.1. Dependent variables

The dependent variables used in cross-sectional as well as prospective regression modelling were formed by the baseline and 3-months follow-up data respectively and are described in detail in **2.6.1**. They were comprised by the distinct self-management and self-care (SM/SC) constructs (HEIQ domains, PSEQ, HCUQ).

2.10.3.4.2. Independent variables

The independent variables used in cross-sectional and prospective regression modelling were formed by the baseline data collected for the outcome measures described in detail in **2.6.2** and **2.6.3**. They were comprised by the CS indices (QST modalities, pain distribution, and the Central Mechanisms trait), pain (NRS) or disability (RMDQ), depression (HADS), catastrophising (PCS), fatigue (FSS) and quality of life (EQ-5D-5L) measures.

2.10.3.4.3. Cross-sectional and prospective associations between baseline CS indices and SM/SC at 3-months follow-up adjusted for other factors

Variables found to demonstrate significant ($p<0.05$) or nearly significant ($p<0.1$) association in cross-sectional as well as prospective bivariate unadjusted correlations with either the dependent (SM/SC) or the independent variables (CS indices,

secondary outcomes) were considered eligible for inclusion into linear regression models. SM/SC data were entered into regression models as the dependent variable, with CS indices and secondary outcome scores serving as independent variables. The total number of covariates in the model could not exceed 7, due to limitations imposed by the baseline sample size (n=97).

Different models were explored for each of the CS indices and SM/SC constructs. Each model built to explore the association of CS indices and SM/SC constructs was adjusted for six variables; age, sex, pain, depression, catastrophising and fatigue that have been previously found to be predictors of SM/SC (Akin-Akinyosoye et al., 2019, Miles et al., 2011). Because constructs from 4 out of those 6 variables (painDETECT, depression, catastrophising, fatigue) were among those used for the derivation of the Central Mechanisms trait score, each model for the Central Mechanisms trait was adjusted only for pain, age, sex and quality of life.

As pain and disability have been found to be closely interlinked (Di Iorio et al., 2007) and disability appears not to be prevalent without the presence of pain (Lackner et al., 1996), inclusion of both constructs in a multivariable model could mediate their relationship with the dependent variable. Therefore, both in cross-sectional as well as prospective exploratory analyses, disability was used instead of pain in all multivariable analyses, comprising models featuring QST, age, sex, disability, depression, catastrophising and fatigue as independent variables.

Follow-up SM/SC outcomes were also entered into prospective regression models as the dependent variable, with baseline CS indices as well as age, sex, pain, depression, catastrophising and fatigue serving as independent variables. Exploratory analyses of multivariable models featuring the baseline score of each SM/SC outcome as an additional independent variable were reiterated to explore whether baseline CS indices are implicated in the development of poor SM/SC outcomes in the long-term. Multivariable models featuring disability instead of pain, as well as the baseline score of each SM/SC outcome as an additional independent variable, were reiterated to further explore whether the inclusion of disability acts as confounder in the relationship between baseline CS indices and long-term SM/SC outcomes or whether disability might be also implicated in the development of poor SM/SC outcomes.

Multivariable models featuring program participation as an independent variable were also explored, to examine whether associations between CS indices and SM/SC

outcomes may be generalizable between participants who participated in either the PT or MDT group intervention programme.

Goodness of model fit and the explanatory power of each regression model was established with the calculation of the coefficient of determination (adjusted R^2) statistic. The value of R^2 ranges from 0 (indicating that the independent variables do not explain the outcome) to 1 (the independent variables completely account for the outcome) and can be multiplied by 100 to express a percentage of the variance in the dependent variable explained by the independent variables (Katz, 2011).

All analyses were carried out with R Free Software (version 3.4.2) (R Core Team, 2017) and p-values of ≤0.05, after adjusted for multiple comparisons, were taken to indicate statistical significance. All significant correlations or associations were marked with bold font. In correlation or regression matrixes with multiple significant associations, statistically significant values were further highlighted in colour according to strength of correlation/association; $r/\beta<0.50$ in yellow whereas $r/\beta\geq0.50$ in green.

3. QUANTITATIVE SENSORY TESTING AND PREDICTING OUTCOMES FOR MUSCULOSKELETAL PAIN, DISABILITY, AND NEGATIVE AFFECT: A SYSTEMATIC REVIEW AND META-ANALYSIS

3.1. Introduction

Musculoskeletal disease is a worldwide phenomenon and one of the most frequent reasons for seeking healthcare assistance (May, 2010). Pain is paramount in a range of symptoms associated with musculoskeletal pathology which contribute to functional limitation (Picavet and Schouten, 2003, Urwin et al., 1998). The most prevalent musculoskeletal conditions that transition into chronicity include OA, LBP or neck pain, and rheumatoid arthritis (Bergman, 2007). Chronic pain may be initiated by musculoskeletal pathology, but is frequently also augmented by modulation of sensory inputs by the peripheral and central nervous system (Nijs et al., 2015, Daenen et al., 2013a, Staud et al., 2005, Banic et al., 2004, Price et al., 2002).

Central sensitisation refers to neurophysiological processes that can occur throughout the CNS leading to changes in the spinal cord as well as in supraspinal centres such as the brainstem, cerebral cortex, thalamus and the limbic system (Latremoliere and Woolf, 2009). Central sensitisation is implicated in pain chronification, manifested by pain hypersensitivity (augmentation) and spread to sites beyond those directly affected by musculoskeletal pathology (Woolf, 2011). Sustained activation of peripheral nociceptive pathways due to musculoskeletal pathology (e.g. trauma or inflammation) drives pain hypersensitivity (Fornasari, 2012), which may be maintained by neuroplastic changes in the CNS (Pelletier et al., 2015). Pain hypersensitivity is influenced by physical, genetic, psychological and environmental factors (Phillips and Clauw, 2011). Researchers have suggested that cognitive factors such as maladaptive beliefs (catastrophising, fear of movement, expectations of treatment outcomes) might contribute to pain hypersensitivity (Zusman, 2002, Ursin and Eriksen, 2001). Pain-specific cognitions such as catastrophisation influence endogenous pain modulation in healthy participants (Traxler et al., 2018). The presence of pain hypersensitivity complicates the clinical picture of chronic musculoskeletal conditions (Nijs et al., 2009), may cause or contribute to the transition from acute to chronic pain, and may be a barrier to achieving optimal treatment outcomes (Edwards et al., 2016b, Goldsmith et al., 2012, Jull et al., 2007).

Clinically important pain hypersensitivity is not present in all individuals with chronic pain (Smart et al., 2011, Woolf, 2011), contributing to heterogeneity in prognosis and treatment outcomes. It has been suggested that people with centrally driven pain hypersensitivity might better respond to education, exercise and cognitive behavioural therapy (CBT) than to treatments focusing on reducing nociceptive triggers alone (NICE, 2016, NICE, 2014, Lee et al., 2013).

Detection and measurement of hypersensitivity is challenging in human research and clinical practice, and there is not yet consensus on the most appropriate tools for use in chronic musculoskeletal pain (Girbés et al., 2013, Murphy et al., 2011). Identifying optimal indices of hypersensitivity is required to develop targeted treatment strategies that can improve patient outcomes. Self-report questionnaires may identify risk factors and symptoms that are commonly associated with central sensitisation (Akin-Akinyosoye et al., 2018). Qualities of pain in people with central sensitisation however overlap substantially with pain qualities in people with predominant nociceptive pain mechanisms. QST is a psychophysical approach through which stimuli are applied under standardised testing protocols, and the participants' self-reported sensory experience is quantified (Hall et al., 2015). QST can explore mechanisms responsible for the development or maintenance of local and widespread pain in musculoskeletal disorders (Courtney et al., 2010, Pavlaković and Petzke, 2010). QST utilises simple tools for the assessment of the perception of touch, vibration, proprioception, pinprick/blunt pressure sensitivity or sensitivity to cold or heat stimuli (Cruz-Almeida and Fillingim, 2014). The various QST modalities can provide important information about pain mechanisms (Courtney et al., 2010, Arendt-Nielsen and Yarnitsky, 2009), and can be used to quantify sensory alterations to healthy individuals and patients alike (Rolke et al., 2006a). However, the exact neurophysiological mechanisms that underline QST responses are not yet fully established.

The predictive capacity of QST has been previously explored in non-musculoskeletal pain states. Baseline sensory measurements have been associated with analgesic consumption in patients with chronic pancreatitis (Olesen et al., 2013) and in healthy individuals with experimental pain (Eisenberg et al., 2010), with the clinical course of painful temporomandibular disorder (Slade et al., 2014), and with tension-type headaches (Buchgreitz et al., 2008). In musculoskeletal conditions QST measures before surgery have been associated with acute post-operative outcomes (Sangesland et al., 2017, Abrishami et al., 2011, Edwards et al., 2005, Werner et al., 2010), however the capacity of QST to predict long-term post-operative outcomes and outcomes in non-surgical contexts has not been fully investigated. Potential

influences from QST modality and musculoskeletal diagnosis, and prediction of different pain-related experiences, such as pain severity, reduced functional capacity (disability), anxiety and depression (negative affect) (Smart et al., 2012a) remain uncertain. A greater understanding of the role of pain hypersensitivity in prognosis and how QST might predict musculoskeletal outcomes should help better predict those who are most likely to gain benefit from treatments aiming to reduce ongoing pain, distress or disability.

3.2. Aims and Objectives

3.2.1. Aims

The aim of this systematic review was to determine the ability of QST to predict musculoskeletal outcomes

3.2.2. Objectives

- To systematically search the literature to identify titles that have reported prospective associations between QST and pain, disability and negative affect or between QST and SM/SC outcomes.
- To extract relevant data and conduct meta-analysis to explore whether QST collectively is correlated with distinct musculoskeletal outcomes and establish the strength of such correlations.
- Perform subgroup analyses to explore the capacity of distinct QST modalities or groups of operationally similar modalities to predict musculoskeletal outcomes.

3.3. Methods

The full methodological details of this study are given in the Methods chapter (**METHODS**) and only a brief outline is presented in this chapter whenever necessary.

3.3.1. Systematic literature search

A systematic online search was conducted in 6 databases (CENTRAL, MEDLINE, EMBASE, AMED, CINAHL and PubMed) from 1948 until April 2018. Information

regarding the composition of the search strategy and methodological steps of literature search are given in **2.8.1.1**. A list of the search terms and their combinations is demonstrated in **Appendix 1**.

3.3.2. Inclusion/Exclusion criteria

Information about the inclusion/exclusion process as well as the eligibility criteria studies had to satisfy to be included in the systematic review are offered in **2.8.1.2** and in **Table 5** respectively.

3.3.3. Data extraction

Two independent reviewers undertook the data extraction procedure. Any disagreement on the extracted data or their interpretation achieved via consensus after discussion. Information about the type of data extracted is given in **2.8.1.3**.

3.3.4. Quality and content assessment

The quality of included studies was appraised by the Quality In Prognosis Studies (QUIPS) Tool (Hayden et al., 2006) for observational cohort studies as well as RCTs.

3.3.5. Data synthesis and analysis

Data coding was conducted by one reviewer and were validated independently by a second reviewer. Extracted data were grouped according to type and were entered into meta-analysis models. The full methodological details regarding data coding and meta-analysis are given in **2.9.1.1** and **2.10.1.1** respectively.

3.4. Results

3.4.1. Characteristics of included studies

The study selection process is shown in **Figure 10** (37 studies) and a summary of the study characteristics in **Table 7**. An overview of study data is given in **Table 8** and full study details in **Appendix 2**, **Appendix 3** and **Appendix 4**.

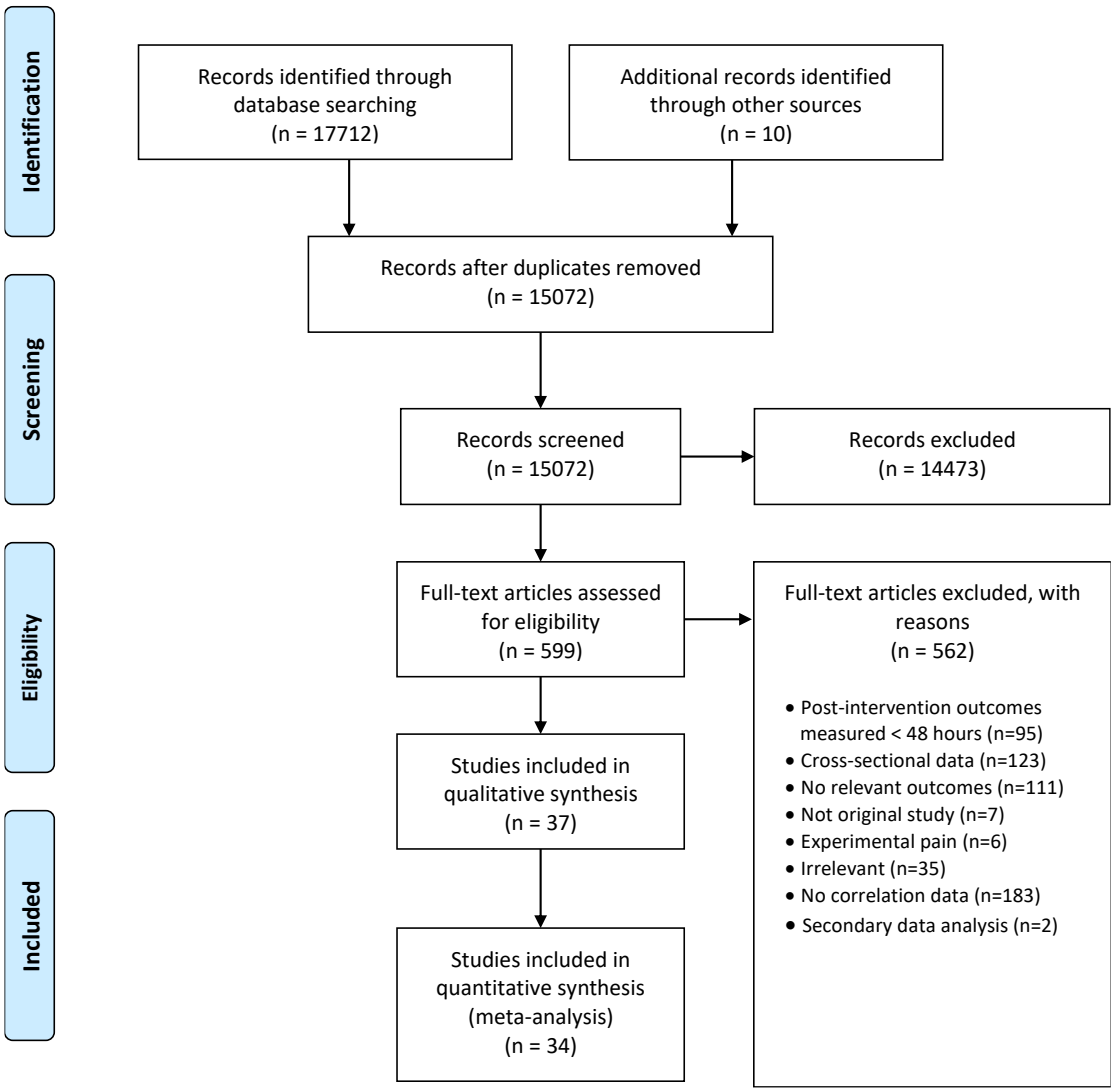

Figure 10. PRISMA flow chart of the study selection process

Of the 37 studies that met the inclusion criteria, 32 (Izumi et al., 2017, Luna et al., 2017, Bar Ziv et al., 2016, Dubois et al., 2016, Edwards et al., 2016a, Pedler et al., 2016, Petersen et al., 2016, Thomazeau et al., 2016, Coombes et al., 2015, Coronado et al., 2015b, Petersen et al., 2015, Wylde et al., 2015, Goodin et al., 2014, Noiseux et al., 2014, Valencia et al., 2014, Davis et al., 2013, LeResche et al., 2013, Mlekusch et al., 2013, Wylde et al., 2013, Rakel et al., 2012, Sterling et al., 2012, Gwilym et al., 2011, Sterling et al., 2011, Walton et al., 2011, Aasvang et al., 2010, Wilder-Smith et al., 2010, Lundblad et al., 2008, Yarnitsky et al., 2008, Martinez et al., 2007, Sterling et al., 2005, Werner et al., 2004) were prospective cohort studies and 5 were RCTs (Arendt-Nielsen et al., 2016, Mendonca et al., 2016, Coronado et al., 2015a, Henriksen et al., 2014, Jull et al., 2013).

The total number of participants was 3810, of whom 277 were in RCTs. Women comprised 58% of all participants and the average age of participants in each study ranged from 36 to 72 years (**Table 7**). Interventions offered in the RCTs were NSAIDs (Arendt-Nielsen et al., 2016), cervical and shoulder manipulation (Coronado et al., 2015a), exercise therapy (Henriksen et al., 2014), Transcranial Magnetic Stimulation (TMS) (Mendonca et al., 2016) and a combination of analgesic medication, physiotherapy, education and psychological support (Jull et al., 2013).

Knee osteoarthritis (OA) was the most commonly studied condition (13/37 studies), and post-operative pain after thoracotomy, subacromial decompression, total hip (THR) or knee replacement (TKR) and abdominal surgery comprised the second commonest condition (8/37). The remaining studies focused on Whiplash Associated Disorders (WAD) (6/37), low back pain (LBP) (4/37), shoulder pain (3/37), epicondylitis (1/37), fibromyalgia (FM) (1/37) and anterior cruciate ligament tear (1/37).

Table 7. Summary of study characteristics

	Cohort	RCT	All Studies
No. studies	32	5	37
No. subjects	3583	277	3860
Mean age (years)	55	54	55
Female%	56	72	58
Setting			
Hospital	6	2	8
Community	6	2	8
Unclear	20	1	21
Diagnosis			
OA	15	2	17
MSK Injury	1	0	1
Whiplash	5	1	6
Neck Pain	1	0	1
Low Back Pain	4	0	4
Fibromyalgia	0	1	1
Shoulder pain	3	1	4
Postoperative pain	3	0	3
Affected Site			
Knee	15	2	17
Hip	1	0	1
Neck	6	1	7
Low Back	4	0	4
Shoulder	3	1	4
Thorax	1	0	1
Abdominal area	2	0	2
Widespread body pain	0	1	1
MSK Outcome Measure			
Pain	26	4	30
Disability	10	1	11
Depression	2	0	2
Anxiety	1	0	1
QST Stimulus Modality†			
Mechanical			
Pressure	21	5	26
Punctate	8	0	8
Movement	1	0	1
Electrical	5	1	6
Thermal			
Heat	13	1	14
Cold	14	1	15
QST Outcome Measure†			
Pain Detection Threshold	28	4	32
Pain Tolerance Threshold	6	0	6
Sensation Detection Threshold	5	1	6
Pain Intensity	8	0	8
Conditioned pain modulation	11	1	12
Temporal Summation	5	3	8
Spatial Summation	1	0	1
QST Test Sites†			
Affected joint	20	4	24
Distal to affected joint	10	3	13
Remote	20	3	23

MSK: Musculoskeletal, **OA:** Osteoarthritis, **QST:** Quantitative Sensory Testing, **RCT:** Randomised Controlled Trail

†One study may involve more than one QST modality, outcome measure and test site

Table 8. Overview of study characteristics and association data

Author	Study Design	Sample size	Country	Mean Age	Female %	Setting	Was QST vs Outcome the primary research question?	Diagnosis	Site of Pathology	Baseline QST Predictor	QST Testing site	Outcome	Outcome Measures	Statistical Analysis Method
Aasvang et al., 2010	Prospective Cohort	442	Denmark, Germany	55.2	0%	Hospital	Yes	Postoperative pain	Groin	WDT, HPT, THPR	**WDT, HPT, THPR:** 2 cm lateral to the pubic bone, 2 cm above the inguinal ligament and anterior on the ipsilateral forearm approximately 10 cm distal to the cubital crease	Disability	AAS	Multivariate Logistic Regression Analysis
Arendt-Nielsen et al., 2016	RCT	37	Denmark	63.3	61%	Hospital	No	OA	Knee	PPT, SS, CPM, TS	**CPM:** Tourniquet cuff at contralateral arm to most sensitive knee and then PPT to Knee, Arm, Leg; **PPT:** Patella of most painful knee (superior edge, medial edge and lateral midpoint); **SS:** PPT at leg (tibialis anterior) and arm (radialis longus); **TS:** Patella of most painful knee (superior edge, medial edge and lateral midpoint) and Tibialis anterior	Pain	BPI (0-11), WOMAC	Correlation analysis
Bar Ziv et al., 2016	Prospective Cohort	48	Israel	72	71%	Unclear	Yes	OA	Knee	PPI	**PPI:** Tourniquet cuff distal to the elbow crease on the proximal forearm	Pain	KSS	Correlation analysis
Coombes et al., 2015	Prospective Cohort (Data analysis)	41	Australia	49.9	42%	Unclear	Yes	Lateral Epicondylalgia (Tennis Elbow)	Elbow	PPT, CPT	**PPT, CPT:** Perpendicular to the tissues at the tendinous origin of the wrist extensor muscles	Pain, Disability	PRTEE	Multivariate Linear Regression Analysis

Coronado et al., 2015a	RCT	63	USA	57.9	59%	Hospital	No	General Shoulder Pain	Shoulder	PPT, HPT, TS	**PPT:** Dominant side acromion; bilateral tibialis anterior (TA) muscle belly; **HPT, TS:** Anterior forearm	Pain	BPI (0-11)	Correlation analysis
Coronado et al., 2015b	Prospective Cohort	68	USA	39	46%	Unclear	Yes	LBP	Low Back	PPT	**PPT:** Each side of the L1 spinous process	Pain, Disability	BPI, ODI	Multivariate Logistic Regression analysis
Davis et al., 2013	Prospective Cohort	31	United Kingdom	51.4	52%	Unclear	Yes	Postoperative pain	Shoulder	EPT	**EPT:** Grip the Pain Matcher device with hand contra-lateral to that side undergoing surgery	Pain	VAS	Univariate and Multivariate Linear Regression Analysis
Dubois et al., 2016	Prospective Cohort	77	Canada	37	49%	Community	Yes	LBP	Low Back	HPT (Thr.), HPT (Tol.), HHNCS	**HPT** (Thr.)**, HPT** (Tol.)**:** Midline between L4-L5 spinous processes and middle point between medial condyle of humerus and styloid process if ulna; **HNCS:** Midline between L4-L5 spinous processes and left hand	Disability	RMDQ	Correlation analysis and Multivariate Linear Regression analysis
Edwards et al., 2016	Prospective Cohort (Case-control)	35	USA	57.9	67%	Unclear	No	OA	Knee	PPT, TS, CPA, CPM (via PPT+CPT), CPT	**PPT:** Trapezius muscle, the patella, and the metacarpophalangeal joint of the thumb; **TS:** Dorsum of hand; **CPA:** Gastrocnemius muscle; **CPT:** Right hand	Pain	DPI, KOOS	Correlation and Stepwise Linear Regression analysis
Goodin et al., 2014	Prospective Cohort	225	USA	57.1	68%	Unclear	Yes	OA	Knee	TS (Mechanical), TS (Heat)	**Mech TS:** Patella of the index knee and the back of the ipsilateral hand; **Heat TS:** index knee and ipsilateral volar forearm	Pain, Depression	NRS, CES-D	Correlation analysis and Multivariate Linear Regression Analysis
Gwilym et al., 2011	Prospective Cohort	17	United Kingdom	55	59%	Unclear	Yes	Subacromial	Shoulder	MPT (Punctate	**MPT:** Deltoid insertion of deltoids bilaterally	Pain	OSS	Correlation analysis

	(Case-control)							Impingement		Sharpness Threshold)				
Henriksen et al., 2014	RCT	48	Denmark	63.7	81%	Community	Yes	OA	Knee	CPA via PPT, TS	**PPT, TS:** Calf (gastrocnemius muscle of affected leg)	Pain	KOOS	Correlation analysis
Izumi et al., 2017	Prospective Cohort (Case-control)	40	Denmark	68	50%	Unclear	Yes	OA	Hip	PPT, cPPT, cPTT, TS, SS, PinPS, CDT, CPT, WDT, CPM	**PPT:** 6 Hip locations: m. gluteus medius; 3 cm proximal to the tip of the greater trochanter (Hip-1). M. gluteus maximus; 3 cm posterior to the posterior edge of the greater trochanter (Hip-2). M. vastus lateralis; 3 cm distal to the distal edge of the greater trochanter (Hip-3). M. tensor fascia latae; 3 cm anterior to the anterior edge of the greater trochanter (Hip-4). M. tibialis anterior; 5 cm distal to the tibial tuberosity (TA). M. extensor carpi radialis longus; 5 cm distal to the lateral epicondyle of the humerus (arm). 11 body areas bilaterally; (1) groin, (2) greater trochanter, (3) buttocks, (4) anterior thigh, (5) posterior thigh, (6) lateral thigh, (7) medial thigh, (8) knee, (9) lower leg, (10) foot, and (11) lumbar; **cPPT, cPTT:** Proximal thigh (below inguinal crease); **TS, SS:** Proximal thigh; **PinPS:** Hip-2, Hip-4; **CDT, CPT, WDT:** Hip-4; **CPM:** Contralateral biceps brachii muscle, Hip-2, Hip-4	Pain	VAS	Correlation analysis

Jull et al., 2013	RCT	97	Australia	35.9	61%	Community	No	WAD	Neck	CPT, PPT	**PPT:** Over the cervical region and over a remote site (tibialis anterior); **CPT:** Over the cervical region using a thermode	Disability	NDI	Multivariate Linear Regression Analysis
LeResche et al., 2013	Prospective Cohort	147	USA	47.4	62%	Community	Yes	LBP	Low Back	PPT, TS (VFPI), CPM, CPP	**PPT:** Right and left side of low back, middle of right thenar emminence; **TS:** Volar aspect of non-dominant forearm; **CPM:** Volar aspect of dominant forearm and immersion of non-dominant hand in cold water; **CPP:** Immersion of non-dominant hand in cold water	Pain	NRS	Multivariate Linear Regression Analysis
Luna et al., 2017	Prospective Cohort	60	Denmark	67	62%	Hospital	Yes	OA	Knee	CPTol, PPT, EPThr, EPTol	**CPTol:** Immersion of non-dominant hand and wrist in cold water; **PPT:** Volar surface of the dominant forearm and on medial site of both knees; **EPThr, EPTol:** Fingertips of non-dominant hand	Pain	NRS	Multivariate Logistic Regression Analysis
Lundblad et al., 2008	Prospective Cohort	69	Sweden	68	51%	Hospital	Yes	Postoperative pain	Knee	EPT, EST	**EPT, EST:** Hold electrodes between the thumb and index finger of their hand	Pain	VAS	Correlation analysis and logistic regression analysis

Martinez et al., 2007	Prospective Cohort	20	Canada	69	95%	Hospital	No	OA	Knee	PPT, Tac.Alo, HPT, CPT	**PPT, TA:** Operated patella & adjacent non-inflamed area in the proximal direction; **HPT, CPT:** Operated patella; Control measurements for mechanical and thermal pain threshold were performed in two remote sites: the contralateral knee (stimulation on the patella; 1 cm lateral to the midline) and the palmar aspect of the right hand	Pain	VAS	Correlation analysis
Mendonca et al., 2016	RCT	32	Brazil	47.5	97%	Unclear	No	Fibromyalgia	Widespread Pain	PPT	**PPT:** Thenar region of the hand and Tibialis Anterior	Pain	VNS	Correlation analysis and Univariate Linear Regression Analysis
Mlekusch et al., 2013	Prospective Cohort	169	Switzerland	49.5	42%	Community	Yes	LBP, Neck Pain	Low Back, Neck	PTT, CPM, CTT	**PTT:** Center of the pulp of the second toe; **CPM, CTT:** Hand in cold water	Pain	BPI	Multivariate Linear Regression analysis
Noiseux et al., 2014	Prospective Cohort (Data analysis)	215	USA	61.7	58%	Unclear	Yes	OA	Knee	MPS (VFPI), HPT, PPT	**MPS, HPT, PPT:** All measurements were performed on three sites medial to the center of the patella	Pain	0-20 Scale: Mild, Moderate, Severe Resting & Movement Pain	Multivariate Logistic Regression Analysis
Pedler et al., 2016	Prospective Cohort	91	Australia	39.7	72%	Unclear	No	WAD	Neck	CPT, PPT	**CPT:** Mid & Lower Cervical spine, **PPT:** Spinous process of the 2nd Cervical verterbra	Pain, Disability	VAS, NDI	Correlation analysis and Multivariate Linear Regression Analysis

Petersen et al., 2015	Prospective Cohort	78	Denmark	70	59%	Unclear	Yes	OA	Knee	PPT, TS, CPM (via CPT)	**PPT, TS:** Most affected knee joint (peripatellar region, tibialis anterior (TA), extensor carpi radialis longus (arm); **CPM:** hand contralateral to the most affected knee	Pain	VAS	Correlation analysis and Multivariate Linear Regression Analysis
Petersen et al., 2016	Prospective Cohort	103	Denmark	69.2	73%	Unclear	Yes	OA	Knee	PDT, PTT, TS, CPM, PPT	**PDT, PTT, TS, CPM:** Head of gastrocnemius muscle (most affected); **PPT:** Most affected knee, TA, arm	Pain	VAS	Correlation analysis and Multivariate Linear Regression Analysis
Rakel et al., 2012	Prospective Cohort	215	USA	61.7	58%	Hospital	No	Postoperative pain	Knee	MPS (VFPI), HPT, PPT	**MPS, HPT, PPT:** All measurements were performed on three sites 4cm apart and 4cm medial to the center of the patella bilaterally	Pain	0-20 Scale: Mild, Moderate, Severe Resting & Movement Pain	Correlation analysis and Multivariate Linear Regression analysis and Logistic Regression Analysis
Sterling et al., 2005	Prospective Cohort	76	Australia	36.3	74%	Community	No	WAD	Neck	PPT, HPT, CPT, BPPT	**PPT:** bilaterally over the articular pillars of C2/3 and C5/6, over the three main peripheral upper limb nerve trunks and at TA; **HPT, CPT:** over the cervical spine; **BPPT:** Arm	Disability	NDI (6 months)	Stepwise Multivariate Linear Regression analysis and Logistic Regression Analysis
Sterling et al., 2011	Prospective Cohort	155	Australia	36.9	63%	Community	No	WAD	Neck	PPT, CPT	**PPT:** Neck: C5 spinous process, Arm: Median nerve at elbow (bilaterally); **CPT:** Over the mid	Disability, PTSD	NDI, PDS	Multivariate Linear Regression analysis

	(Data analysis)										to lower regions of the cervical spine			
Sterling et al., 2012	Prospective Cohort	225	Australia, Iceland, Canada	33.6	79%	Unclear	No	WAD	Neck	CPT	**CPT:** Mid cervical spine	Disability	NDI	Multivariate Linear Regression analysis
Thomazeau et al., 2016	Prospective Cohort	103	France	69.2	70%	Hospital	Yes	Postoperative pain	Knee	EST, EPT	**EST, EPT:** Median nerve territory	Pain	BPI	Correlation analysis and Multivariate linear regression analysis
Vaegter et al., 2016	Prospective Cohort	14	Denmark	66.3	50%	Unclear	Yes	OA	Knee	PPT, PTT, CPT, CPM (via CPT and PPT)	**PPT:** Both legs, the arm, the shoulder and lower leg of the affected side. **CPM:** Foot of the nonaffected leg	Pain	NRS	Correlation analysis
Valencia et al., 2014	Prospective Cohort	78	USA	47.3	28%	Unclear	Yes	Shoulder Pain due to multiple conditions	Shoulder	SHPR, HPT, CPM	**SHPR:** Thenar eminence of surgical and non-surgical side; **HPT:** Forearm of surgical and non-surgical side; **CPM:** Surgical side hand in cold water	Pain, Disability, Depression	BPI, DASH, PHQ-9	Correlation analysis and Multivariate linear regression analysis
Walton et al., 2011	Prospective Cohort	45	Canada	38	73%	Community	Yes	WAD	Neck	PPT	**PPT:** Angle of upper fibers of the trapezius muscle and from the belly of tibialis anterior	Disability	NDI	Correlation analysis and Multivariate linear regression analysis
Werner et al., 2004	Prospective Cohort	20	Denmark	28	30%	Unclear	Yes	ACL Tear	Knee	MPT, MPP, HPT, HPP	**MPT, MPP, HPT, HPP:** Medial aspect of the calf contralateral to the surgical side	Pain	VAS	Correlation analysis

Wilder-Smith et al., 2010	Prospective Cohort	20	Netherlands	53	54%	Unclear	Yes	Postoperative pain	Abdomen	ePTT, pPTT, DNIC via ePTT or pPTT and immersion to cold water	**ePTT, pPTT:** Lateral to the point of the (planned) surgical incision on the abdomen and in the L2 dermatome of the leg opposite to operated side; **DNIC:** ePTT or pPTT and immersion of the hand on the side of operation in cold water	Pain	VAS	Correlation analysis
Wylde et al., 2013	Prospective Cohort	51	United Kingdom	68	57%	Unclear	Yes	OA	Knee	PPT, HPT	**PPT, HPT:** Volar surface of right forearm and medial side of index knee	Pain	WOMAC	Correlation analysis
Wylde et al., 2015	Prospective Cohort (Data analysis)	493	United Kingdom	67.8	27%	Unclear	Yes	Postoperative pain	Knee/Hip	PPT	**PPT:** Pain-free site distant to the painful joint	Pain	WOMAC	Multivariate Linear Regression Analysis
Yarnitsky et al., 2008	Prospective Cohort	62	Israel	61.8	39%	Unclear	Yes	Postoperative pain	Thorax	DNIC via heat noxious stimuli and immersion to hot water	**HPT:** Thenar part of right forearm; **DNIC:** Stimulus to volar aspect of dominant forearm and immersion of non-dominant hand in hot water	Pain	NRS	Correlation analysis and Logistic Regression Analysis

Note: A negative correlation or β-coefficient value indicates that a low QST value is associated with a higher level of pain, disability or depression.

AAS: Activity Assessment Scale, **AUC:** Area Under Curve, **β:** Beta Coefficient of Regression, **BPI:** Brief Pain Inventory, **BPPT:** Brachial Plexus Provocation Tess, **CDT:** Cold Detection Threshold, **CES-D:** Center for Epidemiological Studies-Depression Scale, **CPA:** Cuff Pressure Algometry, **CPM:** Conditioned Pain Modulation, **CPP:** Cold Pressor Pain, **cPPT:** Cuff Pain Pressure Threshold, **CPT:** Cold Pain Threshold, **CPTol:** Cold Pain Tolerance Threshold, **cPTT:** Cuff Pain Tolerance Threshold, **CTT:** Cold Tolerance Time, **DNIC:** Diffuse Noxious Inhibitory Control, **DPI:** Daily Pain Intensity, **EPT:** Electrical Pain Threshold, **EPThr:** Electrical Pain Detection Threshold, **EPTol:** Electrical Pain Tolerance Threshold, **ePTT:** Electrical Pain Tolerance Threshold, **EST:** Electrical Sensation Threshold, **HNCS:** Heterotopic Noxious Counter-Stimulation, **HPP:** Heat Pain Perception, **HPT:** Heat Pain Threshold, **HPT (Thr.):** Heat Pain Detection Threshold, **HPT (Tol.):** Heat Pain Tolerance Threshold, **KOOS:** Knee Injury and Osteoarthritis Outcomes Score, **kPa:** Kilopascal, **KSS:** Knee Society Score, **LBP:** Low Back Pain, **MEP:** Motor-Evoked Potentials, **MPS:** Mechanical Pain Sensitivity, **MPP:** Mechanical Pain Perception, **MPT:** Mechanical Pain Threshold, **NDI:** Neck Disability Index, **NRS:** Numerical Rating Scale, **OA:** Osteoarthritis, **ODI:** Oswestry Disability Index, **OR:** Odds Ratio, **OSS:** Oxford Shoulder Score, **PDS:** Post-traumatic Stress Diagnostic Scale, **PinPS:** Pinprick Pain Sensitivity, **PPI:** Pain Pressure Intensity, **PPT:** Pressure Pain Detection Threshold, **PTT:** Pain Tolerance Threshold, **pPTT:** Pressure Pain Tolerance Threshold, **PRTEE:** Patient Rated Tennis Elbow Evaluation, **PTSD:** Post-traumatic Stress Disorder, **r:** Pearson's r Correlation-Coefficient, **ρ:** Spearman's ρ Rank-Order Correlation, **RCT:** Randomised Controlled Trial, **RMDQ:** Roland-Morris Disability Questionnaire, **SHPR:** Suprathreshold Heat Pain Response, **SS:** Spreading Sensitisation, **TA:** Tibialis Anterior, **Tac.Alo:** Tactile Allodynia, **THPR:** Tonic Heat Pain Response, **TS:** Temporal Summation, **VAS:** Visual Analogue Scale, **VFPI:** Von Frey Pain Intensities, **VNS:** Visual Numeric Scale, **WAD:** Whiplash Associated Disorders, **WDT:** Warmth Detection Threshold, **WOMAC:** Western Ontario and McMaster Universities Osteoarthritis Index, **χ^2:** Determinant of significant difference

Thirty-four studies reported data that could be included in meta-analysis; 22 studies reported correlation coefficients and 24 studies reported β-coefficients or Odds Ratios (OR). Two studies each reported on either 2 separate conditions (LBP and neck pain) or 2 interventions (THR and TKR) and each provided data for their different populations that allowed fitting within a single meta-analysis model. All 37 identified studies demonstrated good methodological quality with most of them (25/37) displaying low risk of bias (**Table 9**). Out of the 34 studies included in meta-analysis models 24 were considered of low risk of bias and 10 of moderate risk (**Table 9**).

Table 9. Assessment of risk of bias

		Quality in Prognosis Studies (QUIPS)						
	Studies & Risk of Bias Tool Criteria	*Study Participation*	*Study Attrition*	*Prognostic Factor Measurement*	*Outcome Measurement*	*Study Confounding*	*Statistical Analysis and Reporting*	**Overall Risk of Bias**
RCTs	Arendt-Nielsen et al., 2016	Low	Low	Low	Low	Low	Low	**Low**
	Coronado et al., 2015a	Moderate	Low	Low	Low	Low	Low	**Low**
	Henriksen et al., 2014	Low	Low	Low	Low	Low	Low	**Low**
	Jull et al., 2013	Low	Low	Low	Low	Low	Low	**Low**
	Mendonca et al., 2016	Low	Low	Low	Low	Moderate	Low	**Low**
Prospective Observational	Aasvang et al., 2010	Low	Low	Low	Low	Low	Low	**Low**
	Bar Ziv et al., 2016	Low	Low	Low	Low	Moderate	Low	**Low**
	Coombes et al., 2015	Low	Low	Low	Low	Moderate	Low	**Low**
	Coronado et al., 2015b	Low	Low	Low	Low	Low	Low	**Low**
	Davis et al., 2013	Low	Low	Low	Low	Moderate	Moderate	**Moderate**
	Dubois et al., 2016	Low	Low	Low	Low	Moderate	Low	**Low**
	Edwards et al., 2016	Low	Low	Low	Low	Moderate	Low	**Moderate**
	Goodin et al., 2014	Low	Low	Low	Low	Moderate	Low	**Low**
	Gwilym et al., 2011	Low	Low	Low	Low	Moderate	Low	**Moderate**
	Izumi et al., 2017	Low	Low	Low	Low	Low	Low	**Low**
	LeResche et al., 2013	Low	Low	Low	Low	Low	Low	**Low**
	Luna et al., 2017	Low	Low	Low	Low	Moderate	Low	**Moderate**
	Lundblad et al., 2008	Low	Low	Low	Low	Moderate	Low	**Moderate**
	Martinez et al., 2007	Low	Low	Low	Low	Moderate	Low	**Low**
	Mlekusch et al., 2013	Low	Low	Low	Low	Moderate	Low	**Low**

Noiseux et al., 2014	Low	Low	Low	Low	Moderate	Low	**Low**
Pedler et al., 2016	Low	Low	Low	Low	Low	Low	**Low**
Petersen et al., 2015	Low	Low	Low	Low	Low	Low	**Low**
Petersen et al., 2016	Low	Low	Low	Low	Low	Moderate	**Low**
Rakel et al., 2012	Low	Low	Low	Low	Moderate	Low	**Moderate**
Sterling et al., 2005	Low	Low	Low	Low	Moderate	Low	**Moderate**
Sterling et al., 2011	Low	Low	Low	Low	Low	Low	**Low**
Sterling et al., 2012	Low	Low	Low	Low	Low	Low	**Low**
Thomazeau et al., 2016	Low	Low	Low	Low	Moderate	Low	**Low**
Vaegter et al., 2016	High	Low	Low	Low	Low	Low	**Moderate**
Valencia et al., 2014	Low	Low	Low	Low	Low	Low	**Low**
Walton et al., 2011	Low	Low	Low	Low	High	Low	**Moderate**
Werner et al., 2004	Moderate	Low	Low	Low	High	Low	**Moderate**
Wilder-Smith et al., 2010	Low	Low	Low	Low	Moderate	Low	**Low**
Wylde et al., 2013	Low	Low	Low	Low	Moderate	Low	**Moderate**
Wylde et al., 2015	Low	Low	Low	Low	Low	Low	**Low**
Yarnitsky et al., 2008	Low	Low	Low	Low	Low	Low	**Low**

RCT: Randomised Controlled Trial, **QUIPS:** Quality in Prognosis Studies risk of bias tool

3.4.2. QST modalities, outcomes and test sites

The majority of the studies (30/37) reported on more than one QST modality. Mechanical pressure (assessed by various pressure algometers) was the most common stimulus (26/37) followed by cold (15/37), heat (14/37) and punctate (pinprick: 5/37, von-Frey monofilaments 3/37) stimuli. Nearly all studies (32/37) reported pain detection and tolerance thresholds as a QST outcome. Conditioned pain modulation (12/37), temporal summation (8/37), pain intensity (8/37), sensation detection threshold (6/37), and spatial summation (1/37) was each reported in a minority of studies. Most publications reported more than one anatomical site for QST assessment (**Table 8**). The commonest were the site of clinical pain (24/37) followed equally by sites that were distal to (13) or remote from (13) the site of reported pain.

3.4.3. Reliability of QST

Four studies (Izumi et al., 2017, Pedler et al., 2016, Wylde et al., 2013, Yarnitsky et al., 2008, Martinez et al., 2007) reported that QST applications were performed by the same individual but did not report test-retest reliability. One study (Coronado et al., 2015b) referenced a previous study to report QST reliability. Only 1 of the 37 studies (Rakel et al., 2012) reported intraclass correlation coefficient (ICC) for test-retest and inter-rater reliability with values 0.92-0.97 for mechanical pain sensitivity (MPS), 0.70-0.92 for heat pain threshold (HPT) and 0.87-0.97 for PPT. In 2 studies (Arendt-Nielsen et al., 2016, LeResche et al., 2013) assessments were conducted by multiple individuals but no ICC or inter-observer variability was reported. One of the studies (LeResche et al., 2013) reported this as a methodological limitation.

3.4.4. Ability of QST to predict outcomes in people with musculoskeletal conditions

Baseline QST demonstrated a statistically significant association with worse musculoskeletal-related outcomes (pain or disability) in 35 of the 37 studies (**Appendix 2**). Presentation of associations varied between studies as Pearson's r, β-coefficients, ORs, AUC and Chi^2. Several studies (Dubois et al., 2016, Edwards et al., 2016a, Mendonca et al., 2016, Pedler et al., 2016, Petersen et al., 2016, Petersen et al., 2015, Goodin et al., 2014, Valencia et al., 2014, Walton et al., 2011, Yarnitsky et al., 2008) reported both correlation-coefficient and regression coefficient values. Two studies (Noiseux et al., 2014, Martinez et al., 2007) narratively reported (without presenting data) no observed correlation between baseline QST, measured using mechanical and thermal modalities (sensitivity and pain threshold) and follow-up pain. Five other studies used mechanical stimuli (Izumi et al., 2017, Arendt-Nielsen et al., 2016, Petersen et al., 2016, Yarnitsky et al., 2008) or electrical stimuli (Thomazeau et al., 2016), but presented data only partially, favouring data that supported association.

Twenty-five studies reported regression analyses of which 22 reported baseline factors used for statistical adjustment. **Table 10** describes these 22 studies, their outcomes and the factors they adjusted for. Pain alone was the outcome in 13/22 and disability alone in 6/22. Both pain and disability were reported in 2/22, and disability and negative affect were reported together in 1/22. Fifteen studies reported baseline pain scores, of which 11 adjusted outcomes for baseline pain, and 4 for factors other than pain measured at baseline (pain catastrophising, depression, age, sex, ethnicity,

analgesia requirement, pain duration, genetic factors). Adjustment for baseline disability was reported in 3/9 of the studies reporting disability as an outcome. The single study that measured negative affect did not adjust for baseline negative affect.

Table 10. Adjustment of prognostic models for baseline variables

Study	QST Modality	Musculoskeletal Outcome	Adjusted for
Aasvang et al., 2010	THPR	Disability (AAS)	Baseline disability Detection threshold change Postoperative pain
Coronado et al., 2015b	PPT	Pain (BPI) Disability (ODI)	Baseline catastrophisation Baseline catastrophisation
Davis et al., 2013	EPT	Pain (VAS)	Extra analgesia requirement
Dubois et al., 2016	CPM	Disability (RMDQ)	Pain at 6 months
Edwards et al., 2016	CPM	Pain (KOOS)	Baseline pain
Goodin et al., 2014	TS	Pain (NRS)	Ethnicity
Jull et al., 2013	CPT	Disability (NDI)	Baseline pain
LeResche et al., 2013	PPT CPM TS CPP	Pain (NRS)	Baseline pain Age and Sex
Luna et al., 2017	PPT	Pain (NRS)	Pre-surgery catastrophisation
Lundblad et al., 2008	EPT	Pain (VAS)	Baseline pain
Mlekusch et al., 2013	PTT CPM	Pain (BPI)	Baseline pain Baseline catastrophisation Baseline depression Pain duration Age and Sex Opioid intake
Pedler et al., 2016	CPT PPT	Pain (VAS) Disability (NDI)	Baseline pain Baseline depression
Petersen et al., 2015	TS	Pain (VAS)	Baseline pain
Petersen et al., 2016	PDT MPS	Pain (VAS)	Baseline pain Baseline pain
Rakel et al., 2012	HPT PPT	Pain (NRS)	Baseline depression Baseline anxiety
Sterling et al., 2005	CPT	Disability (NDI)	Baseline disability
Sterling et al., 2011	CPT PPT	Disability (NDI) Post-traumatic Stress Disorder (PDS)	Baseline pain
Sterling et al., 2012	CPT	Disability (NDI)	Baseline disability
Thomazeau et al., 2016	EST EPT	Pain (BPI)	Baseline pain (rest) Baseline depression Genetic factors Follow-up opioid intake
Walton et al., 2011	PPT	Disability (NDI)	Baseline pain Gender
Wylde et al., 2015	PPT	Pain (WOMAC)	Preoperative pain
Yarnitsky et al., 2008	DNIC	Pain (NRS)	Baseline pain

AAS: Activity Assessment Scale, **BPI:** Brief Pain Inventory, **CPM:** Conditioned Pain Modulation, **CPP:** Cold Pressor Pain, **CPT:** Cold Pain Threshold, **DNIC:** Diffuse Noxious Inhibitory Control, **EPT:** Electrical Pain Threshold, **EST:** Electrical Sensation Threshold, **HPT:** Heat Pain Threshold, **KOOS:** Knee Injury and Osteoarthritis Outcomes Score, **MPS:** Mechanical Pain Sensitivity, **NDI:** Neck Disability Index, **NRS:** Numerical Rating Scale, **ODI:** Oswestry Disability Index, **PDS:** Post-traumatic Stress Diagnostic Scale, **PDT:** Pain Detection Threshold, **PPT:** Pressure Pain Detection Threshold, **PTT:** Pain Tolerance Threshold, **RMDQ:** Roland-Morris Disability Questionnaire, **THPR:** Tonic Heat Pain Response, **TS:** Temporal Summation, **VAS:** Visual Analogue Scale, **WOMAC:** Western Ontario and McMaster Universities Osteoarthritis Index

3.4.5. Outcome prediction by baseline QST (primary analysis)

3.4.5.1. Prediction of clinical pain by baseline QST

Unadjusted (r) correlation data were available from 18 studies that permitted meta-analysis examining the ability of QST to predict follow-up pain. The pooled unadjusted r value among the included studies was 0.31 (95%CI: 0.23 to 0.38) (**Figure 11**). I^2 calculations indicated 36% of heterogeneity (p=0.07). Funnel plot for studies reporting unadjusted correlations were symmetrical suggesting little or no bias (Egger's test = -1.0, p=0.32) (**Figure 12A**).

Subgroup analyses according to risk of bias showed that unadjusted correlations for studies with low (r = 0.28, 95%CI 0.19 to 0.38; I^2=53%, p=0.01) and moderate (r=0.34, 95%CI: 0.17 to 0.48; I^2=0%, p=0.50) risk of bias were similar to those reported in **Figure 11**.

Seven studies (Izumi et al., 2017, Edwards et al., 2016a, Mendonca et al., 2016, Petersen et al., 2016, Vaegter et al., 2016, Henriksen et al., 2014, Wylde et al., 2013) reported unadjusted correlation data between baseline QST and change in pain as observed between two time points (baseline and follow-up). When pooled they yielded an overall r=0.32 (95%CI: 0.19 to 0.44) and heterogeneity of 29% (p=0.21).

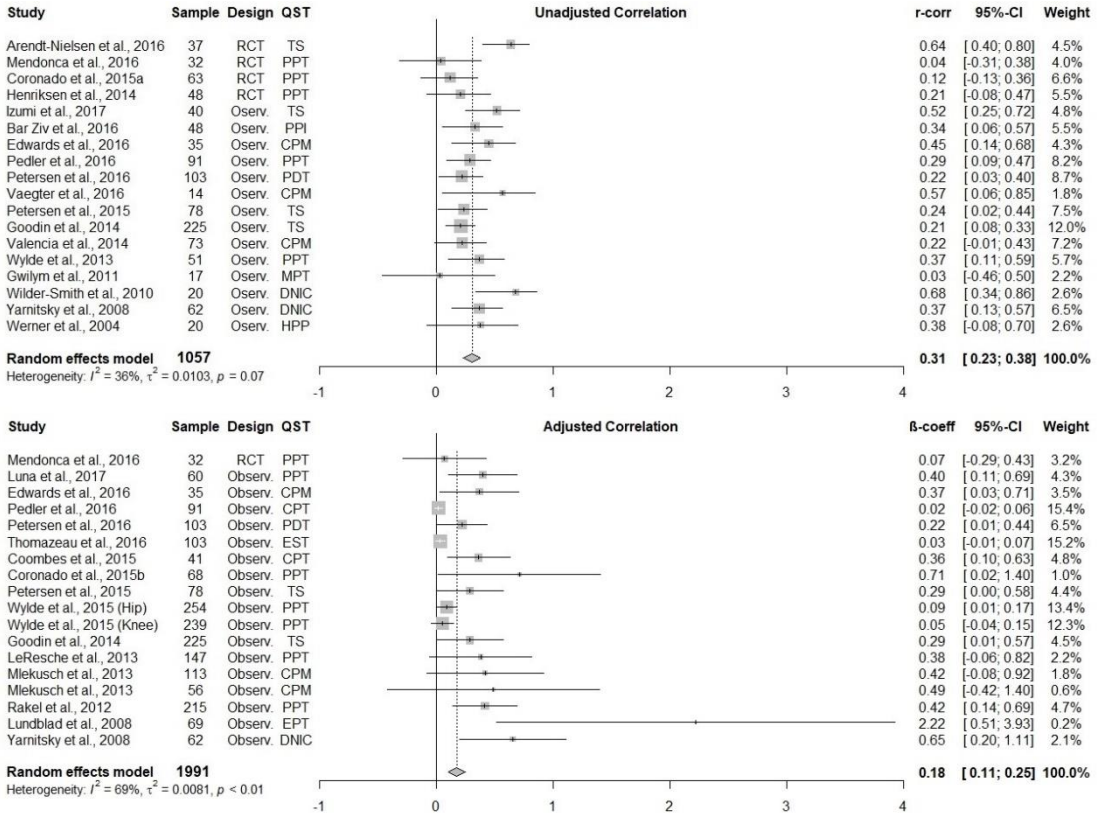

Figure 11. Forest plots showing the overall association (r-correlations and β-coefficients) between QST and follow-up pain.

CI: *Confidence Interval,* ***CPM:*** *Conditioned Pain Modulation,* ***CPT:*** *Cold Pain Detection Threshold,* ***DNIC:*** *Diffuse Noxious Inhibitory Control* ***EPT:*** *Electrical Pain Threshold,* ***EST:*** *Electrical Sensation Threshold,* ***MEP:*** *Motor Evoked Potentials,* ***Observ.:*** *Observational Cohort Study* ***PDT:*** *Pressure Detection Threshold,* ***PPT:*** *Pressure Pain Detection Threshold,* ***QST:*** *Quantitative Sensory Testing,* ***RCT:*** *Randomised Controlled Trial* ***TS:*** *Temporal Summation.*

Forest plot showing the pooled unadjusted (0.31, 95%CI: 0.23 to 0.38) and adjusted correlation (0.18, 95%CI: 0.11 to 0.25) of QST modalities with musculoskeletal pain. The Unadjusted Correlation plot has been derived through the incorporation of Correlation-Coefficient data (Pearson's or Spearman's r) expressing a univariate association unadjusted by other factors whereas the adjusted correlation plot has been derived through the incorporation of β-coefficient data from linear or logistic regressions expressing a multivariate association.

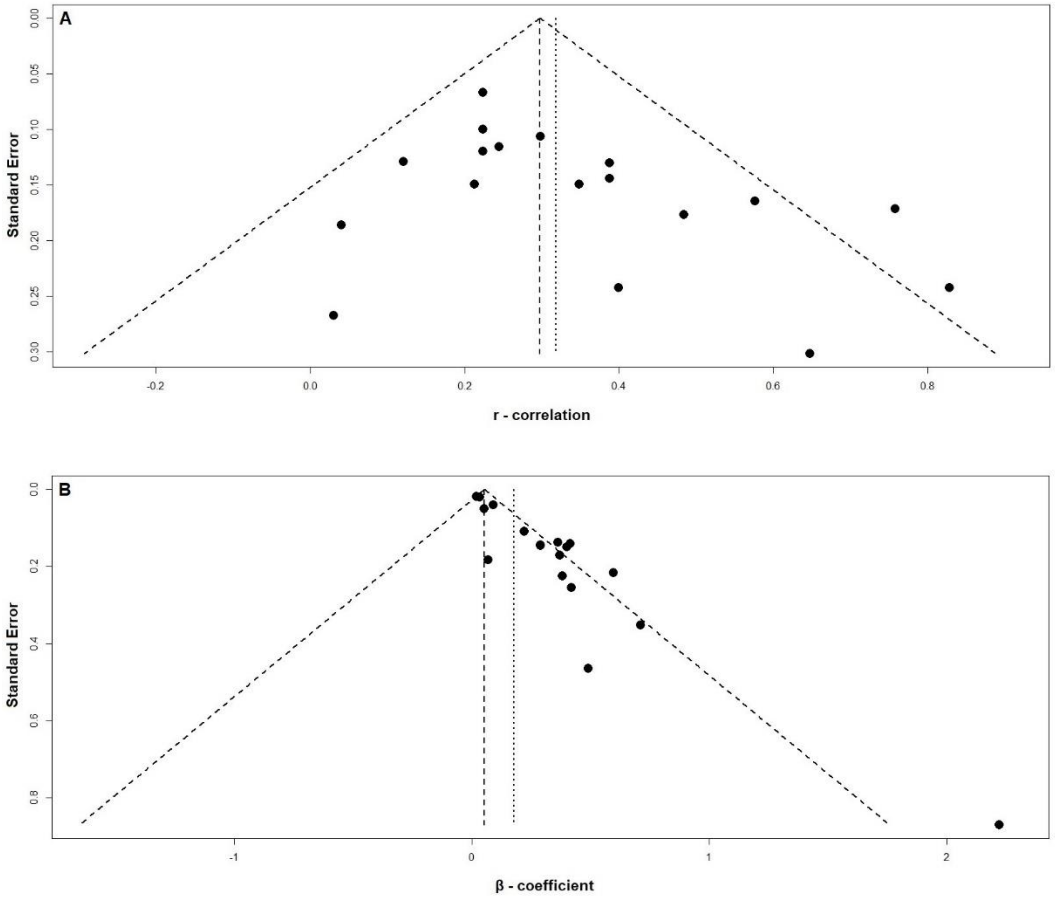

Figure 12. Funnel plots for QST studies (n=18) examining the capacity of pain hypersensitivity (as measured by QST) to predict or associate with pain at follow-up depicting (A): unadjusted (r-correlation) data with little or no indication of publication bias due to their symmetrical presentation and (B) adjusted (β-coefficient) data with an indication of publication bias due to asymmetry. The axes on both graphs are different scales.

3.4.5.2. Prediction of disability by baseline QST

Eleven studies (Dubois et al., 2016, Pedler et al., 2016, Coombes et al., 2015, Coronado et al., 2015b, Valencia et al., 2014, Jull et al., 2013, Sterling et al., 2012, Sterling et al., 2011, Walton et al., 2011, Aasvang et al., 2010, Sterling et al., 2005) reported disability outcomes and most of these (7/11) included participants with WAD. Meta-analysis revealed a mean unadjusted correlation between baseline QST and disability outcome of 0.30 (95%CI: 0.19 to 0.40) (**Figure 13**). I^2 calculations indicated heterogeneity of 0% (p=0.72) for the unadjusted correlation subset. Funnel plot and Egger's test did not show significant asymmetry for the unadjusted dataset (Egger's test =-0.10, p=0.93) (**Figure 14A**).

Subgroup analyses according to risk of bias showed that studies with low risk of bias yielded similar unadjusted correlation (r=0.30, 95%CI 0.18 to 0.41; I^2=0%, p=0.51) to the correlation reported in **Figure 13**. Meta-analysis of unadjusted correlation data from the small number of studies with moderate risk of bias was not feasible for disability.

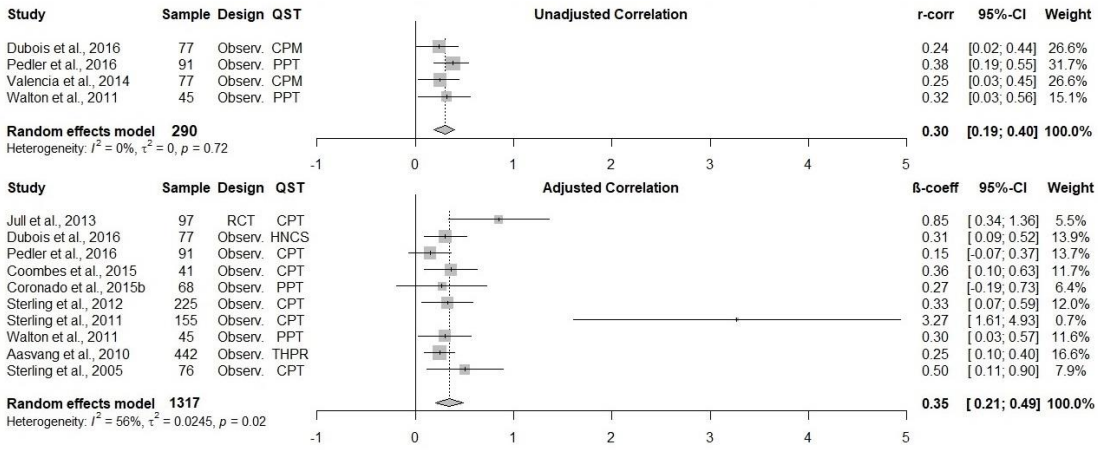

Figure 13. Forest plot showing the overall association (r-correlations and β-coefficients) between QST and follow-up disability in musculoskeletal conditions.

***CI:** Confidence Interval, **CPM:** Conditioned Pain Modulation, **CPT:** Cold Pain Detection Threshold, **EPT:** Electrical Pain Threshold, **HNCS:** Heterotopic Noxious Counter-Stimulation, **Observ.:** Observational Cohort Study **PPT:** Pressure Pain Detection Threshold, **QST:** Quantitative Sensory Testing, **RCT:** Randomised Controlled Trial, **THPR:** Tonic Heat Pain Response.*

Forest plot showing the pooled unadjusted (0.30, 95%CI: 0.19 to 0.40) and adjusted correlation (0.35, 95%CI: 0.21 to 0.49) of QST modalities with musculoskeletal disability. The unadjusted correlation plot has been derived through the incorporation of correlation-coefficient data (Pearson's or Spearman's r) expressing a univariate association unadjusted by other factors, whereas the adjusted correlation plot has been derived through the incorporation of β-coefficient data from linear or logistic regressions expressing a multivariate association.

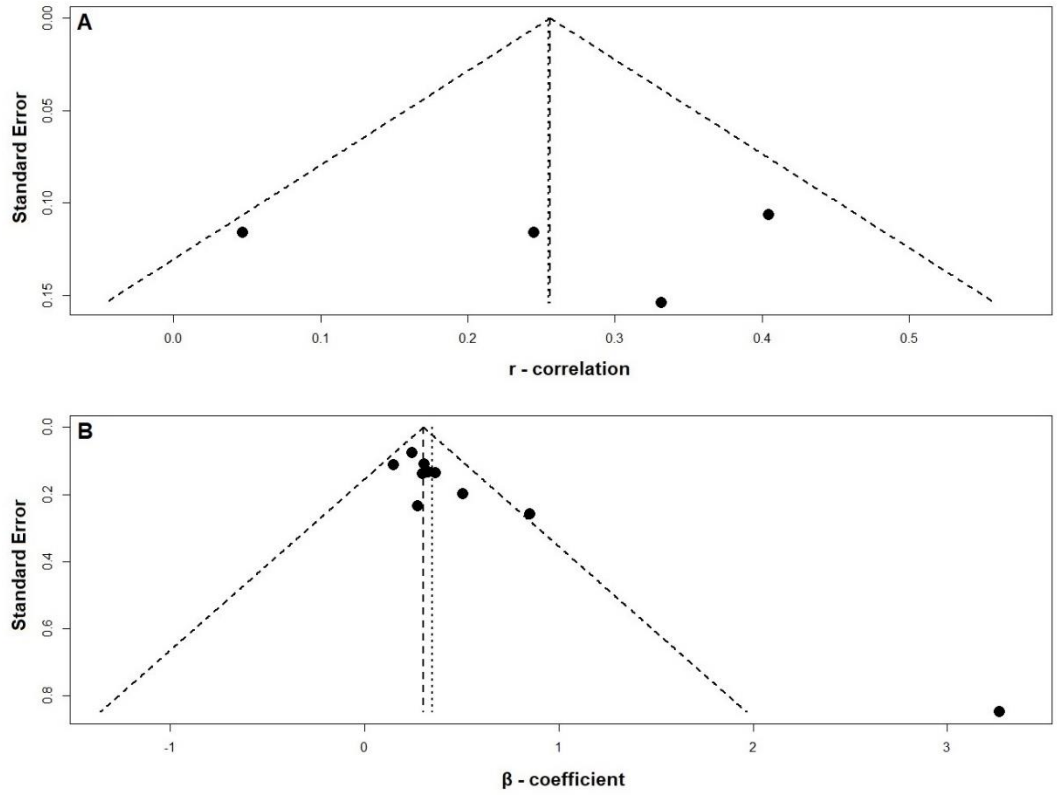

Figure 14. Funnel plot for QST studies examining the capacity of pain hypersensitivity (as measured by QST) to predict or associate with disability at follow-up in musculoskeletal conditions depicting A: unadjusted (r-correlation) data and B: adjusted (β-coefficient) data. The axes on both graphs are different scales.

3.4.5.3. Prediction of negative affect by baseline QST

Three studies (Goodin et al., 2014, Valencia et al., 2014, Sterling et al., 2011) examined whether QST can predict pain-related negative affect (depression or anxiety), of which 2 studies (2011, 2014) reported statistically significant prediction of depression and post-traumatic stress disorder (**Appendix 2**).

3.4.5.4. Effect of study design on outcome prediction by baseline QST

Post-hoc subgroup analyses for unadjusted correlations and heterogeneity were similar for cohort studies (r=0.31, 95%CI 0.24 to 0.38; I^2=14%, p=0.30) as for RCTs (r=0.27, 95% CI: -0.02 to 0.52; I^2=73%, p=0.01), and therefore similar to the overall models presented in **Figure 11**.

3.4.5.5. Effect of QST anatomical site on outcome prediction by baseline QST

We tested the levels of association between QST application site (sites of pathology or remote) and pain. Subgroup analyses of unadjusted correlation data for application of QST at the site of pathology showed r=0.30, 95%CI: 0.22 to 0.38; I^2=43%, p=0.03. Unadjusted correlation data for application of QST at a remote site showed r=0.19, 95%CI: 0.07 to 0.30; I^2=23%, p=0.27. Pooling of distal site data for subgroup analysis was not feasible.

3.4.5.6. Effect of QST modality on outcome prediction by baseline QST

Multiple subgroups were analysed to examine whether specific QST modalities (or similar groups of those) could predict follow-up pain or pain-related disability (**Table 11**). In terms of pain, the pooled unadjusted correlation results between static modalities such as pain detection threshold (mechanical, thermal, electrical) and clinical pain outcomes were lower (0.20) than the 0.31 presented in **Figure 11** yielding a range from 0.14 to 0.20 and I^2 values of 0%. PPT demonstrated a pooled unadjusted correlation (r) of 0.20 (95%CI: 0.11 to 0.29) and heterogeneity of 0% (p=0.57). Higher pooled unadjusted correlation values were given by TS (0.37, 95% CI: 0.17 to 0.54) and CPM alone (0.36, 95% CI: 0.20 to 0.50) and by a model including only dynamic modalities (CPM and TS) (0.38, 95% CI: 0.26 to 0.49) with displayed heterogeneity of 69% (p=0.02), 43% (p=0.12), and 53% (p=0.02) respectively. Post-hoc subgrouping of CPM according to conditioning stimulus (cold water immersion), and subgrouping of predictive capacity in axial (LBP and Neck pain) or peripheral (OA) pathologies yielded similar pooled unadjusted correlations as those presented in **Table 11**. CPM-related post-hoc subgroup analyses displayed either complete absence (0%) or non-statistically significant (p>0.05) heterogeneity. Pooling of TS data according to type of stimulus for post-hoc subgroup analysis was not feasible. Insufficient data precluded also meta-analysis of unadjusted correlation values for disability outcomes.

Table 11. Associations between QST modalities and pain or disability

	QST Modality	Clinical Outcome	Number of Studies	Sample Size	Overall Correlation	95% Confidence Intervals	I^2	I^2 p-value
Unadjusted Correlation (r)	**PDT (All)***	Pain	10	576	0.19	0.10 – 0.27	0%	0.75
	PDT (Mechanical)	Pain	9	503	0.20	0.11 – 0.28	0%	0.73
	PDT (Thermal)	Pain	5	298	0.16	0.04 – 0.27	0%	0.88
	PPT	Pain	7	466	0.20	0.11 – 0.29	0%	0.57
	HPT	Pain	4	207	0.14	0.00 – 0.27	0%	0.82
	CPM	Pain	6	282	0.36	0.20 – 0.50	43%	0.12
	TS	Pain	4	380	0.37	0.17 – 0.54	69%	0.02
	Dynamic Mods†	Pain	9	943	0.38	0.26 – 0.49	53%	0.03
	PDT (All)	Disability	3	213	0.25	0.03 – 0.45	63%	0.07
Adjusted Correlation (β)	**PPT**	Pain	11	1378	0.14	0.05 – 0.23	72%	<0.01
	CPT	Pain	3	279	0.14	-0.10 – 0.37	68%	0.04
	CPM	Pain	5	413	0.35	0.15 – 0.54	0%	0.41
	TS	Pain	3	450	0.26	0.08 – 0.44	0%	0.83
	Dynamic Mods†	Pain	7	716	0.33	0.19 – 0.47	0%	0.80
	PDT (All)	Pain	13	1488	0.17	0.08 – 0.26	68%	<0.01
	PDT (Thermal)	Pain	4	461	0.13	-0.01 – 0.30	74%	0.01
	PDT (All)	Disability	8	1195	0.35	0.16 – 0.55	63%	<0.01
	PDT (Thermal)	Disability	7	1127	0.37	0.16 – 0.58	69%	<0.01
	CPT	Disability	6	685	0.48	0.19 – 0.77	74%	<0.01
	PPT	Disability	3	256	0.02	-0.01 – 0.05	86%	<0.01

CPM: Conditioned Pain Modulation, **CPT:** Cold Pain Threshold, **HPT:** Heat Pain Threshold, **PDT:** Pain Detection Threshold, **PPT:** Pressure Pain Detection Threshold, **TS:** Temporal Summation

*PDT (All) includes all pain detection threshold modalities such as Pain Pressure Detection Threshold, Pain Pressure Tolerance Threshold, Electrical Pain Threshold, Cold Pain Detection Threshold, Heat Pain Detection Threshold

†Dynamic modalities include Conditioning Pain Modulation and Temporal Summation data taken across studies and fit into the same model.

3.4.6. Adjusted associations of musculoskeletal outcomes with baseline QST (secondary analyses)

3.4.6.1. Association of clinical pain with baseline QST

Adjusted (β) correlation data were also available from 18 studies that permitted meta-analysis examining the association of baseline QST with follow-up pain when other variables are taken into account. The pooled adjusted correlation among the included studies was 0.18 (95%CI: 0.11 to 0.25) out of which, in post-hoc analysis, studies that adjusted for baseline pain (13/18) displayed a pooled adjusted correlation of 0.13 (95%CI: 0.06 to 0.20). I^2 calculations indicated 69% heterogeneity ($p<0.01$) for the adjusted correlation dataset and 72% ($p<0.01$) for the subset that adjusted for baseline pain. Funnel plots for adjusted correlations deviated to the right (0.00 to 2.22) (Egger's test = 10.0, $p<0.0001$) (**Figure 12B**) indicating publication bias.

Subgroup analyses according to risk of bias showed that adjusted correlation for studies with low risk of bias were similar (β=0.12, 95%CI 0.06 to 0.18; I^2=61%, $p<0.01$) to those reported in **Figure 11** but higher (β=0.43, 95%CI 0.21 to 0.65; I^2=32%, p=0.22) for studies with moderate risk of bias.

3.4.6.2. Association of disability with baseline QST

Meta-analysis revealed a mean adjusted correlation between baseline QST and disability outcome of 0.35 (95%CI: 0.21 to 0.49) with I^2 calculations yielding heterogeneity of 59% ($p = 0.01$) (**Figure 13**). Funnel plot and Egger's test indicated significant asymmetry for the adjusted dataset (Egger's test = 4.3, $p < 0.01$) (**Figure 14B**).

Subgroup analyses according to risk of bias showed that studies with low risk of bias yielded similar pooled adjusted correlation (β = 0.35, 95%CI 0.18 to 0.52; I^2 = 64%, $p < 0.01$) to the correlation reported in **Figure 13**. Meta-analysis of adjusted correlation data from studies with moderate risk of bias was not feasible for disability.

3.4.6.3. Effect of site of clinical pain on association with baseline QST

Multiple subgroup analyses according to musculoskeletal condition revealed similar pooled adjusted correlation for OA (β=0.30, 95%CI: 0.18 to 0.42; I^2=0%, p=0.89), higher adjusted correlation for LBP (β=0.46, 95%CI: 0.16 to 0.75; I^2=0%, p=0.72) and

lower adjusted correlation for post-operative pain (β=0.13, 95%CI: 0.02 to 0.24; I^2=77%, p<0.01). Meta-analysis of the subgroup of studies reporting an association between baseline QST and WAD-related disability indicated an adjusted correlation of 0.47 (95%CI: 0.18 to 0.76) and significant heterogeneity (I^2=74%, p<0.01). Post-hoc subgroup analysis was also carried out to explore the degree of association between QST and clinical pain according to its anatomical site as adjusted correlation data were available both for peripheral joint and axial pain. Models for adjusted correlation were similar for peripheral joint pain (β=0.22, 95%CI: 0.12 to 0.32) as for axial pain (β=0.22, 95%CI: 0.00 to 0.43) and therefore similar to the overall models presented in **Figure 11**. Heterogeneity in the axial pain model (I^2=49%, p=0.08) was slightly lower than for peripheral joints (I^2=73%, p=<0.01).

3.4.6.4. Effect of QST modality on association with baseline QST

Pooled adjusted correlation results between static modalities such as pain detection threshold (mechanical, thermal, electrical) and clinical pain outcomes approximated the 0.20 presented in **Figure 11** yielding a range from 0.13 to 0.17 and I^2 values from 68% to 74% (**Table 11**). PPT as a stand-alone modality demonstrated a pooled adjusted correlation (β) of 0.14 (95%CI: 0.05 to 0.23) and heterogeneity of 72% (p<0.01). Pooled adjusted correlation values by CPM alone (0.35, 95%CI: 0.15 to 0.54) and by a model including only dynamic modalities (CPM and TS) (0.33, 95% CI: 0.19 to 0.47) yielded higher values than those in the overall models in **Figure 11**. Both analyses displayed heterogeneity of 0% (p=0.41 and 0.80 respectively).

For studies reporting disability as the clinical outcome, subgroup analysis of thermal pain detection threshold modalities showed a pooled adjusted correlation of 0.37 (95% CI: 0.16 to 0.58), and the subset using cold as the thermal stimulus revealed pooled adjusted correlation for cold pain detection threshold (CPT) of 0.48 (95% CI: 0.19 to 0.77). Both subgroups showed statistically significant heterogeneity (p<0.01) with I^2 scores of 69% and 74%.

3.5. Discussion

This systematic review and meta-analysis demonstrates a predictive relationship between baseline QST, a measure of pain hypersensitivity, and musculoskeletal pain and disability at follow-up. This is demonstrated across multiple musculoskeletal conditions (OA, LBP, WAD, post-operative pain) affecting different anatomical sites

(knee, hip, low back, neck, shoulder), and across different QST modalities and study contexts (cohort studies and RCTs). The results of this review show that pain hypersensitivity contributes to prognosis. QST might help identify people who could most benefit from interventions aiming to improve pain and disability.

Previous systematic reviews have been less conclusive on the ability of QST to predict longitudinal outcomes in patients with peripheral musculoskeletal conditions, heathy volunteers, surgical patients and patients with chronic pain (O'Leary et al., 2017, Sangesland et al., 2017, Grosen et al., 2013). The present systematic review extends these reports by demonstrating longitudinal prediction of several outcomes across a range of musculoskeletal conditions, and addressing through meta-analysis the limited power of individual studies. It was shown that QST might predict other outcomes beyond pain and disability such as depression in people with musculoskeletal pain. Depression and chronic pain may share similar brain activation pathways as shown by magnetic resonance imaging (Sheng et al., 2017, Han and Pae, 2015, Mutschler et al., 2012), and shared mechanisms might explain shared predictive factors. Future studies might explore whether QST can predict additional outcomes such as ability to self-care or absenteeism/presenteeism. Our findings also indicate that prediction of poor outcomes by QST evidence of pain hypersensitivity is not disease specific, applying similarly to axial and non-axial musculoskeletal pain. QST can also predict acute post-operative pain (Abrishami et al., 2011, Werner et al., 2010, Edwards et al., 2005).

No significant differences in outcome prediction by QST were found between data from cohort studies and those from RCTs, supporting generalisation of conclusions from our findings. Treatments received by participants might be similar between cohort studies and RCTs, and generalisation of our findings to other treatment contexts should be cautious. Future research might explore whether baseline QST evidence of hypersensitivity can predict *good* response to novel treatments that more effectively reverse hypersensitivity, such as antidepressant (i.e. duloxetine) or anticonvulsant (i.e. Gabapentin) drugs (Häuser et al., 2012), psychotherapy (i.e. CBT) (Salomons et al., 2014) and paced exercises (Erickson et al., 2011).

The primary purpose of this study was to investigate outcome prediction in people with musculoskeletal pain. Those destined to experience worse outcomes stand to gain more from effective interventions. Predictors of poor outcomes might also shed some light on mechanisms and potential targets for interventions aiming to improve outcome. Univariate prediction is important for identifying people at risk of poor

outcome, but provides only very limited mechanistic understanding. Multiple regression provides greater insight into causal relationships by adjusting for other factors in order to reduce confounding (Hayden et al., 2009, Mallen et al., 2007) and bias (Hayden et al., 2006, Herbert, 2014). Outcome prediction by QST appeared stronger in unadjusted than in adjusted correlation analyses but the magnitude of these two values should not be compared directly as they are measured through different scales. However, weaker associations in adjusted analyses might be expected in light of the cross-sectional associations between QST and outcome measures at baseline (Fingleton et al., 2015, Suokas et al., 2012, Goldsmith et al., 2012), and the well-recognised prediction of an outcome measure by its baseline value. Significant outcome prediction by QST in adjusted analyses suggests a direct effect of pain hypersensitivity on musculoskeletal outcome.

Pain hypersensitivity has been identified in multiple reports of chronic pain conditions as an underlying pathophysiology (Smart et al., 2012b, Banic et al., 2004, Sterling et al., 2003b) and has been associated with the development of additional symptoms, such as fatigue and mood disturbance (Akin-Akinyosoye et al., 2018), that can further impact on prognosis (Bourke et al., 2015, Woolf, 2011). QST can identify the presence of pain hypersensitivity in people with OA (Fingleton et al., 2015, Suokas et al., 2012) and WAD (Goldsmith et al., 2012). Our findings that QST can predict clinical outcomes in people with musculoskeletal pain indicate that pain hypersensitivity could be investigated as a mechanism for poorer prognosis. This is further supported by a recent study (O'Leary et al., 2018), published after our database search end-date, showing that patients with knee OA and higher TS responded poorly to exercise programs.

Possible mechanisms by which pain hypersensitivity might lead to worse outcomes include alterations in pain processing which can persist despite treatment (Arendt-Nielsen et al., 2018, Baliki et al., 2006, Sterling et al., 2003b). Pain hypersensitivity might also pose a barrier to gaining benefit from current treatments, for example by reducing treatment uptake or engagement (Bushnell et al., 2013, Smart et al., 2012a, Jull et al., 2007). Interventions targeting hypersensitivity might have benefit across a range of musculoskeletal conditions.

Various QST modalities have been designed to address different mechanisms of hypersensitivity, body regions or medical conditions and therefore might differentially predict outcome. Pain hypersensitivity may be due to changes in the peripheral or central nervous system. Alterations in pain thresholds using deep stimuli, such as

those used for pressure pain detection thresholds at sites local to musculoskeletal pathology, might predominantly reflect peripherally-driven pain hypersensitivity. However, dynamic QST modalities such as CPM or TS were most strongly associated with musculoskeletal pain and disability, suggesting a possible role for centrally-driven pain hypersensitivity (Arendt-Nielsen and Yarnitsky, 2009). CPM reflects cerebral processes that are implicated in depressive or psychological disorders even in the absence of nociceptive drive (Arendt-Nielsen et al., 2018). CPM might therefore be associated with psychological mechanisms contributing to chronic musculoskeletal pain. Thermal pain and pain in response to punctate stimulation are mediated by cutaneous nerves, rather than those localised within musculoskeletal tissues. It was established that thermal modalities in general, and cold pain thresholds in particular, were associated with pain-related disability. Data leading to these conclusions were predominantly from studies of whiplash-related pain and disability (Goldsmith et al., 2012), and condition-specific injury mechanisms might be responsible for disturbances to the nervous system that differ between conditions. Further research might explore whether a contribution of thermal QST modalities to worst outcomes might also apply to other musculoskeletal diagnoses.

Centrally-driven pain hypersensitivity has also been associated with reduced pain detection thresholds at sites remote from the site of pathology (Nijs et al., 2010, Schliessbach et al., 2010, Herren-Gerber et al., 2004), whereas increased sensitivity at the site of pathology might reflect peripheral sensitisation alone plus augmentation by central sensitisation (Suokas et al., 2012). Our findings that hypersensitivity at a remote site can predict worse musculoskeletal outcomes further supports a contribution from central sensitisation. However, pain thresholds at the site of clinical pain also predicted outcomes and a contribution of peripheral sensitisation to prognosis deserves further study.

Interpretation of our findings is subject to a number of limitations. Outcome prediction can be influenced by the type of therapeutic intervention that participants receive, and the effect of treatments on pain hypersensitivity cannot be determined from the available data. Significant heterogeneity between studies in several of our subgroup analyses was found suggesting that factors additional to those explored here might influence the ability of QST predict musculoskeletal outcomes. Funnel plots displayed significant asymmetry suggesting possible publication bias, particularly for adjusted analyses. However, 26 of the 37 reports were judged to be of low risk of bias and the remaining studies of only moderate risk. Sensitivity analyses showed that the levels of bias did not have a significant effect on our main findings. Our search strategy was

intentionally broad, but it remains possible that not all relevant studies have been identified. Small numbers of studies and participants limit our ability to exclude differences between some subgroups and our use of a small number of studies in several analyses might limit generalisability. The current meta-analyses suggests relatively weak predictive ability (Portney and Watkins, 2009) for QST, with correlations only sometimes and marginally above 0.30, a threshold considered to be clinically meaningful (Revicki et al., 2008). However, what consists a meaningful deviation from that threshold was not established and analyses regarding the magnitude of those deviations were not performed. Inferences in relation to pooled predictive values must be drawn with caution. A systematic review with meta-analysis of cross-sectional studies (Hübscher et al., 2013) also indicated that pain detection thresholds might not present a clinically important correlation with pain or disability in spinal pain. However, the significant association even in adjusted analyses between QST and musculoskeletal outcomes might suggest underlying mechanisms and potential targets for intervention. Other prognostic factors, including psychological factors such as depression or anxiety (Burke et al., 2015) and maladaptive beliefs such as catastrophising or fear avoidance (Edwards et al., 2016b) might complement outcome prediction by QST.

3.6. Conclusion

Identifying which patients might be at particular risk of poor outcome is important in order to identify those who are most likely to benefit from treatment. QST modalities with stimuli applied at the site of clinical pain, dynamic modalities such as CPM and TS, and thermal pain detection thresholds appeared to have the greatest potential. PPTs have advantages of ease of application in clinical settings, low cost and high user and patient acceptability. Further refinement of QST and adoption of standardised QST protocols are recommended. Important methodological variation between published studies was identified, particularly reflected by the range of stimulus types used in dynamic modalities. Studies which used blunt pressure as a testing stimulus and hand immersion in cold water as a conditioning stimulus contributed most to evidence that CPM can predict musculoskeletal outcomes. However, available data did not enable the drawing of robust conclusions on superiority between different stimulus types for TS. Additional confirmatory research is required in larger and more homogenous populations, inside and outside the musculoskeletal spectrum. Translation into clinical practice requires also feasibility in clinical contexts, acceptability to patients, and evidence that implementation improves

patient outcomes. Future studies should aim to define reliability of specific QST approaches and establish clinically meaningful thresholds in specific pathologies in order to translate QST from a research tool into a clinical decision aid for musculoskeletal conditions.

In conclusion, it was shown that QST, an index of pain hypersensitivity, can predict worse musculoskeletal outcomes of pain, disability and negative affect. Our findings are consistent with important contributions from hypersensitivity to outcome, and reducing pain hypersensitivity has potential to improve outcome for people with musculoskeletal conditions.

4. RELIABILITY AND VALIDITY OF A QST PROTOCOL IN HEALTHY PARTICIPANTS AND INDIVIDUALS WITH CLBP

4.1. Introduction

Quantitative sensory testing is an umbrella term for a battery of different modalities (methods of specific tissue stimulation under specific application protocols) that can provide important information about different types of pain processing (Courtney et al., 2010, Arendt-Nielsen and Yarnitsky, 2009) and is widely used to assess how subjects perceive touch, vibration, proprioception, and sharp/blunt pressure sensitivity or sensitivity to cold/heat stimuli (Cruz-Almeida and Fillingim, 2014). QST is considered an optimal method to explore the involvement of central mechanisms responsible for the development or maintenance of local or widespread pain in various musculoskeletal disorders (Arendt-Nielsen et al., 2015, Uddin et al., 2014, Courtney et al., 2010, Fernández-Carnero et al., 2010, Pavlaković and Petzke, 2010, Sterling et al., 2003a, Staud, 2002).

Calculation and presentation of the reliability of clinical tests when pain tests are used to follow up patients or to investigate the effect of any particular treatment facilitates the translation of QST techniques from the laboratory to a clinical setting (Manresa et al., 2011). The use of QST is not common practice in clinical settings because its test-retest reliability has not been very well established in patient populations (Wylde et al., 2011b). Test-retest and inter-rater reliability findings for PPT, TS and CPM are extensively reported in healthy individuals and in people suffering from a chronic condition including CLBP (Balaguier et al., 2016, Kong et al., 2013, Paungmali et al., 2012, Geber et al., 2011, Manresa et al., 2011, Park et al., 2011, Wylde et al., 2011b, Cathcart et al., 2009, Chesterton et al., 2007, Cathcart and Pritchard, 2006, Nussbaum and Downes, 1998). Nevertheless, the development and adaptation of reliable QST protocols is critical to pain-related research in order to standardise the QST testing procedures (Hall et al., 2015).

Even though QST is a method used extensively for the quantification of sensory alterations and the identification of widespread pain in musculoskeletal disorders, inferences regarding statistical associations between modalities (internal validity) as well as associations between modalities and alternative methods of hypersensitivity measurements (external validity) are not very well established or are under-reported.

Widespread pain has been researched by other groups that propose alternate to QST methods for its identification and classification of individuals. The distribution of pain across different body sites and regions along with the existence of comorbidities such as fatigue, unrefreshed sleep, cognitive difficulties, depression, headaches and abdominal pain/cramps are proposed by ACR as a valid widespread pain identification and classification method (Wolfe et al., 2019). The distribution of pain across different scoring grids on a pain manikin has also been recommended as a marker of widespread pain sensitisation (Wylde et al., 2011a, Croft et al., 1996).

Widespread pain is a common characteristic of centrally driven pain hypersensitivity across musculoskeletal conditions (Woolf, 2011). It could be therefore anticipated that pain spread disproportionately and distally to the primary area of pain should be identifiable by all approaches that have been appropriately developed for that purpose. Evidence of associations between different methods of widespread pain identification have been demonstrated in OA, where PPT thresholds at a distal site have been associated with the number of painful sites determined by classifying pain distribution reported by people using a body manikin (Akin-Akinyosoye et al., 2018). Decreased pain thresholds at sites distant to the site of pathology is considered a feature of central sensitisation in many musculoskeletal disorders where central hyperexcitability is prevalent (Hidalgo-Lozano et al., 2010, Schliessbach et al., 2010, Fernández-Carnero et al., 2009, Freeman et al., 2009).

4.2. Aims and Objectives

4.2.1. Aims

The primary aim of this study was to establish the test-retest and inter-rater reliability of PPT, TS and CPM in healthy volunteers and patients with CLBP.

Secondary aims were to ascertain the means of calculating TS and CPM that demonstrate the higher reliability as well as to examine the internal validity of the QST protocol.

In the form of secondary analyses, the study further aimed to determine the external validity of the QST protocol

4.2.2. Objectives

- To establish the intraclass-correlation coefficient (ICC) and concordance correlation coefficient (CCC) for each QST modality (PPT, TS, CPM) for the same rater, between raters and for both study populations (healthy, patients with CLBP).
- To ascertain the limits of agreement (LoA) for each QST modality (PPT, TS, CPM) for the same rater, between raters and for both study populations (healthy, patients with CLBP).
- To establish the internal validity of the protocol by exploring the correlation relationship between the three discrete QST modalities.
- To establish the external validity of the protocol by exploring the correlation between each QST modality with different self-reported pain ratings and the number of painful sites as shaded on a pain manikin.

4.3. Methods

The full methodological details of this study are given in the Methods chapter (**METHODS**) and only a brief outline is presented in this chapter whenever necessary.

4.3.1. Sample size considerations

Calculations yielded that a sample size of 25 individuals for each group of participants could offer enough power for robust statistical analysis. Information on the sample size calculation methods for this study are given in section **2.5** of the methods chapter.

4.3.2. Study participants and recruitment

The study part featuring the healthy participants was approved by the Faculty of Medicine & Health Sciences Research Ethics Committee of the University of Nottingham, United Kingdom (ERN: 264-1803). The study part featuring individuals with LBP was approved by the East Midlands - Nottingham 1 Research Ethics Committee of the Health Research Authority, United Kingdom (REC: 18/EM/0049). Healthy participants were from various departments of the University of Nottingham, whereas patients were recruited from a list of individuals undertaking treatment at the

Sherwood Forest Hospitals NHS Foundation Trust. Details on recruitment as well as inclusion and exclusion criteria for each study group are given in **2.7.1**.

4.3.3. Clinical assessment and application methods

Individuals of both groups were invited to participate in two identical sessions (baseline/follow-up) separated by at least a week and no more than 15 days in a temperature controlled and quiet room. All participants had QST (PPT, TS and CPM) applied to the forearm of their dominant hand. For the purposes of CPM, a blood pressure cuff was applied to their non-dominant hand and was inflated to produce ischaemic pain (conditioning stimulus). Data about self-reported measures of pain and pain distribution were also taken from individuals with CLBP. Details about the university and hospital settings, clinical assessment, QST application methods and self-reported outcome measures are given in **2.8.2.1**, **2.6.2.1**, **2.6.2.2** and **2.6.3.1**.

4.3.4. Data analyses and analytical procedures

Data normality exploration procedures are detailed in **2.9.2.1**. Details on QST calculation and interpretation methods are given in **2.9.2.3**. ICCs were calculated through two-way random effects absolute agreement models to express the test-retest and inter-rater reliability for each modality across both groups as well as for measurements of operationally similar modalities (e.g. comparison of overall PPT value with unconditioned CPM stimulus). Bland and Altman plots to establish the LoA for each modality for a single rater as well as between raters were also explored. Details on ICC calculation and LoA derivation methods are given in **2.10.2.1**. The degree of correlation between each QST modality as well as between QST modalities and self-reported pain intensity and pain distribution methods were examined to establish the internal and external validity of the protocol. Details on internal and external validity testing are given in **2.10.2.2**.

4.4. Results

4.4.1. Data management and transformation

Details about data distribution as well as the results of the log-transformation process can be found in **Appendix 5** and **Appendix 6**. Not all variables demonstrated normal distribution upon testing. Baseline and follow-up TS^{WUD} and TS^{WUR} for the combined

population, follow-up TS^{WUD} and TS^{WUR} for the patient population, and follow-up CPM^{Unc} for the combined population demonstrated distributions significantly different to normal after logarithmic transformation.

4.4.2. Demographics data and clinical characteristics

Participants' demographic data and clinical characteristics are given in **Table 12**. Overall, 28 individuals were invited to participate in the healthy group whereas 40 were invited to participate in the patient cohort. The final sample for each group comprised 25 healthy individuals and 25 patients with LBP forming a total sample size of 50 participants. No drop-outs were observed and participants of both groups completed the protocol in full. There was only one occasion where a participant had to reschedule an appointment due to an unplanned consumption of painkillers. For patients, QST assessments and self-reported outcome measures were taken once and on the same day when the reliability baseline session coincided with the baseline session for the observational study (n=12). In all other cases (n=13) the outcome measures were taken upon entry to the observational study. In regards to CPM, 80% of the patient population indicated pain 4/10 at 209±51mm/Hg without ball gripping whereas all of the healthy participants (100%) performed repeated ball-gripping to induce ischaemic pain.

Table 12. Demographic and clinical data of study subjects in their original groups as well as combined.

	Variables (Unit, Value, Range, %)	Healthy Mean ± SD Median (IQ Range)	Patients Mean ± SD Median (IQ Range)	Combined Mean ± SD Median (IQ Range)
Demographics	**No. Participants**	25	25	50
	Age (y)	31 (28 to 46)	57 (48 to 65)	48 (31 to 58)
	Baseline to Follow-up (d)	8 ± 1	8 ± 1	8 ± 1
	BMI	22.5 (20.8 to 24.4)	27.1 (26.0 to 31.6)	24.5 (22.0 to 28.5)
	Female (%)	60%	68%	64%
	Pain (0-10)	-	6 (5 to 7)	-
	Neuropathic Pain (painDETECT)	-	17 (13 to 21)	-
	Now (0-10)	-	6 (4 to 7)	-
	Strongest (0-10)	-	8 (8 to 9)	-
	Average (0-10)	-	7 (6 to 8)	-

	Quantitative Sensory Testing	Rater 1 Median (IQ Range)	Rater 2 Median (IQ Range)	Rater 1 Median (IQ Range)	Rater 1 Median (IQ Range)	Population Difference t-test (p-values)
Baseline	**PPT** (kPa)	222.0 (176.9 to 249.5)	206.3 (147.0 to 275.4)	271.5 (195.5 to 305.3)	244.5 (185.2 to 306.7)	-0.43 (0.67)
	TS^{WUD} (0-10)	1.2 (0.5 to 2.2)	1.4 (0.5 to 2.2)	1.5 (0.5 to 2.5)	1.3 (0.5 to 2.4)	0.11 (0.91)
	TS^{WUR} (Ratio)	2.5 (1.9 to 3.8)	3.6 (2.0 to 5.4)	5.0 (2.3 to 9.5)	3.0 (2.0 to 5.9)	-2.68 (0.01)
	CPM^{Unc} (kPa)	92.1 (37.2 to 163.6)	120.5 (30.3 to 213.6)	47.0 (-6.9 to 98.0)	53.2 (5.7 to 135.3)	1.67 (0.10)
	$CPM^{PPT\text{-}mean}$ (kPa)	87.2 (50.4 to 119.9)	109.3 (42.1 to 173.0)	55.2 (24.2 to 91.8)	66.5 (35.0 to 118.2)	1.24 (0.22)
Follow-up	**PPT** (kPa)	224.0 (178.4 to 251.9)	-	216.5 (164.6 to 281.6)	228.2 (172.8 to 271.8)	0.31 (0.76)
	TS^{WUD} (0-10)	0.9 (0.3 to 2.0)	-	1.3 (0.4 to 2.3)	1.1 (0.3 to 2.2)	-0.61 (0.55)
	TS^{WUR} (Ratio)	2.6 (1.7 to 4.6)	-	3.5 (2.1 to 7.5)	2.7 (2.0 to 5.8)	-1.31 (0.20)
	CPM^{Unc} (kPa)	55.9 (5.9 to 95.0)	-	38.2 (11.8 to 81.4)	52.9 (11.2 to 82.1)	0.68 (0.50)
	$CPM^{PPT\text{-}mean}$ (kPa)	66.6 (36.9 to 131.0)	-	62.7 (31.0 to 99.3)	65.6 (32.1 to 125.5)	0.47 (0.64)

$CPM^{PPT\text{-}mean}$: 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, **IQ:** Inter-quartile, **kPa:** kiloPascals, **PPT:** Pain Pressure Detection Threshold, **SD:** Standard Deviation, **TS^{WUD}:** Temporal Summation calculated as a difference, **TS^{WUR}:** Temporal Summation calculated as a ratio

4.4.3. Test-retest and inter-rater reliability of QST measures

Intraclass correlation coefficient, CCC, paired t-test and paired Wilcoxon signed-rank test values for all combinations of raters and sessions are presented in **Table 13**. Overall, test-retest ICC values for PPT were 0.77 and 0.89 for the healthy and the patient cohort respectively. TS^{WUD} demonstrated test-retest ICCs of 0.76 for the healthy and 0.78 for the patient group whereas TS^{WUR} ICCs were 0.49 for the healthy and 0.71 for the patient group. Test-retest ICCs for both calculation methods of CPM ($CPM^{PPT\text{-}mean}$, CPM^{Unc}) ranged from -0.10 to 0.50 across both groups. Associated CCCs ranged from 0.77 to 0.87 ($p \leq 0.01$) for all modalities and both types of TS and CPM calculation in the healthy cohort. In the patient cohort, CCCs for PPT, TS^{WUD} and TS^{WUR} ranged from 0.84 to 0.87 ($p<0.01$) whereas CCCs for CPM of both calculation methods did not reach statistical significance ($p>0.05$) for Rater 1. Paired t-tests for PPT, TS^{WUD} and TS^{WUR} and paired Wilcoxon signed-rank tests for both CPM calculation methods ranged from -2.65 to 2.20 (t-test) and 141 to 200 (Wilcoxon test) without reaching statistical significance ($p>0.05$) indicating that there are no significant differences in QST measurements between baseline and follow-up measurements both in the healthy and patient populations.

No significant differences were observed between the two populations as unpaired t-tests between the healthy and the patient group for each modality ranged from -0.43 to 1.24 without reaching statistical significance ($p>0.05$) indicating that the null hypothesis (similarity) is true. Combined patient and healthy participant (n=50) values for ICC, CCC, paired t-test and paired Wilcoxon signed-rank test are also presented in **Table 13**. Test-retest ICCs were 0.84 for PPT, 0.76 and 0.63 for TS^{WUD} and TS^{WUR} respectively, 0.29 for CPM^{Unc} and 0.35 for $CPM^{PPT\text{-}mean}$. CCC values ranged from 0.74 to 0.86 ($p<0.01$) for all modalities and calculation methods apart from CPM^{Unc} (CCC=0.70, p=0.14), indicating no statistically significant concordance. Paired t-tests and paired Wilcoxon signed-rank values for the combined cohorts showed no significant differences between baseline and follow-up measurements.

Table 13. Within and between observer reliability of QST

	Sample Size	**Healthy Test-Retest Reliability (Rater 1** Baseline with **Rater 1** Follow-up)						
	n=25	**ICC**	**95% CI**	**CCC**	**95% CI**	**CCC p-value**	**t/w**	**t/w p-value**
Healthy Group	**PPT**	0.77	0.54 - 0.89	0.83	0.75 - 1.0	<0.01	-0.34	0.74
	TS^{WUD}	0.76	0.52 - 0.89	0.87	0.80 - 1.0	<0.01	1.96	0.06
	TS^{WUR}	0.48	0.11 - 0.73	0.82	0.74 - 1.0	<0.01	-0.08	0.94
	CPM^{Unc}	0.50	0.15 - 0.74	0.77	0.68 - 1.0	0.01	200	0.33
	$CPM^{PPT\text{-}mean}$	0.43	0.06 - 0.70	0.77	0.68 - 1.0	0.01	-0.51	0.61
	Sample Size	**Healthy Inter-rater Reliability (Rater 1** Baseline with **Rater 2** Baseline)						
	n=25	**ICC**	**95% CI**	**CCC**	**95% CI**	**CCC p-value**	**t**	**t p-value**
	PPT	0.86	0.72 - 0.94	0.87	0.80 - 1.0	<0.01	1.46	0.16
	TS^{WUD}	0.88	0.75 - 0.94	0.93	0.87 - 1.0	<0.01	-0.39	0.70
	TS^{WUR}	0.71	0.45 - 0.86	0.83	0.75 - 1.0	<0.01	-2.67	0.01
	CPM^{Unc}	0.55	0.21 - 0.77	0.81	0.72 - 1.0	<0.01	-2.10	0.05
	$CPM^{PPT\text{-}mean}$	0.46	0.09 - 0.72	0.78	0.70 - 1.0	0.01	-1.67	0.11
	Sample Size	**Patient Test-Retest Reliability (Rater 1** Baseline with **Rater 1** Follow-up)						
	n=25	**ICC**	**95% CI**	**CCC**	**95% CI**	**CCC p-value**	**t/w**	**t/w p-value**
Patient Group	**PPT**	0.92	0.83 - 0.96	0.87	0.80 - 1.0	<0.01	2.20	0.08
	TS^{WUD}	0.78	0.56 - 0.86	0.86	0.79 - 1.0	<0.01	0.32	0.75
	TS^{WUR}	0.71	0.44 - 0.86	0.84	0.76 - 1.0	<0.01	1.85	0.08
	CPM^{Unc}	-0.10	-0.44 - 0.27	0.64	0.57 - 1.0	0.70	144	0.63
	$CPM^{PPT\text{-}mean}$	0.44	0.07 - 0.71	0.73	0.65 - 1.0	0.07	141	0.58
	Sample Size	**Combined Population Test-Retest Reliability (Rater 1** Baseline with **Rater 1** Follow-up)						
	n=50	**ICC**	**95% CI**	**CCC**	**95% CI**	**CCC p-value**	**t/w**	**t/w p-value**
Combined Group	**PPT**	0.84	0.74 - 0.91	0.86	0.80 - 1.0	<0.01	0.77	0.44
	TS^{WUD}	0.76	0.62 - 0.86	0.86	0.81 - 1.0	<0.01	1343.5	0.52
	TS^{WUR}	0.63	0.42 - 0.77	0.83	0.77 - 1.0	<0.01	1313.0	0.67
	CPM^{Unc}	0.29	0.01 - 0.52	0.70	0.64 - 1.0	0.14	1361.0	0.44
	$CPM^{PPT\text{-}mean}$	0.35	0.09 - 0.57	0.74	0.69 - 1.0	<0.01	1246.5	0.98

CCC: Concordance Correlation Coefficient, **CI:** Confidence Interval, **CPM:** Conditioning Pain Modulation, **PPT:** Pain Pressure Detection Threshold, **ICC:** Intra-class Correlation Coefficient, **SD:** Standard Deviation, **t/w:** paired t-test or paired Wilcoxon-signed rank test, **TS:** Temporal Summation

Test-retest Bland-Altman plots for each QST modality across study populations, alongside the corresponding ±LoA (95%CIs) between baseline and follow-up measurements are shown in **Figure 15** and **Figure 16** and display no significant differences from zero. In healthy participants, the mean difference between the measurements of PPT was -0.02 (-LoA: -0.71, 95%CI: -0.96 to -0.46; +LoA: 0.66, 95%CI: 0.41 to 0.91), of TS^{WUD} was 0.29 (-LoA: -1.18, 95%CI: -1.72 to -0.64; +LoA: 1.77, 95%CI: 1.23 to 2.31) and of CPM^{Unc} was 19.59 (-LoA: -199.75, 95%CI: -279.76 to -119.74; +LoA: 238.94, 95%CI: 158.93 to 318.95). In patients with LBP, the demonstrated mean differences were 0.09 (-LoA: -0.30, 95%CI: -0.44 to -0.16; +LoA: 0.47, 95%CI: 0.33 to 0.61) for PPT, 0.05 (-LoA: -1.52, 95%CI: -2.09 to -0.94; +LoA: 1.62, 95%CI: 1.05 to 2.19) for TS^{WUD} and -16.89 (-LoA: -271.24, 95%CI: -364.02 to -178.46; +LoA: 237.46, 95%CI: 114.68 to 330.24) for CPM^{Unc}.

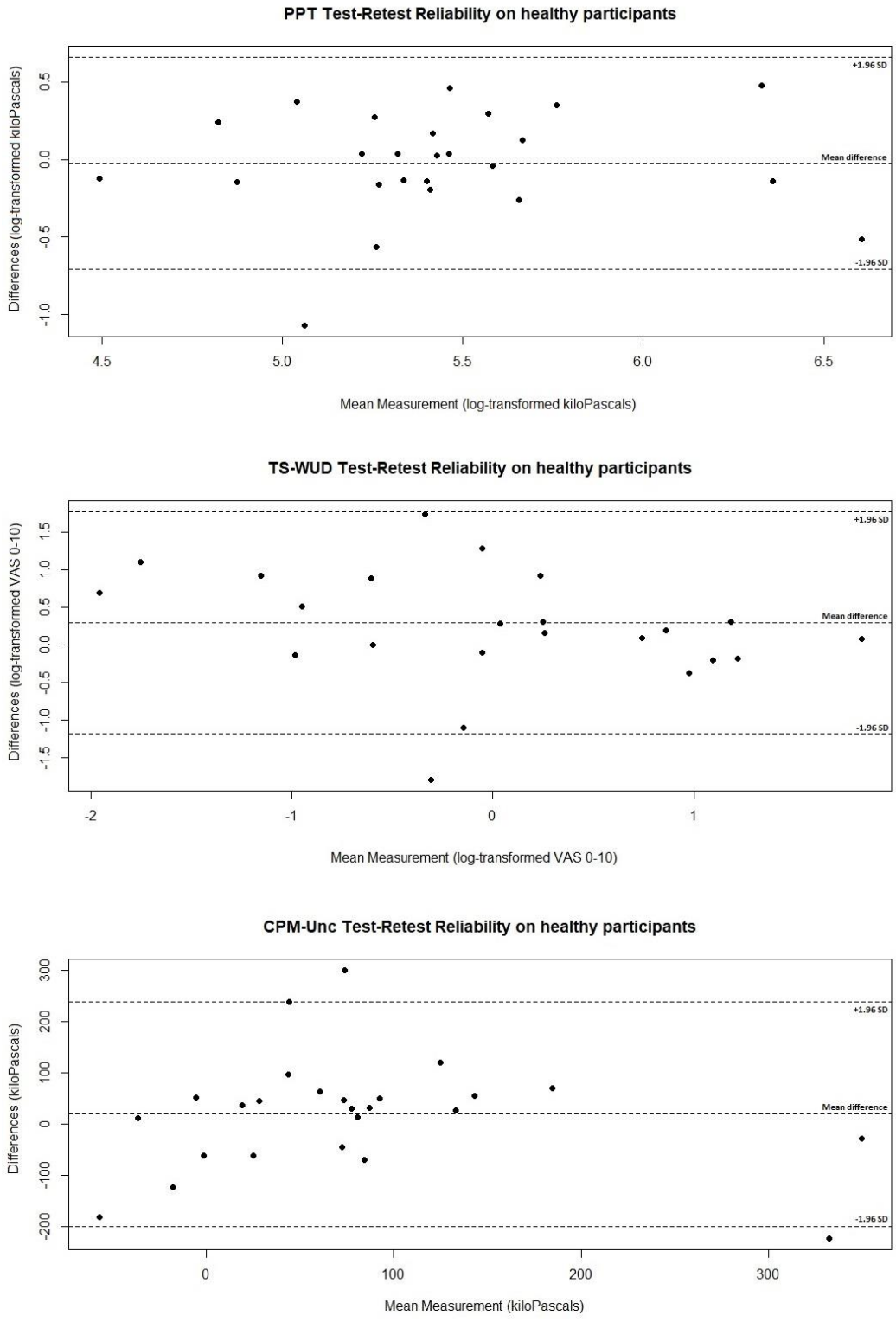

Figure 15. Bland-Altman plots for PPT, TS^{WUD} and CPM^{Unc} after measurements on healthy participants at baseline and follow-up.

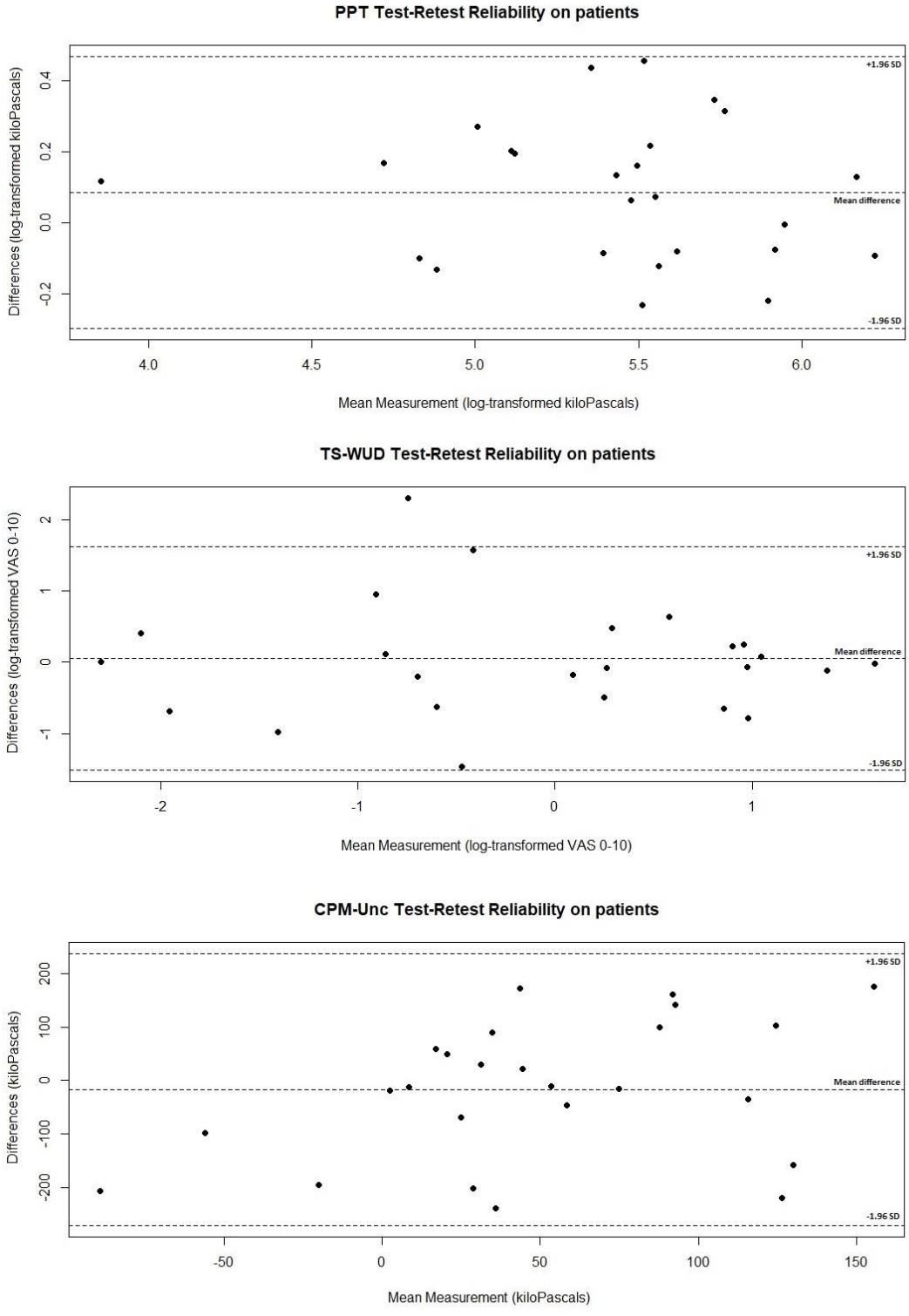

Figure 16. Bland-Altman plots for PPT, TS^{WUD} and CPM^{Unc} after measurements on patients with LBP at baseline and follow-up.

Graphically similar Bland-Altman plots are observed in for TS^{WUR} and $CPM^{PPT\text{-}mean}$ across populations (mean differences for healthy: TS^{WUR} = -0.01, $CPM^{PPT\text{-}mean}$ = -10.06; mean differences for patients: TS^{WUR} =0.05, $CPM^{PPT\text{-}mean}$ = -19.25), revealing no signs of measurement bias (**Figure 17** and **Figure 18**).

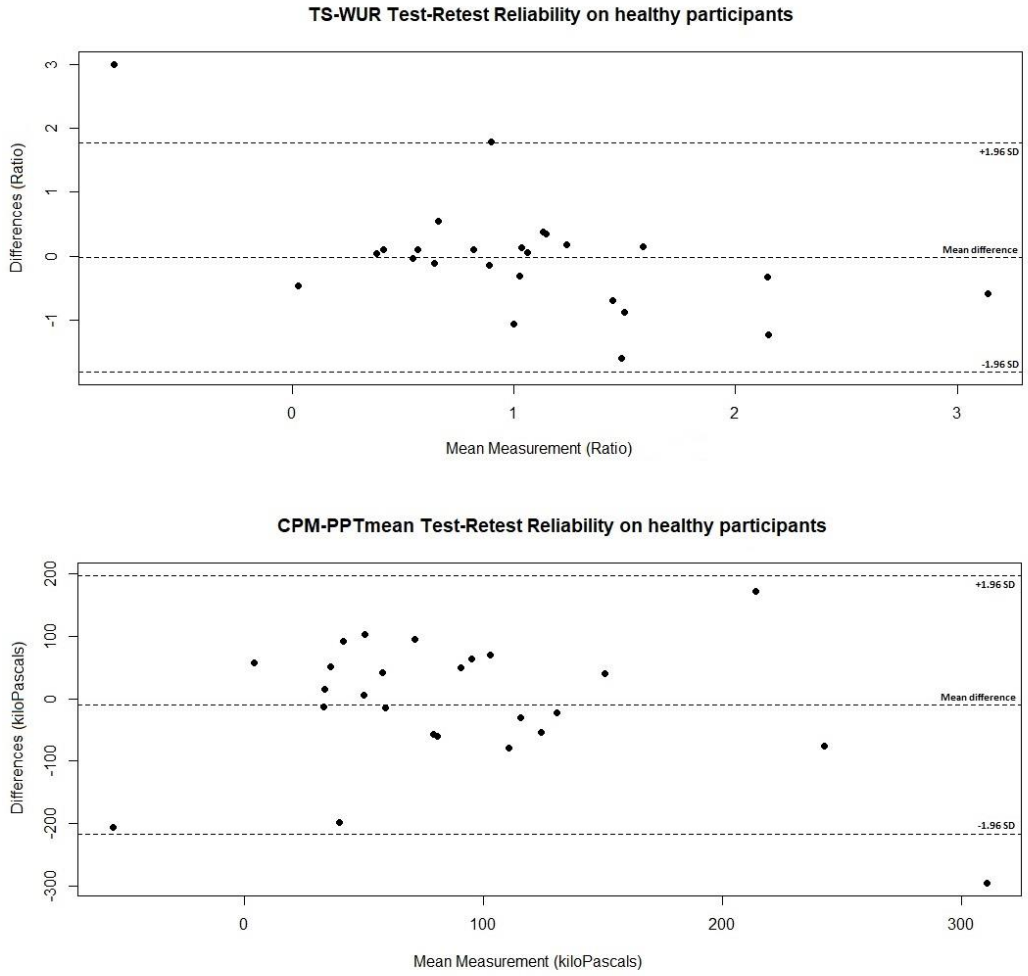

Figure 17. Bland-Altman plots for TS^{WUR} and $CPM^{PPT\text{-}mean}$ after measurements on healthy participants at baseline and follow-up.

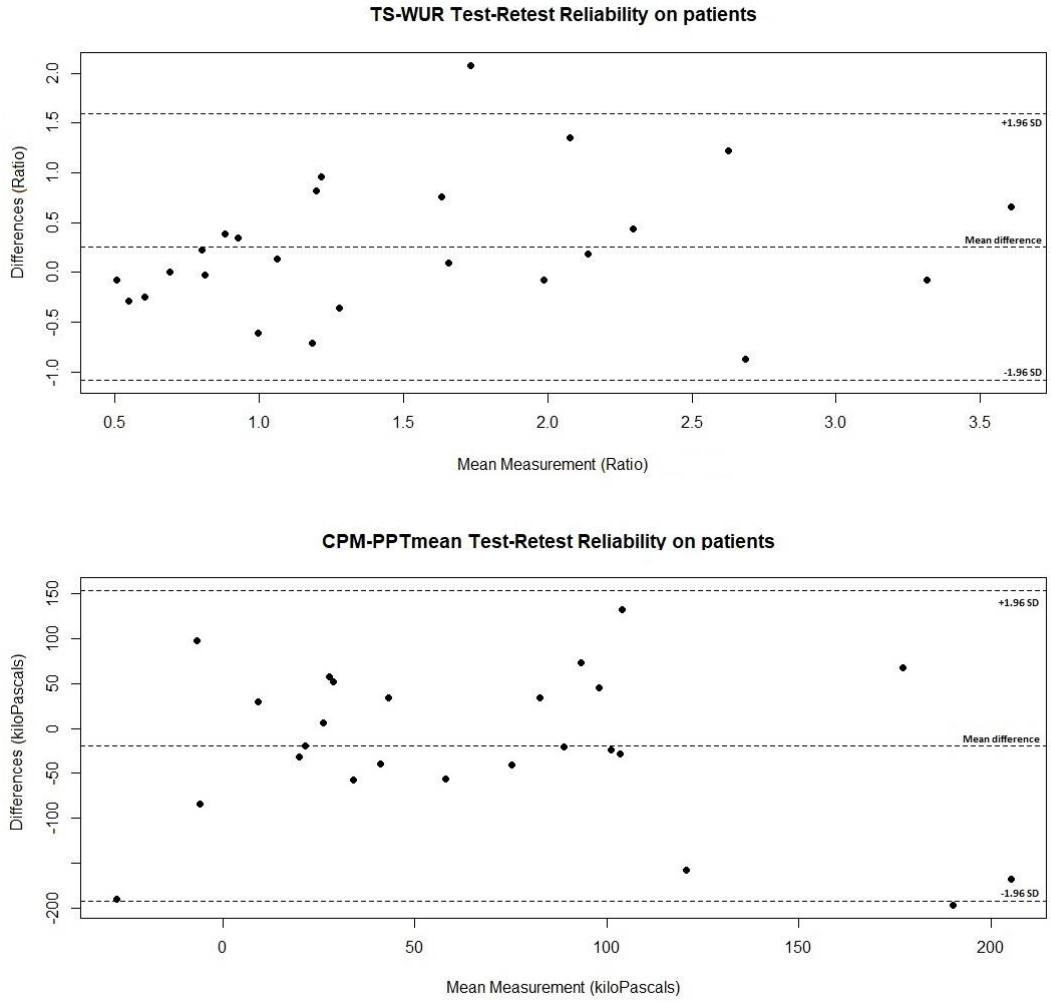

Figure 18. Bland-Altman plots for TS^{WUR} and $CPM^{PPT\text{-}mean}$ after measurements on patients with LBP at baseline and follow-up.

Test-retest Bland-Altman plots and corresponding ±LoA (95%CIs) for the combined population of participants (n=50) demonstrate no significant differences to the plots presented in **Figure 15** and **Figure 16** (mean differences: PPT=0.03, TS^{WUD}=0.17, TS^{WUR}=0.11, CPM^{Unc}=1.35 and CPM^{PPT}= -14.66) indicating no signs of measurement bias (**Figure 19** and **Figure 20**).

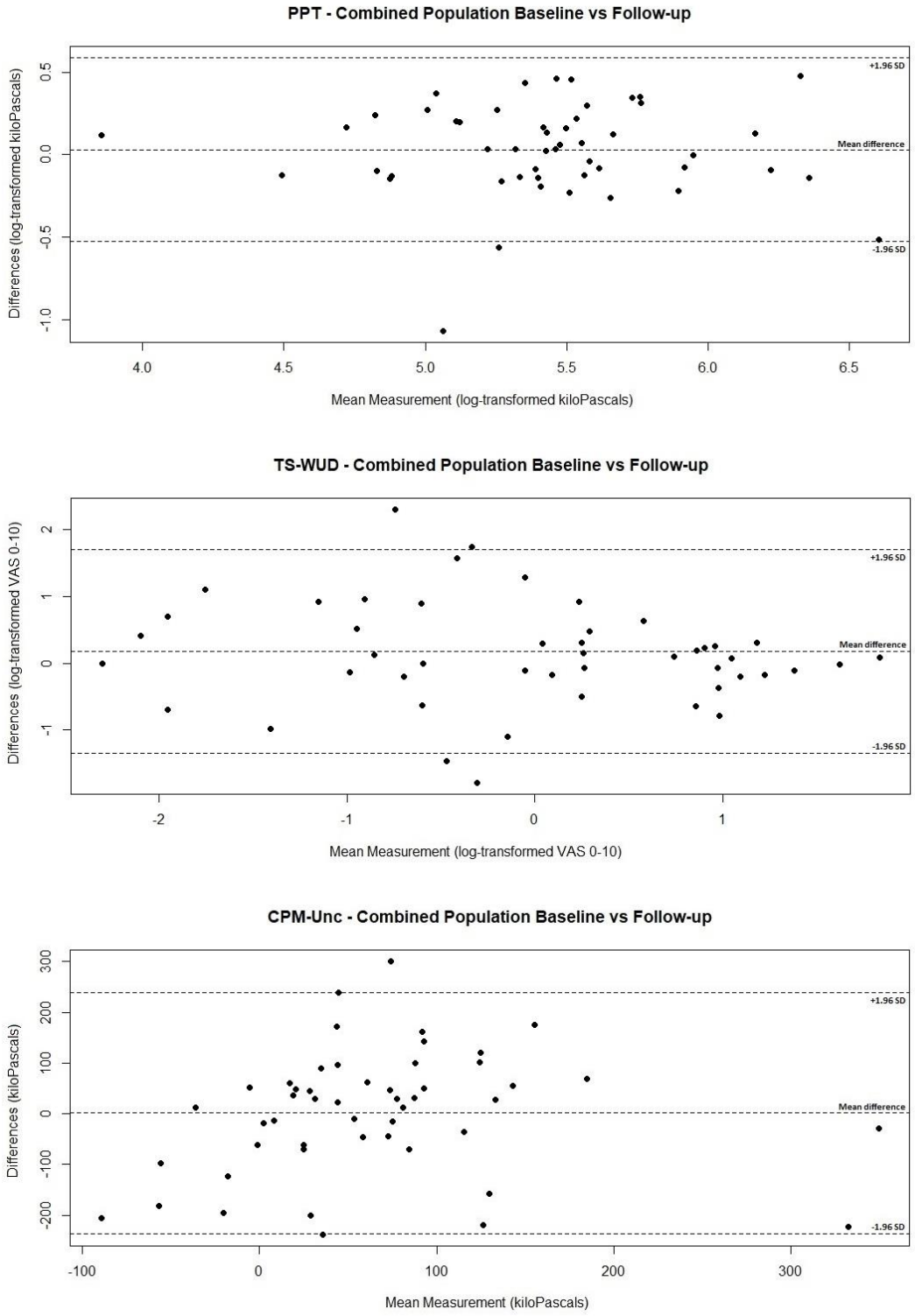

Figure 19. Bland-Altman plots for PPT, TS^{WUD} and CPM^{Unc} after measurements on the combined population at baseline and follow-up.

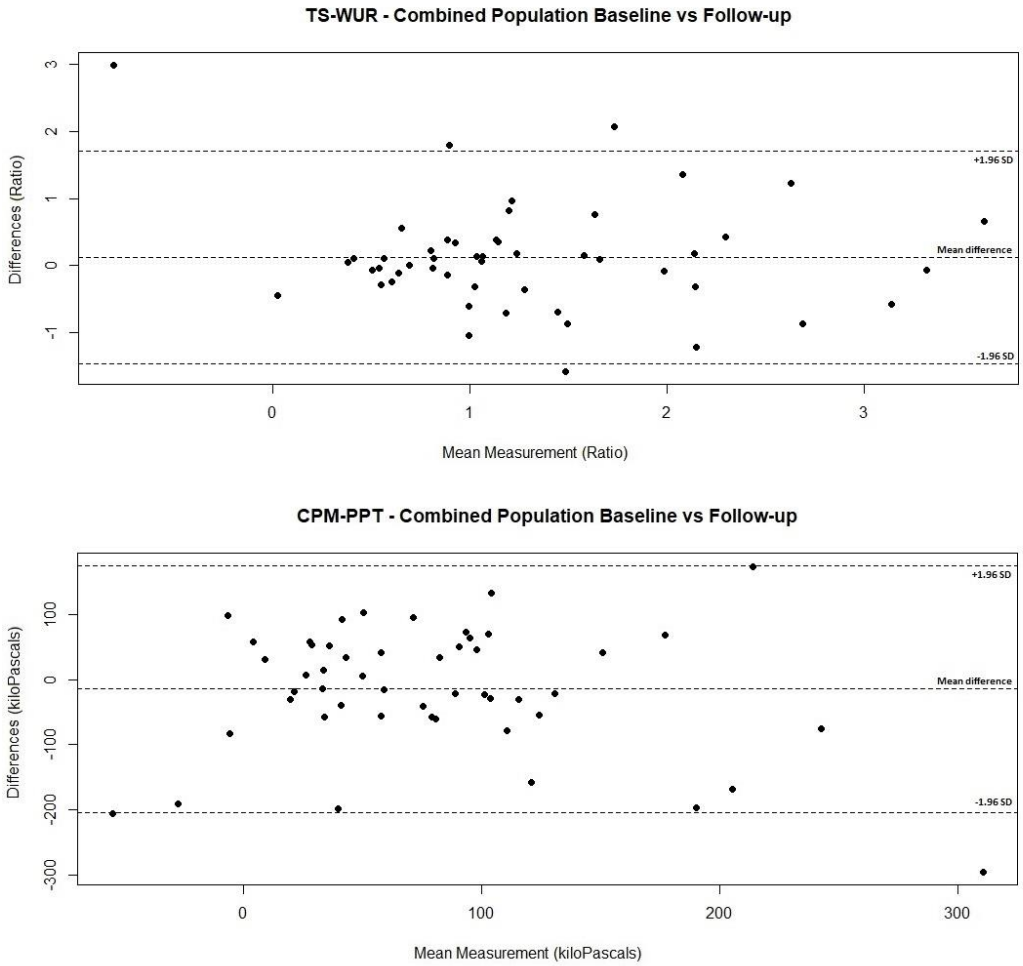

Figure 20. Bland-Altman plots for TS^{WUR} and $CPM^{PPT\text{-}mean}$ after measurements on the combined population at baseline and follow-up.

Inter-rater reliability ICCs for PPT were 0.75, for TS^{WUD} 0.88, TS^{WUR} 0.71, for CPM^{Unc} 0.55 and for CPM^{PPT} 0.35. Inter-rater CCCs ranged from 0.78 to 0.93 ($p \leq 0.01$) confirming a strong association between the measurements conducted by the two raters. Paired t-tests indicated that there were no significant differences ($p>0.05$) between the measurements of the two raters on the healthy population for all modalities with the exception of TS^{WUR}. T-tests for TS^{WUR} revealed significant differences ($t = -2.67$, $p=0.01$) between the measurements of the two raters which was not demonstrated in TS^{WUD}.

Inter-rater Bland-Altman plots and corresponding ±LoA (95%CIs) for each modality are shown in **Figure 21** and display no significant differences from zero.

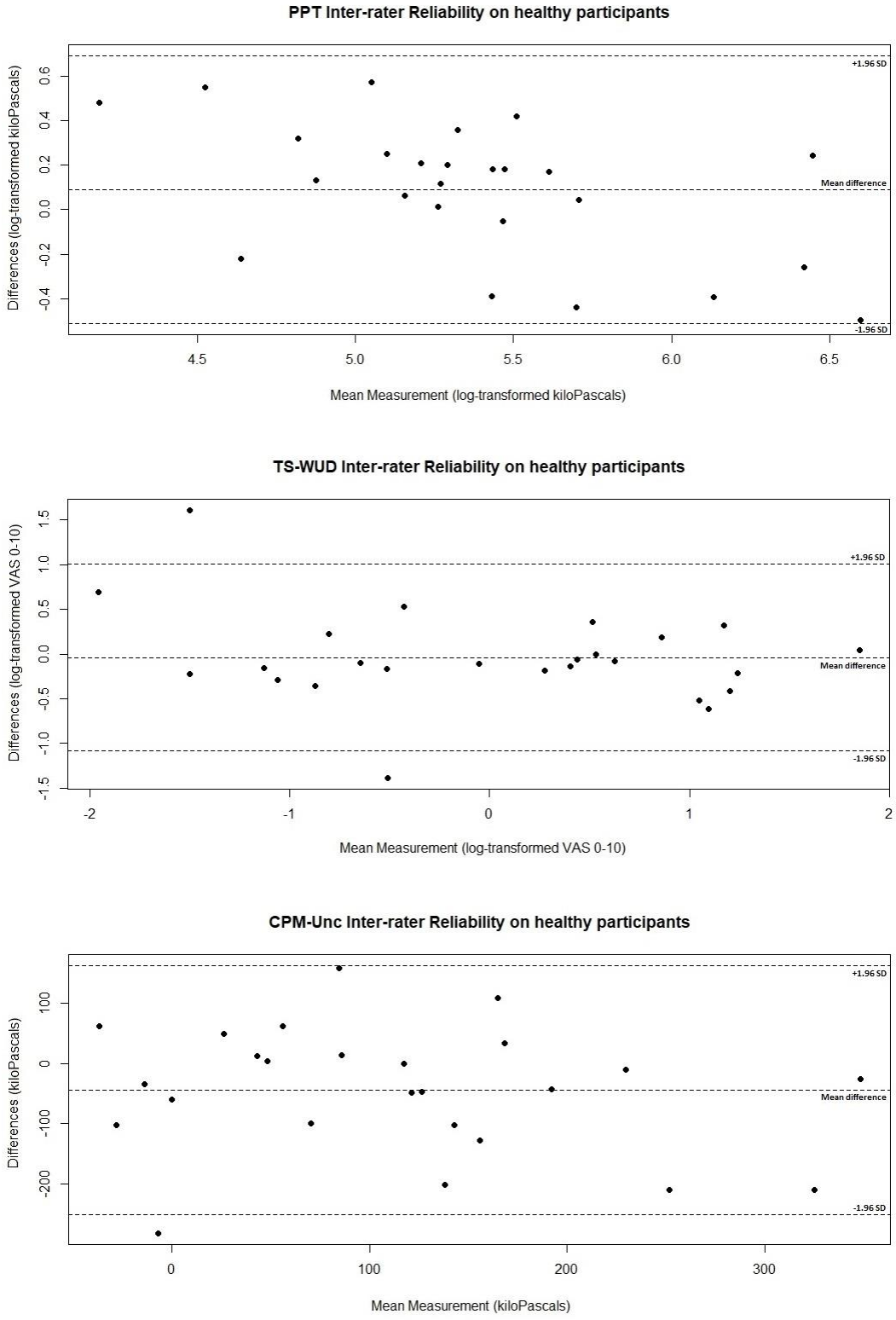

Figure 21. Bland-Altman plots for PPT, TS^{WUD} and CPM^{Unc} after measurements from both raters on healthy participants at baseline.

PPT demonstrated mean difference of 0.09 (-LoA: -0.51, 95%CI: -0.73 to -0.29; +LoA: 0.69, 95%CI: 0.47 to 0.91), TS^{WUD} of -0.04 (-LoA: -1.09, 95%CI: -1.47 to -0.70; +LoA: 1.00, 95%CI: 0.62 to 1.38) and CPM^{Unc} of -44.34 (-LoA: -251.03, 95%CI: -326.42 to -175.64; +LoA: 162.34, 95%CI: 86.95 to 237.74). No signs of bias along with similar mean differences and LoA can be observed in Bland-Altman plots for TS^{WUR} and $CPM^{PPT\text{-}mean}$ measurements of both raters (mean differences: TS^{WUR}=-0.24, CPM^{PPT}= -19.20) (**Figure 22**).

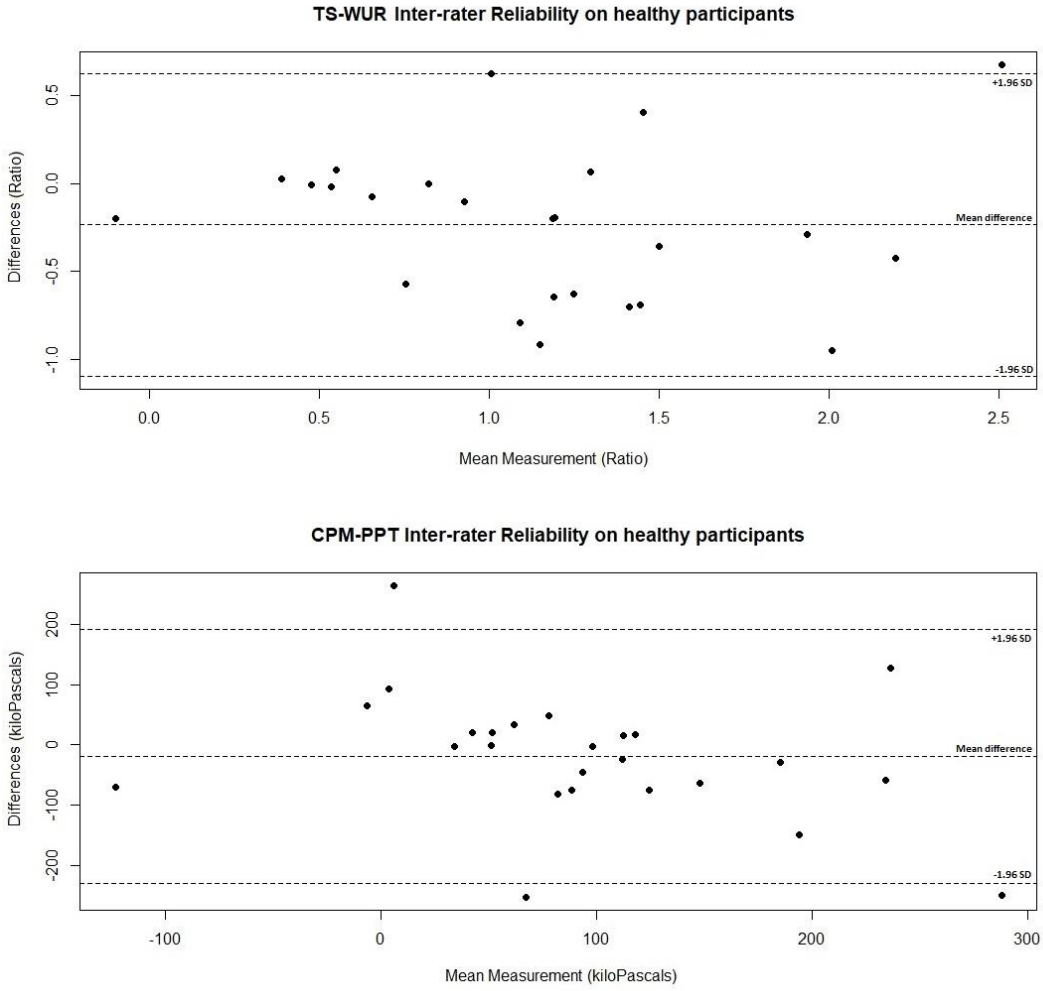

Figure 22. Bland-Altman plots for TS^{WUR} and $CPM^{PPT\text{-}mean}$ after measurements from both raters on healthy participants at baseline.

4.4.4. Reliability of using PPT means for CPM calculations

Analyses for degrees of association between the PPT mean values and the unconditioned PPT stimulus delivered for the purposes of CPM yielded good to excellent correlation-coefficients (r) and high to very high intraclass correlation-coefficients (ICC). Across populations and sessions, correlation values ranged from 0.79 to 0.89 ($p<0.01$) and ICCs from 0.84 (95%CI: 0.69 to 0.93) to 0.92 (95%CI: 0.83 to 0.96) for Rater 1 whereas correlation was 0.64 (<0.01) and ICC was 0.71 (95%CI: 0.45 to 0.86) for Rater 2 on the healthy population. Similarly, in the combined population analysis (n=50), correlation values were 0.84 ($p<0.01$) at baseline and 0.92 ($p<0.01$) at follow-up. ICCs followed the same pattern with values of 0.84 (95%CI: 0.74 to 0.91) at baseline and 0.91 (95%CI: 0.84 to 0.95) at follow-up. Paired t-tests revealed no significant differences between the two measurements both in the original sample format (n=25) and in the combined (n=50).

Median values and median difference of the different PPT measurements at each time-point for each rater as well as results from Kruskal–Wallis one-way analysis of variance and Pairwise Wilcoxon comparisons are presented in **Table 14**. Overall, Kruskal–Wallis analysis demonstrated significant heterogeneity between the 3 PPT measures (PPT^{MEAN}, PPT^{Unc} and PPT^{Con}) in the healthy population at baseline ($H=11.02$, $p<0.01$) and follow-up ($H=9.41$, $p<0.01$) for Rater 1 and at baseline ($H=7.57$, $p=0.02$) for Rater 2 but no significant heterogeneity for Rater 1 in the patient sample. Pairwise Wilcoxon comparisons revealed that PPT^{MEAN} was not significantly different from PPT^{Unc} ($p>0.05$) across populations, raters and timepoints, not different from PPT^{Con} in the patient population ($p>0.05$) but significantly different from PPT^{Con} on the healthy population at both time-points ($p=0.02$). PPT^{Unc} and PPT^{Con} were significantly different in the healthy population at baseline for both raters ($p=0.03$) but not significantly different at follow-up for healthy participants ($p>0.05$) and across timepoints for the patient population ($p>0.05$).

Table 14. PPT measurements descriptives accompanied by Kruskal-Wallis analysis of variance and pairwise Wilcoxon comparisons

Sample Size n=25 and n=50	Median values			Median difference			Kruskal-Wallis		Pairwise Wilcoxon Comparisons (Bonferroni corrected)		
	†PPT^{MEAN} (kPa)	PPT^{Unc} (kPa)	PPT^{Con} (kPa)	†PPT^{MEAN} vs PPT^{Unc} (kPa)	†PPT^{MEAN} vs PPT^{Con} (kPa)	PPT^{Unc} vs PPT^{Con} (kPa)	†PPT^{MEAN} with PPT^{Con} and PPT^{Unc}		†PPT^{MEAN} vs PPT^{Unc}	†PPT^{MEAN} vs PPT^{Con}	PPT^{Unc} vs PPT^{Con}
	Median (IQ Range)	Median (IQ Range)	Median (IQ Range)	Median (IQ Range)	Median (IQ Range)	Median (IQ Range)	H	p-value	Adj. p-value	Adj. p-value	Adj. p-value
Rater 1 (Healthy)											
Baseline	222.0 (176.9 to 249.5)	202.9 (142.1 to 290.1)	311.6 (247.9 to 393.0)	9.8 (-23.9 to 51.3)	-87.2 (-119.9 to -50.4)	-92.1 (-163.6 to -37.2)	9.80	<0.01	1.0	0.02	0.03
Follow-up	224.0 (178.4 to 251.9)	239.1 (200.9 to 294.0)	305.8 (244.0 to 337.1)	-14.0 (-52.2 to 8.2)	-66.6 (-131.03 to -36.9)	-55.9 (-95.0 to -5.9)	8.41	0.01	1.0	0.02	0.09
Rater 1 (Patients)											
Baseline	271.5 (195.5 to 305.3)	253.8 (204.8 to 355.7)	306.7 (197.0 to 402.8)	2.0 (-43.4 to 42.8)	-55.2 (-91.8 to -24.2)	-47.0 (-98.0 to 6.9)	2.15	0.34	1.0	0.39	1.0
Follow-up	216.5 (164.6 to 281.6)	275.4 (154.8 to 350.8)	312.6 (214.6 to 373.4)	-22.2 (-52.3 to 9.5)	-62.7 -99.3 to -31.0)	-38.2 (-81.4 to -11.8)	3.82	0.14	1.0	0.17	0.66
Rater 1 (Combined)											
Baseline	244.5 (185.2 to 306.7)	230.3 (155.6 to 353.8)	311.2 (247.3 to 400.4)	5.9 (-38.8 to 46.7)	-66.5 (-118.2 to -35.0)	-53.2 (-135.3 to -5.65)	8.42	0.01	1.0	0.02	0.08
Follow-up	228.2 (172.8 to 271.8)	243.5 (179.6 to 314.2)	308.7 (241.8 to 368.0)	-16.5 (-52.3 to 9.4)	-65.6 (-125.5 to -32.1)	-52.9 (-82.1 to -11.2)	10.97	<0.01	1.0	<0.01	0.07
Rater 2 (Healthy)											
Baseline	206.3 (147.0 to 275.4)	188.2 (114.7 to 224.4)	328.3 (192.1 to 448.8)	7.8 (-7.9 to 42.8)	-112.8 (-157.2 to -44.4)	-120.5 (-213.6 to -30.3)	7.82	0.02	1.0	0.11	0.03

PPT^{MEAN}: Use of all three initial measurements to calculate Pain Pressure Detection Threshold, **PPT^{Unc}:** Pain Pressure Detection Threshold measurement undertaken before the application of a conditioned stimulus, **PPT^{Con}:** Pain Pressure Detection Threshold undertaken under a simultaneous application of a conditioning stimulus

†PPT^{MEAN} is taken to be the reference PPT measurement

4.4.5. Internal validity

No statistically significant correlations were observed between modalities ($p>0.05$) for the healthy participants across raters (Rater 1/Rater 2) and time-points (baseline/follow-up). In the patient population, out of the three modalities, PPT negatively correlated with TS^{WUD} at baseline ($r=-0.43$, $p=0.03$) and at follow-up ($r=-0.52$, $p=0.01$) but did not significantly correlate with TS^{WUR} at the same timepoints. Analysis conducted in the combined population sample ($n=50$), yielded a statistically significant correlation between only PPT and TS^{WUD} ($r=-0.42$, $p<0.01$) only at follow-up. No significant correlations ($p>0.05$) were identified between PPT and $CPM^{PPT\text{-}mean}$ and between TS^{WUD} and $CPM^{PPT\text{-}mean}$ across raters, populations, population sizes and time-points. Correlation values and corresponding p-values are presented in **Table 15**.

4.4.6. External validity

External validity analyses were conducted only on the patient population as data from self-reported pain outcomes (NRS, PainDETECT, Body Manikin) were available only from those individuals. No statistically significant correlations ($p>0.05$) were observed between QST modalities (of any calculation type) and self-reported numerical pain outcomes (NRS, painDETECT). Correlation values for external validity are presented in **Table 15**.

Table 15. Correlation matrix between QST modalities as well as between QST modalities and indices of pain and central sensitisation

		Sample Size n=25	PPT		TS		†CPM	
			Cor	p-value	Cor	p-value	Cor	p-value
			Baseline					
Internal Validity	**Rater 1** (Healthy)	**PPT**	-	-	-	-	-	-
		TS	-0.05	0.79	-	-	-	-
		†CPM	-0.002	0.99	-0.05	0.79	-	-
			Follow-up					
		PPT	-	-	-	-	-	-
		TS	-0.31	0.14	-	-	-	-
		†CPM	0.19	0.36	-0.03	0.90	-	-
	Rater 1 (Patients)		**Baseline**					
		PPT	-	-	-	-	-	-
		TS	-0.42	0.03	-	-	-	-
		†CPM	0.26	0.21	-0.03	0.90	-	-
			Follow-up					
		PPT	-	-	-	-	-	-
		TS	-0.53	0.01	-	-	-	-
		†CPM	0.26	0.22	-0.19	0.38	-	-
	Rater 2 (Healthy)		**Baseline**					
		PPT	-	-	-	-	-	-
		TS	-0.04	0.86	-	-	-	-
		†CPM	0.39	0.05	-0.28	0.18	-	-

		Sample Size n=50	PPT		TS		†CPM	
			Cor	p-value	Cor	p-value	Cor	p-value
			Baseline					
Internal Validity	**Rater 1** (Combined)	**PPT**	-	-	-	-	-	-
		TS	-0.26	0.07	-	-	-	-
		†CPM	-0.11	0.45	0.19	0.18	-	-
			Follow-up					
		PPT	-	-	-	-	-	-
		TS	-0.42	<0.01	-	-	-	-
		†CPM	-0.15	0.31	-0.14	0.35	-	-

		Sample Size n=25	PPT		TS		†CPM	
			Cor	p-value	Cor	p-value	Cor	p-value
			Baseline					
External Validity	**Rater 1** (Patients)	**NRS**	0.27	0.20	0.08	0.71	0.13	0.53
		painDETECT	-0.20	0.35	0.13	0.54	-0.18	0.40
		ACR	-0.32	0.12	-0.14	0.51	-0.17	0.40
		7 Sites	-0.27	0.20	-0.19	0.36	-0.23	0.28
		24 Sites	-0.12	0.57	-0.17	0.43	-0.32	0.12

7 Sites: 7-site body manikin classification method, **24 Sites:** 24-site body manikin classification method, **ACR:** American College of Rheumatology, **CPM:** Conditioned Pain Modulation, **PPT:** Pain Pressure Detection Threshold, **TS:** Temporal Summation, **Cor:** Spearman's Rank Order Correlation

†$CPM^{PPT\text{-}mean}$ is taken to be the reference method of CPM calculation

4.5. Discussion

This study attempted to establish the reliability and validity of a QST protocol. It demonstrates that PPT and TS are reliable and valid tools to measure discrete aspects of central pain processing in healthy participants and patients alike. This is demonstrated by a single rater across both populations as well as between two raters on the healthy population. None of the QST modalities displayed a significant correlation with pain outcomes or pain distribution indicators. CPM demonstrated only acceptable reliability across healthy participants for both raters but it appeared to be less reliable on patients. The results of this study suggest that a combination of static and dynamic QST modalities within a single protocol is a reliable and valid way to quantify sensory function in healthy participants and pain hypersensitivity in individuals with chronic low back pain.

The present study demonstrated that PPT is a reliable approach across populations, time-points, and raters. This study extends previous studies of a similar research question that have demonstrated high reliability (ICC = 0.75 – 0.94) of PPT in healthy participants (Park et al., 2011, Chesterton et al., 2007, Nussbaum and Downes, 1998, Fabio Antonaci, 1998, Chung et al., 1992), in different patient populations (Geber et al., 2011, Wylde et al., 2011b) and in individuals with LBP (Paungmali et al., 2012). Apart from the high reliability, PPT results of this study demonstrate little variability as their ICC values were consistently high. This finding is in agreement with past findings where PPT reliability was consistently higher than other QST modalities when used in healthy and patient participants (Wylde et al., 2011b). The hight test-retest and inter-rater reliability of PPT could be also attributed to the 1-week period between sessions, which tends to reduce the variability of PPT measurements (Suokas et al., 2012).

Temporal summation test-retest and inter-rater reliability were also found to be high in this study, which was demonstrated across raters, populations, and time-points. Findings of this study are consistent with past research on healthy participants that demonstrated moderate to high TS test-retest reliability (ICC = 0.67 – 0.87) (Graven-Nielsen et al., 2015, Cathcart et al., 2009, Kong et al., 2013) but different to the findings of research on patients (Geber et al., 2011, Pigg et al., 2010) where poor test-retest (ICC = 0.43) and inter-rater reliability (ICC = 0.41) were demonstrated.

Test-retest and inter-rater reliability for CPM were found to marginally exceed the acceptability cut-off point (0.50) when calculated with an interim PPT stimulus as an

unconditioned stimulus in the healthy population. Such values are consistent with past literature on healthy participants where moderate test-retest reliability (ICC = 0.57) was demonstrated (Cathcart et al., 2009). When CPM was calculated with an unconditioned stimulus in the patient population, the test-retest reliability was negative (ICC = -0.10), indicating very little correlation between the two time-points, which is similar to past test-retest findings (ICC = -0.40) with ischaemic pain as the conditioning stimulus (Lewis et al., 2012a). However, when CPM was calculated with the mean PPT value as a conditioning stimulus, intraclass-correlation coefficients between time-points were low (ICC = 0.43 – 0.44) across raters, study populations and time-points. Such findings are marginally different from findings of past research conducted on patients with CLBP (ICC = 0.59) (Martel et al., 2013) or shoulder pain (ICC = 0.54) (Valencia et al., 2014) and better than findings on participants suffering from chronic pancreatitis (ICC = 0.10) (Olesen et al., 2012).

This study demonstrated that TS is more reliable and valid if calculated as a difference (TS^{WUD}) rather than as a ratio (TS^{WUR}), in contrast to the recommendations of the German Research Network on Neuropathic Pain (DFNS) (Rolke et al., 2006a, Rolke et al., 2006b). Conceptually, ratio calculations describe the excitability of spinal cord neurons as it reaches a plateau after frequent stimulation (Rolke et al., 2006a) and can be easily utilised in routine clinical assessment (Rolke et al., 2006b). However, TS^{WUD} was superior to WUR with respect to ICC and CCC suggesting a statistical and methodological advantage over TS^{WUR}. In terms of CPM, a consistently strong correlation (ICCs and correlation-coefficients) between the mean PPT value and the interim unconditioned PPT stimulus across populations and time-points was observed. It could be argued that the mean PPT value at the beginning of the protocol can operate as an unconditioned stimulus in its own accord. Forfeiting the interim PPT stimulus could minimise the chances of moderated patient response to stimuli by increased or intolerable pain. Similar to TS, $CPM^{PPT\text{-}mean}$ demonstrates statistical, methodological, and application advantages over CPM^{Unc}.

Findings from PPT and TS in this study are consistent with existing evidence that consider these modalities sensitive and reliable to assess CS in patient populations in general and CLBP in particular (Arendt-Nielsen et al., 2018). In terms of CPM, consistent research findings (O'Brien et al., 2018, Kennedy et al., 2016, Lewis et al., 2012b, Yarnitsky, 2010) suggest that deficient CPM is a characteristic of populations with chronic pain although, whether that is due to a fully activated endogenous inhibition or because of a reduced ability to modulate pain is still unclear. Nevertheless, obtaining CPM reliability is an elusive concept (Kennedy et al., 2016)

and establishing the association between CPM responses and clinical manifestations of pain merits further investigation (Fernandes et al., 2019). A large amount of the patient population (80%) indicated pain without ball-gripping, which could have influenced patient response to stimuli by bringing the pain threshold closer to tolerance levels. Occurrences where the painfulness of stimulation just becomes intolerable have been associated with poor CPM test-retest reliability (Olesen et al., 2012). Another explanation for the poor reliability values, particularly in CPM, could be the lack of true variance between PPT^{MEAN}, PPT^{Unc} and PPT^{Con} that was observed across timepoints. The CPM effect on healthy subjects might also unpredictably influence the test-retest reliability because of the absence of systematic variance induced by the prevalence of a pathology. Repeatability calculations of a non-existent measure in the cases of both the healthy and the patient participants are likely to be poor. Nevertheless, the CPM effect identified in this study (PPT^{Con} - PPT^{Unc}) in healthy and patient participants is along the lines of the CPM effect reported in past studies with CLBP populations and healthy controls (Corrêa et al., 2015, Owens et al., 2015). Generally, pain is a biopsychosocial phenomenon and can be a notoriously difficult construct to measure. Therefore, aiming to measure pain processing on two different occasions through identical pain indices would be unlikely and remarkably elusive, due to the variability and fluctuating nature often observed in musculoskeletal pain (Gooberman-Hill et al., 2007). In general, QST modalities are likely to produce variable results because of the subjective nature of perception (Wylde et al., 2011b). That is particularly prevalent in CPM where participants are asked to subjectively indicate when they feel pain on two separate occasions.

The lack of significant correlation between CPM and the other QST modalities might be due to low reliability in CPM highlighted above. Significant associations were demonstrated between the 2 measures (PPT and TS^{WUD}) with the highest reliability, despite the low sample size. A larger sample size could potentially demonstrate weak associations as significant. However, the weakness of associations between QST modalities suggests that they represent different aspects of pain processing. Central sensitisation results from multiple processes, and different QST modalities might reflect different aspects of central sensitisation, rather than each being estimates of a uniform `central sensitisation' phenomenon. Patients with LBP have been found to have reduced PPTs, increased TS and deficient CPM (Müller et al., 2019, Marcuzzi et al., 2018). Despite the theoretical and operational similarities between different modalities, their association with each other is often not reported. Also, PPT

correlated significantly only with TS^{WUD} and not with TS^{WUR}, which highlights further the increased validity of that calculation type.

None of the QST modalities correlated significantly with any of the self-reported pain outcomes in the patient population. This is consistent with current evidence highlighting the lack of meaningful cross-sectional correlations between pain threshold modalities and pain in individuals with spinal pain (Hübscher et al., 2013) or between CPM and chronic pain (Fernandes et al., 2019). That could be because QST and self-reported pain outcomes measure entirely different constructs or that central sensitisation might not be implicated in the way people perceive, experience and report pain. Also, the self-reported outcome measures had already been gathered from 50% of the patients prior to their baseline reliability assessment as part of their participation to an observational study. Participants could have responded differently if the measurement was on the same day, which could lead to statistically significant correlations. A recent systematic review (Georgopoulos et al., 2019) found significant correlations between baseline dynamic modalities and follow-up pain although correlations were not sufficiently strong to be able to recommend any current QST modality as a stand-alone pain sensitisation identification or prediction tool.

The number of painful sites on a body manikin is used as a measure of widespread pain and ongoing research has added to its validity (construct, content and face validity) over the years (Wolfe et al., 2019, Wolfe et al., 2016, Wolfe et al., 2011). Since central sensitisation is considered a mechanism that drives widespread pain (Arendt-Nielsen et al., 2018) and pain at a distal site to the pathology is also considered to be a feature of central sensitisation (Woolf, 2011), an association between the number of painful sites and QST at a distal site would be anticipated. A recent study demonstrated that PPT at a distal site is associated with the number of painful sites as well as with the shaded distribution of the pain on a body manikin (Akin-Akinyosoye et al., 2018). However, no significant associations between QST and the number of painful sites were found despite the different classification methods available in this study. The relatively small sample size could have influenced that relationship as well or the number of painful sites might simply be a marker of pain density and clinical severity, encapsulating only some of the processes that drive CS.

One of the strengths of the study is that it provides evidence regarding internal and external validity of the QST protocol which are not often or consistently reported in literature. Due to the study set-up there was risk of systemic bias, which was not realised as analyses did not reveal any significant differences between raters. Another

strength was that the delivered protocol is easy to implement in clinical practice through the utilisation of low-cost alternatives. QST could be used in clinic to confirm the involvement of central mechanisms as initially indicated by self-reported outcomes or derived after thorough clinical examination. Confirmation of diagnosis through QST could facilitate referral of patients to appropriate clinicians such as consultant rheumatologists, anaesthetists or physiotherapists, which could enhance patient satisfaction and improve health-care utilisation by minimising unnecessary appointments and enhancing health services navigation. Essential training is required for the successful utilisation of QST in clinical practice as lack of reliability among raters could negatively influence diagnosis. The relatively small sample size might have been an important limitation. Even though it was determined to provide robust test-retest and inter-rater reliability results, it might have been small for the post-hoc analyses carried out for the purposes of external validity and therefore such findings should be viewed with caution.

4.6. Conclusion

A QST protocol consist of PPT (mean of three repeated measurements), TS (calculated as the difference between a single and the average of ten repetitive stimuli), and CPM (calculated as the difference between a conditioned PPT measurement and the mean of the three repeated PPT measurements conducted earlier) is reliable for the assessment of pain on healthy participants and patients with CLBP. PPT and TS^{WUD} demonstrated high test-retest and inter-rater reliability and moderate internal validity. Ratio calculation for TS and the use of an interim unconditioned stimulus for the purposes of CPM calculation appear to be less sensitive for reliability and validity analyses. Forfeiting an operationally similar stimulus from CPM calculations could reduce the examination burden for participants, which could lead to more accurate findings. PPT and TS correlated with each other on the patient population but none of the modalities correlated with self-reported pain intensity or self-reported pain distribution. QST protocols combining ‘static’ and ‘dynamic’ modalities appear to be reliable enough to provide inference about pain hypersensitivity. Further research is needed on larger samples and different populations to confirm the findings in this study.

5. CROSS-SECTIONAL ASSOCIATIONS BETWEEN CENTRAL SENSITISATION INDICES SELF-MANAGEMENT/SELF-CARE OUTCOMES INDIVIDUALS WITH CLBP

5.1. Introduction

Different self-management approaches have demonstrated small to moderate short and long-term effectiveness on pain and disability in individuals with OA (Du et al., 2011) and CLBP (Du et al., 2017). Nevertheless, measuring the effectiveness of strategies that aim to increase the self-management capacity of patients with chronic conditions requires the utilisation of outcome measures with a multidimensional focus that simple pain and disability questionnaires might not possess (Nolte and Osborne, 2013). Self-efficacy, depression, catastrophisation and, to a lesser degree, physical function appear to be linked with self-management outcomes rather than pain and disability levels (Miles et al., 2011). Multi-construct tools measuring physical, social and psychological domains should be prioritised in studies featuring self-management interventions as they have been found to be more consistent and more valid than single-construct scales in the measurement of self-management outcomes (Banerjee et al., 2018).

The overall pathophysiology of CLBP implicates mechanisms in both peripheral and central nervous systems. Constant or repetitive nociceptive input can cause neuroplastic changes in the central nervous system, which can facilitate the CS (Woolf and Salter, 2000). Evidence of CS has been found to be present in individuals with both acute (Vuilleumier et al., 2017) and CLBP (Smart et al., 2012b). CS is a multifactorial phenomenon, driven by a complex combination of physical, mental and emotional features and has been associated with increasing psychological distress, pain, fatigue and catastrophisation (Bourke et al., 2015, Woolf, 2011). Such factors influence the clinical outcomes of CLBP as well as other musculoskeletal conditions (Arendt-Nielsen et al., 2018).

Quantitative sensory testing can explore mechanisms responsible for the development or maintenance of local and widespread pain as well as provide indices of CS in populations with CLBP (Pavlaković and Petzke, 2010). QST can evaluate the functional status of the somatosensory system comprised by the peripheral nervous system and central pathways (Krumova et al., 2012, Hansson et al., 2007, Meier et

al., 2001). For that purpose, different methods such as pain threshold perception, temporal summation and descending pain modulation, have been utilised to assess different features of pain processing and facilitate the quantification of both loss and gain of sensory function (Arendt-Nielsen et al., 2018, Yarnitsky and Granot, 2006).

Pain distribution, as reported by patients on a human body manikin (Wolfe et al., 2019, Wylde et al., 2011a, Croft et al., 1996), and the combination of 8 discrete self-reported traits (anxiety, depression, catastrophising, neuropathic-like pain, fatigue, sleep disturbance, pain distribution and cognitive impact) found to be associated with CS into one latent 'Central Mechanisms' trait (Akin-Akinyosoye et al., 2018) have been also proposed as valid methods to identify CS and facilitate the classification of patients accordingly.

Increased levels of CS might contribute to the relatively poor efficacy of treatments aiming to facilitate self-management given the observable shared link with factors thought to be influencing SM. It is currently unknown whether CS is associated with worse self-management outcomes in people with CLBP following interventions aiming to improve such outcomes.

The hypothesis underlying this work is that CS indices negatively influence the ability of individuals with CLBP to effectively self-manage their condition.

5.2. Aims and Objectives

5.2.1. Aims

The main aim of this chapter was to ascertain whether there is an association between CS indices and self-management in a population with CLBP prior to participation in an intervention programme that aims to improve such outcomes.

5.2.2. Objectives

- To determine whether different CS indices associated specifically with self-management measures.
- To establish a cut off for number of body sites shaded on a self-reported pain manikin that best identifies those with widespread pain and explore whether certain self-reported items taken to indicate central mechanisms involvement contribute to a single latent trait in individuals with CLBP.

- To explore associations between different putative CS indices.
- To explore the relative importance of CS markers compared to other predictors of self-management.

5.3. Methods

The full methodological details of this study are given in the Methods chapter (**METHODS**) and only a brief outline is presented in this chapter whenever necessary.

5.3.1. Sample size considerations

Calculations determined that a sample size of 90-120 participants was enough to allow robust regression analyses including up to 7 variables within each regression model. Details regarding the sample size calculation methods for this study are given in section **2.5** of the methods chapter.

5.3.2. Study participants and recruitment

The study was approved by the East Midlands - Nottingham 1 Research Ethics Committee of the Health Research Authority, United Kingdom (REC: 18/EM/0049). Individuals suffering from CLBP that had been eligible to participate in a CBT-based physiotherapy (PT) or multidisciplinary (MDT) group intervention programme featuring a combination of neuroscience education, psychological support, relaxation techniques, pacing, exercise, medication and goal-setting were invited to participate in the study. Details on recruitment, inclusion/exclusion criteria, intervention and programme allocation criteria are given in **2.7.2**.

5.3.3. Clinical assessment and application methods

Individuals with CLBP were invited to participate in a session right before the start of their intervention (baseline). All participants had QST (PPT, TS and CPM) applied to the forearm of their dominant hand and were invited to complete a questionnaire booklet featuring self-reported outcome measures about self-management (heiQ), pain intensity (NRS), back pain severity (STarT-Back), pain distribution (body manikin), pain components (PD-Q), disability (RMDQ), emotional distress (HADS), maladaptive beliefs (PCS, TSK), fatigue (FSS), quality of life (EQ-5D-5L), fibromyalgia pain (FMSS) and health-care utilisation (HCUQ). Details about the hospital setting,

clinical assessment, QST application methods and self-reported outcome measures are given in **2.8.2.1**, **2.6.1**, **2.6.2.1**, **2.6.2.2** and **2.6.3**.

5.3.4. Data analysis and analytical procedures

Data normality exploration procedures are detailed in **2.9.2.1**. Group differences calculation methods are given in **2.9.2.2**. Details on QST calculation and interpretation methods are given in **2.9.2.3**. The derivation of the different body manikin classification methods that were used as CS indices are detailed in **2.10.3.1** whereas the procedure for the derivation of the Central Mechanisms trait is presented in **2.10.3.2**.

Correlation analyses were undertaken to explore whether the 10 discrete constructs of SM/SC (8 domains of heiQ, PSEQ and HCUQ) are correlated with the different CS indices, demographic details, and the other secondary factors. Correlation analyses were also undertaken to explore whether the discrete CS indices are correlated with each other as well as with outcome measures other than SM/SC. Extensive information regarding correlation analyses is given in **2.10.3.3**.

Regression modelling was undertaken to explore the association between SM/SC outcomes and the CS indices adjusted for other factors. The 10 discrete constructs of SM/SC were the dependent variables in distinct regression models, whereas CS indices (QST, pain distribution, the Central Mechanisms trait), age, sex, pain or disability, depression, catastrophising and fatigue were included in the models as independent variables at the same time. Information on dependent and independent variables for the purposes of regression modelling are given in **2.10.3.4.1** and **2.10.3.4.2**, whereas regression modelling procedures are fully detailed in **2.10.3.4.3**.

5.4. Results

5.4.1. Data management and transformation

Details about data distribution as well as the results of the log-transformation process can be found in **Appendix 7**. Not all primary and secondary outcomes demonstrated normal distribution upon testing. TS, HEIQ-SMI, HEIQ-STA, HEIQ-HSN, PSEQ, HCUQ, NRS, HADS-Dep., FSS, EQ-5D-5L and STarT-Back demonstrated distributions significantly different to normal after logarithmic transformation.

5.4.2. Demographic data and clinical characteristics

Total population demographic data and clinical features for QST and psychological variables as well as data separated according to intervention programme are given in **Table 16** and **Table 17**. Out of 177 eligible individuals with CLBP, 97 (71% females, mean age 56±13 years) agreed to participate whereas, 80 people declined to be included (70% females, mean age 54±14 years) without declaring a reason. From the 97 individuals who agreed to participate in the study, 92 were undertaking their treatment within a hospital setting and 5 within a community setting. QST assessments and self-reported outcome measures were taken on day 1 of their participation into a group intervention programme.

Overall, participants demonstrated moderate ability to self-manage in all heiQ domains (2.5-3.0/4.0) as well as moderate pain (6/10), depression (9/21), anxiety (9/21) and catastrophising (22/52). When the population was separated according to group intervention programme they are under, participants following the MDT programme demonstrated elevated levels of pain (7/10), depression (12/21) and anxiety (12/21) but moderate catastrophising (26-52) and self-management capacity on all heiQ domains (2.4-3.0/4.0). Participants following the PT programme showed moderate pain (5/10), mild depression (5-21), anxiety (7/21) and catastrophising (13/52) and above moderate capacity to self-manage in all heiQ domains (2.8-3.0/4.0). Participants of both programmes demonstrated similarities in terms of demographics whereas, apart from HEIQ-SMI, HEIQ-STA, HEIQ-HSN, HCUQ, TS and CPM, populations were found to be different in all other primary or secondary outcomes.

Table 16. Patient demographic characteristics, self-management/self-care, and pain sensitivity factors overall and according to intervention programmes at baseline.

Variables (Unit, Value, Range, %)	**All Participants** Mean (± SD), Median (IQ Range)	**PT** Mean (± SD), Median (IQ Range)	**MDT** Mean (± SD), Median (IQ Range)	**Group Difference** MWUT (p-value)
No. Participants	97	42	55	
Age (y)	56 (±13)	57 (±13)	55 (±14)	1200 (0.75)
BMI	29.4 (25.7 to 34.6)	28.7 (25.4 to 33.4)	31.6 (26.1 to 36.0)	937 (0.11)
Female (%)	71%	52%	75%	
Setting				
Hospital	92	38	54	
Community	5	4	1	
Primary Outcomes				
Health Education Impact Questionnaire Domains				
Health Directed Behaviour (1-4)	2.5 (2.3 to 3.0)	2.8 (2.4 to 3.0)	2.5 (2.0 to 2.9)	**1469 (0.02)**
Positive Engagement in Life (1-4)	2.6 (2.2 to 3.0)	2.9 (2.6 to 3.2)	2.4 (2.2 to 2.7)	**1813 (<0.01)**
Self-monitoring & Insight (1-4)	3.0 (2.8 to 3.2)	3.0 (2.8 to 3.2)	3.0 (2.8 to 3.1)	1217 (0.65)
Constructive Attitudes & Approaches (1-4)	2.8 (2.2 to 3.0)	3.0 (2.8 to 3.4)	2.6 (2.2 to 2.8)	**1791 (<0.01)**
Skill & Technique Acquisition (1-4)	2.8 (2.5 to 3.0)	2.8 (2.5 to 3.0)	2.8 (2.5 to 3.0)	1307 (0.26)
Social Integration and Support (1-4)	2.8 (2.4 to 3.0)	3.0 (2.6 to 3.2)	2.6 (2.3 to 3.0)	**1573 (<0.01)**
Health Services Navigation (1-4)	3.0 (2.6 to 3.2)	3.0 (2.8 to 3.4)	3.0 (2.6 to 3.0)	1288 (0.33)
Emotional Distress (1-4)	2.8 (2.3 to 3.2)	2.7 (2.0 to 3.4)	3.0 (2.7 to 3.0)	**629 (<0.01)**
Self-Care				
Pain Self-Efficacy Questionnaire (0-60)	27 (20 to 41)	41 (32 to 48)	21 (16 to 27)	**2006 (<0.01)**
Health Care Utilisation Questionnaire (Units)	4 (2 to 6)	3 (2 to 6)	4 (2 to 6)	1142 (0.92)
QST				
Pain Pressure Detection Threshold (kPa)	205.8 (148.2 to 297.6)	240.4 (175.3 to 293.9)	170.5 (110.4 to 249.8)	**1512 (0.01)**
Temporal Summation (0-10)	1.0 (0.4 to 2.8)	0.8 (0.4 to 2.4)	1.1 (0.5 to 3.1)	1020 (0.33)
Conditioned Pain Modulation (kPa)	59.1 (5.6 to 99.3)	66.3 (22.5 to 110.3)	46.7 (4.2 to 88.3)	1350 (0.16)

BMI: Body Mass Index, **IQ:** Interquartile, **kPa:** Kilopascals, **MDT:** Multidisciplinary Intervention Programme, **MWU:** Mann-Whitney U Test, **PT:** Physiotherapy Intervention Programme, **QST:** Quantitative Sensory Testing, **SD:** Standard Deviation

Table 17. Secondary outcome measures overall and according to intervention programmes at baseline.

Variables (Unit, Value, Range, %)	**All Participants** Mean (± SD), Median (IQ Range)	**PT** Mean (± SD), Median (IQ Range)	**MDT** Mean (± SD), Median (IQ Range)	**Group Difference** MWUT (p-value)
Pain Numerical Rating Scale (0-10)	6 (5 to 7)	5 (4 to 6)	7 (6 to 7)	**618 (<0.01)**
painDETECT (0-38)	17 (12 to 24)	15 (11 to 19)	21 (16 to 25)	**658 (<0.01)**
Now (0-10)	6 (4 to 7)	5 (3 to 6)	6 (6 to 7)	**508 (<0.01)**
Strongest (0-10)	8 (8 to 9)	8 (7 to 9)	9 (8 to 9)	**769 (<0.01)**
Average (0-10)	6 (6 to 7)	6 (5 to 7)	7 (6 to 8)	**635 (<0.01)**
Hospital *Anxiety* Scale (0-21)	9 (6 to 13)	7 (5 to 9)	12 (8 to 15)	**575 (<0.01)**
Hospital *Depression* Scale (0-21)	9 (5 to 12)	5 (3 to 9)	12 (8 to 13)	**481 (<0.01)**
Pain Catastrophising Scale (0-52)	22 (11 to 31)	13 (7 to 21)	26 (18 to 35)	**518 (<0.01)**
Tampa Scale of Kinesiophobia (17-68)	38 (33 to 43)	36 (31 to 42)	39 (35 to 44)	**867 (0.04)**
Roland-Morris Disability Questionnaire (0-24)	13 (9 to 18)	10 (5 to 12)	16 (6 to 20)	**374 (<0.01)**
Fatigue Severity Scale (7-63)	42 (29 to 52)	33 (26 to 45)	46 (39 to 53)	**732 (<0.01)**
Fatigue-Visual Analogue Scale (0-10)	5 (3 to 6)	5 (3 to 7)	5 (3 to 6)	**1313 (0.25)**
Fibromyalgia Severity Scale (0-31)	13 (8 to 18)	9 (6 to 14)	16 (12 to 20)	**460 (<0.01)**
EQ-5D-5L (Index)	0.45 (0.23 to 0.64)	0.61 (0.46 to 0.70)	0.32 (0.14 to 0.53)	**1859 (<0.01)**
EQ-Visual Analogue Scale (0-100)	50 (40 to 65)	61 (50 to 70)	40 (35 to 50)	**1814 (<0.01)**
STarT-Back (0-9)	6 (4 to 7)	4 (3 to 6)	6 (5 to 8)	**469 (<0.01)**

EQ-5D-5L: Quality of Life Instrument, **IQ:** Interquartile, **MDT:** Multidisciplinary Intervention Programme, **MWU:** Mann-Whitney U Test, **PT:** Physiotherapy Intervention Programme, **SD:** Standard Deviation, **STarT-Back:** Stratification toot

5.4.3. Pain distribution on a body manikin

Area Under the Curve values with corresponding 95%CIs, optimal cut-off points for number of painful sites, sensitivity and specificity values for predicting low PPT at brachioradialis, low CPM and high TS are presented in **Table 18**. Overall, AUC values for each modality and different widespread pain classification methods demonstrated low accuracy. PPT demonstrated the highest AUC values across all modalities and calculation methods that ranged from 0.62 to 0.67. The 24-site quantification method demonstrated AUC of 0.67 (95%CI: 0.55 to 0.80) (**Figure 23**), optimal number of painful sites ≥9, sensitivity of 64.0% and specificity of 69.4%, which is marginally below the operational cut-off percentage of 75% but the highest combination of sensitivity and specificity percentage across all modalities, calculation methods, and site quantification approaches.

Table 18. Cut-off points for the optimal number of painful sites needed to classify low PPT or CPM (1st quartile) and high TS (4th quartile) based on a-priori binary manikin classifications.

	PPT			
	AUC (95% CI)	Optimal cut-off point	Sensitivity (TPR)	Specificity (TNR)
7-site quantification				
Forearm	0.62 (0.49 – 0.75)	>=5	32.00%	93.05%
24-site quantification				
Forearm	0.67 (0.55 – 0.80)	>=9	64.00%	69.44%
	TS			
	AUC (95% CI)	Optimal cut-off point	Sensitivity (TPR)	Specificity (TNR)
7-site quantification				
Forearm	0.50 (0.37 – 0.63)	>=4	53.85%	50.70%
24-site quantification				
Forearm	0.59 (0.46 – 0.72)	>=12	38.46%	77.46%
	CPM			
	AUC (95% CI)	Optimal cut-off point	Sensitivity (TPR)	Specificity (TNR)
7-site quantification				
Forearm	0.55 (0.42 – 0.67)	>=4	56.00%	51.39%
24-site quantification				
Forearm	0.47 (0.34 – 0.60)	>=6	68.00%	36.11%

AUC: Area Under the Curve, **CPM:** Conditioned Pain Modulation, **PPT:** Pain Pressure Detection Threshold, **TNR:** True Negative Rate, **TPR:** True Positive Rate, **TS:** Temporal Summation

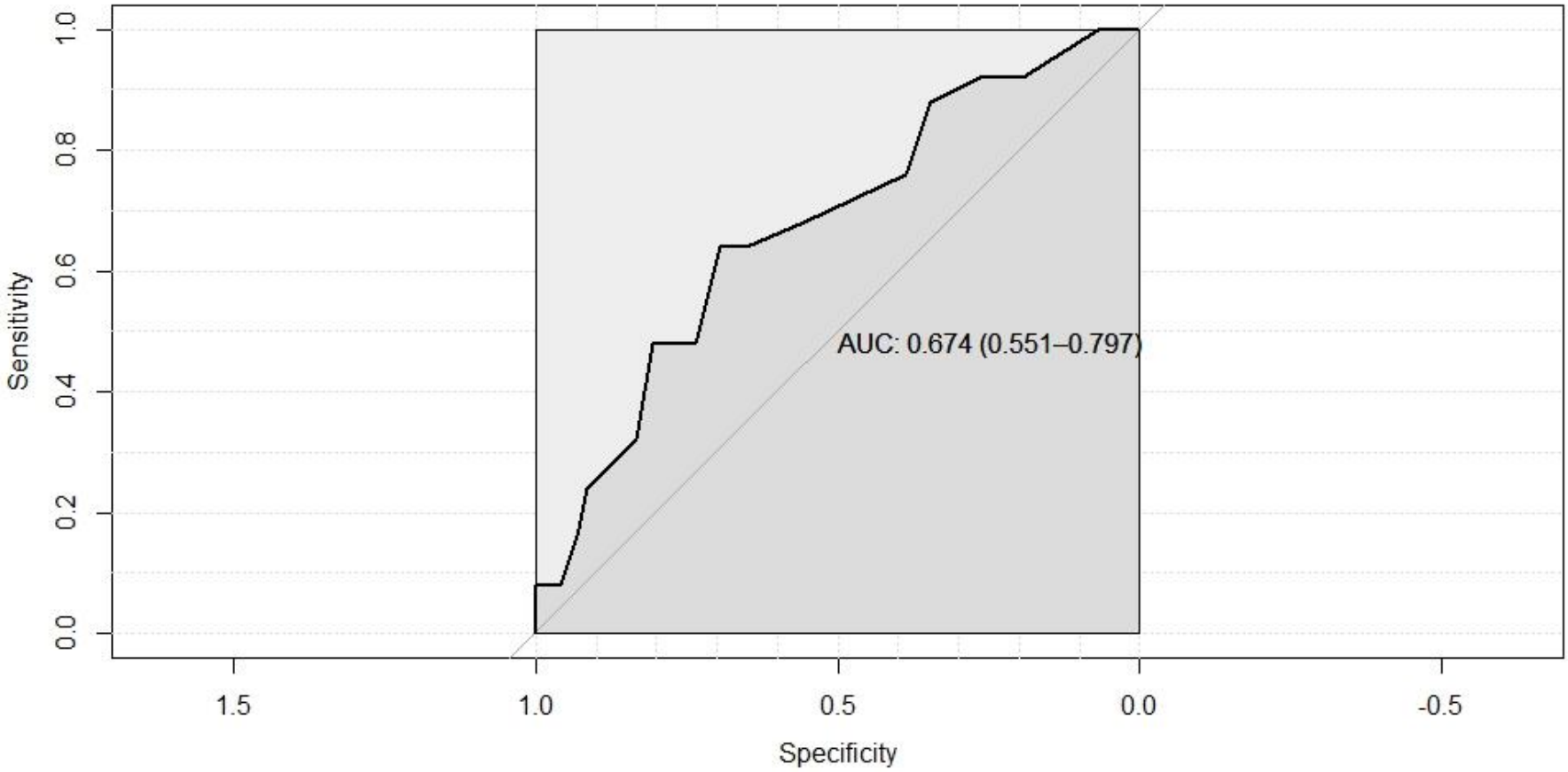

Figure 23. Area Under the Curve (AUC) graph showing that the 24-site quantification approach adequately predicts low PPT (gain-of-function) in the forearm.

5.4.4. Central Mechanisms trait

CFA revealed that all self-reported items loaded on a single factor with loading values (β-coefficients) of 0.45 το 0.79 (**Table 19**). The analysis demonstrated a good fit to the data with CFI=0.92 and TLI=0.88; RMSEA=0.09; SRMR=0.07; χ^2(df)=34.19(20), p=0.03) that allowed the derivation of a single Central Mechanisms trait score.

Table 19. Confirmatory Factor Analysis to explore the existence of an underlying single common factor between 8 self-reported items shown to be traits of Central Sensitisation

Central Mechanisms Trait		Loadings on 1 Factor			1-Factor Model		
		β	SE	p			
Variables	**Anxiety**	*0.701*	0.075	<0.01	Model Diagnostics	**χ^2**	34.19
	Depression	*0.575*	0.088	<0.01		**df**	20
	Neuropathic Pain	*0.485*	0.098	<0.01		**p-value**	0.025
	Fatigue	*0.458*	0.101	<0.01		**TLI**	0.881
	Cognition	*0.791*	0.059	<0.01		**CFI**	0.915
	Pain Distribution (>9/24)	*0.453*	0.120	<0.01		**RMSEA**	0.086
	Catastrophising	*0.491*	0.086	<0.01		**SRMR**	0.068
	Sleep	0.490	0.154	<0.01			

CFI: Comparative Fit Index, **df:** Degrees of Freedom, **RMSEA:** Root Mean Square Error of Approximation, **SE:** Standard Error, **SRMR:** Standardised Root Mean Square Residual, **TLI:** Tucker-Lewis Index,

5.4.5. Cross-sectional unadjusted associations between demographic variables, CS indices and SM/SC outcomes

Correlation values between age, sex, body mass index and CS indices as well as SM/SC outcomes are presented in **Table 20**. Women displayed higher indices of central sensitisation and lower measures of SM/SC. A small number of weak associations between age or BMI and CS or SM/SC scores were detected.

Table 20. Correlation of anthropometric variables with central sensitisation indices and self-management / self-care constructs at baseline.

	Sample Size n=97	Age		Sex		BMI	
		Cor	**p-value**	**Cor**	**p-value**	**Cor**	**p-value**
CS Indices	**PPT**	-0.05	0.60	**-0.31**	**<0.01**	-0.02	0.77
	TS	0.11	0.29	0.11	0.29	0.01	0.87
	CPM	-0.07	0.50	-0.15	0.15	-0.01	0.92
	ACR	-0.13	0.31	**0.33**	**<0.01**	0.07	0.49
	>9/24	-0.07	0.59	**0.24**	**0.04**	**0.23**	**0.05**
	CMT	**-0.37**	**<0.01**	**0.21**	**0.04**	0.10	0.35
Self-management Self-care	**HEIQ_HDB**	0.08	0.41	-0.14	0.18	**-0.20**	**0.05**
	HEIQ-PEL	**0.29**	**<0.01**	-0.19	0.07	-0.04	0.67
	HEIQ-SMI	-0.03	0.77	**-0.26**	**0.01**	0.04	0.70
	HEIQ-CAA	0.17	0.09	-0.07	0.48	-0.16	0.12
	HEIQ-STA	0.14	0.17	-0.10	0.29	0.10	0.31
	HEIQ-SIS	0.15	0.15	**-0.28**	**<0.01**	-0.03	0.74
	HEIQ-HSN	0.11	0.29	**-0.27**	**<0.01**	-0.01	0.91
	HEIQ-ED	**-0.36**	**<0.01**	0.07	0.47	0.18	0.08
	PSEQ	0.18	0.08	-0.12	0.24	-0.08	0.42
	HCUQ	**-0.21**	**0.04**	**-0.22**	**0.03**	0.11	0.25

ACR: American College of Rheumatology, **BMI:** Body Mass Index, **CMT:** Central Mechanisms Trait, **Cor:** Pearson or Spearman Correlation, **CPM**: Conditioned Pain Modulation, **CS:** Central Sensitisation, **CST:** Central Sensitisation Traits, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-Efficacy Questionnaire, **TS:** Temporal Summation, **>9/24:** 24-site body manikin classification method

Values in **bold** indicate statistical significance ($p<0.05$)

Yellow colour indicates correlation of $r<0.50$.

Correlation values between all CS indices are presented in **Table 21**. PPT displayed negative correlations with TS (r=-0.40, p<0.01), >9/24 classification method (r=-0.28, p=0.01) and Central Mechanisms trait (r=-0.19, p=0.03) and CPM correlated with TS (r=-0.22, p=0.03). The ACR classification for widespread pain correlated significantly with the >9/24 classification method (r=0.67, p<0.01) and Central Mechanisms trait (r=0.37, p<0.01) and the Central mechanism was also significantly correlated with the >9/24 classification method (r=0.48, p<0.01).

Correlation coefficient values for all combinations of CS indices and SM/SC outcomes are presented in **Table 22**. Overall, from the QST modalities, PPT did not correlate significantly (r=-0.19-0.21, p>0.05) with SM/SC outcomes at baseline. TS negatively associated with HEIQ-SMI (r=-0.19, p=0.01) and CPM positively correlated with HEIQ-PEL (r=0.22, p=0.03), HEIQ-CAA (r=0.25, p=0.03) and HEIQ-HSN (r=0.24, p=0.04). From the body manikin indices, >9/24 classification method displayed a statistically significant correlation with PSEQ (-0.24, 0.04) whereas the ACR classification method did not display significant association with any SM/SC measure (r=-0.19 to 0.18, p>0.05). The Central Mechanisms trait displayed statistically significant negative correlations with HEIQ-HDB (r=-0.32, p<0.01), HEIQ-CAA (r=-0.46, p<0.01), HEIQ-STA (r=-0.36, p<0.01), HEIQ-SIS (r=-0.27, p=0.01), HEIQ-PEL (r=-0.57, <0.01) and PSEQ (r=-0. 72, <0.01) as well as positive correlations with HEIQ-ED (r=0.68, <0.01).

Correlation values between CS indices and pain, disability and neuropathic pain as well as between CS indices and other secondary outcomes are also presented in **Table 22**. Pain severity measured using NRS showed a significant positive correlation with the Central Mechanism trait (r=0.43, p<0.01). Disability correlated with PPT (r=-0.24, p=0.03), CPM (r=-0.22, p=0.03), the >9/24 classification method (r=0.33, p<0.01) and the Central Mechanisms trait (r=0.57, p<0.01). Neuropathic-like pain correlated significantly with the ACR (r=0.34, p<0.01) and >9/24 (r=0.35, p<0.01) classification methods as well as with the Central Mechanisms trait (r=0.61, p<0.01). The Central Mechanisms trait demonstrated significant correlations with each secondary outcome (r=-0.60 to 0.80). The FMSS significantly correlated with all CS indices (r=-0.26 to 0.80) except CPM. Anxiety (HADS-A) demonstrated a significant negative correlation with PPT (r=-0.24, p=0.03) and positive correlations with TS (r=0.20, p=0.05) and the Central Mechanisms trait (r=0.75, p<0.01). Fatigue (FSS) demonstrated a positive correlation with both the ACR (r=0.25, p=0.03) and >9/24 (r=0.29, p<0.01) classification methods.

Table 21. Correlation matrix between Central Sensitisation Indices at baseline

	Sample Size n=97	PPT		TS		CPM		ACR		>9/24	
		Cor	**p-value**	**Cor**	**p-value**	**Cor**	**p-value**	**Cor**	**p-value**	**Cor**	**p-value**
CS Indices	**PPT**										
	TS	**-0.40**	**<0.01**								
	CPM	0.12	0.22	**-0.22**	**0.03**						
	ACR	-0.14	0.24	-0.06	0.64	-0.12	0.35				
	>9/24	**-0.28**	**0.01**	0.04	0.78	-0.03	0.83	**0.67**	**<0.01**		
	CMT	**-0.19**	**0.03**	0.13	0.22	-0.02	0.84	**0.37**	**<0.01**	**0.48**	**<0.01**

ACR: American College of Rheumatology, **CMT:** Central Mechanisms Trait, **CPM:** Conditioned Pain Modulation, **PPT:** Pain Pressure Detection Threshold, **TS:** Temporal Summation, **Cor:** Spearman's Rank Order Correlation, **>9/24:** 24-site body manikin classification method

Values in **bold** indicate statistical significance (p<0.05), Yellow colour indicates correlation of r<0.50, Green colour indicates correlation of r>0.50

Table 22. Correlation matrix between all CS indices and SM/SC as well as other outcomes at baseline.

	Sample Size n=97	PPT		TS		CPM		ACR		>9/24		CMT	
		Cor	p-value	Cor	p-value	Cor	p-value	Cor	p-value	Cor	p-value	Cor	p-value
Self-management / Self-care	**HEIQ-HDB**	0.06	0.61	-0.01	0.93	0.07	0.52	-0.01	0.92	-0.03	0.81	**-0.32**	**<0.01**
	HEIQ-PEL	0.07	0.61	-0.15	0.17	**0.22**	**0.03**	-0.08	0.56	-0.18	0.14	**-0.57**	**<0.01**
	HEIQ-SMI	0.09	0.40	**-0.19**	**0.01**	0.91	0.48	0.13	0.30	0.03	0.85	-0.06	0.59
	HEIQ-CAA	0.04	0.72	-0.14	0.21	**0.25**	**0.03**	-0.11	0.40	-0.18	0.14	**-0.46**	**<0.01**
	HEIQ-STA	-0.04	0.73	-0.03	0.84	0.04	0.86	0.01	0.99	0.01	0.96	**-0.36**	**<0.01**
	HEIQ-SIS	0.06	0.61	0.01	0.93	0.06	0.59	-0.07	0.60	-0.12	0.36	**-0.27**	**0.01**
	HEIQ-HSN	-0.01	0.95	0.11	0.41	**0.24**	**0.04**	-0.19	0.12	-0.01	0.93	-0.15	0.15
	HEIQ-ED	-0.09	0.48	-0.07	0.52	-0.14	0.25	0.18	0.14	0.18	0.14	**0.68**	**<0.01**
	PSEQ	0.12	0.29	-0.07	0.54	0.15	0.23	-0.13	0.31	**-0.24**	**0.04**	**-0.72**	**<0.01**
	HCUQ	-0.03	0.77	-0.08	0.51	0.19	0.09	0.01	0.99	0.08	0.54	0.11	0.29
Pain	**NRS**	-0.04	0.71	0.10	0.34	0.03	0.76	0.18	0.14	0.23	0.04	**0.43**	**<0.01**
	PDETECT	-0.11	0.27	0.07	0.52	-0.07	0.51	**0.34**	**<0.01**	**0.35**	**<0.01**	**0.61**	**<0.01**
	Now	-0.10	0.34	0.06	0.56	0.01	0.92	0.07	0.57	**0.23**	**0.05**	**0.42**	**<0.01**
	Strongest	0.10	0.35	0.11	0.29	-0.06	0.53	0.13	0.30	0.13	0.28	**0.39**	**<0.01**
	Average	-0.05	0.63	0.10	0.32	-0.04	0.69	0.19	0.11	**0.25**	**0.04**	**0.47**	**<0.01**
Negative Affect	**HADS Anx.**	**-0.24**	**0.03**	**0.20**	**0.05**	0.02	0.85	0.21	0.07	0.20	0.09	**0.75**	**<0.01**
	HADS Dep.	-0.11	0.28	-0.01	0.99	-0.02	0.82	0.13	0.28	**0.31**	**0.01**	**0.78**	**<0.01**
	PCS	**-0.23**	**0.04**	0.15	0.13	-0.03	0.81	0.15	0.22	0.20	0.08	**0.69**	**<0.01**
	TSK	-0.09	0.37	0.11	0.27	-0.10	0.30	-0.12	0.36	-0.03	0.83	**0.33**	**<0.01**
Limiting Factors	**RMDQ**	**-0.24**	**0.02**	0.18	0.08	**-0.22**	**0.03**	0.18	0.14	**0.33**	**<0.01**	**0.57**	**<0.01**
	FSS	-0.14	0.17	0.16	0.11	-0.13	0.19	**0.25**	**0.03**	**0.29**	**<0.01**	**0.52**	**<0.01**
	FSVAS	0.10	0.33	-0.12	0.23	0.07	0.50	-0.21	0.07	-0.11	0.37	**-0.41**	**<0.01**
	FMSS	**-0.26**	**0.02**	**0.25**	**0.01**	-0.10	0.32	**0.48**	**<0.01**	**0.60**	**<0.01**	**0.80**	**<0.01**
Quality of Life	**EQ-5D-5L**	0.09	0.39	-0.13	0.21	0.15	0.13	-0.05	0.71	-0.22	0.06	**-0.65**	**<0.01**
	$EQ1_{Mobility}$	-0.10	0.33	0.14	0.17	-0.14	0.16	0.02	0.90	0.17	0.15	**0.27**	**0.01**
	$EQ2_{Self-care}$	-0.07	0.52	0.12	0.25	-0.15	0.14	0.07	0.59	**0.28**	**0.01**	**0.59**	**<0.01**
	$EQ3_{Activities}$	-0.07	0.51	0.04	0.68	-0.18	0.08	-0.05	0.70	0.14	0.25	**0.34**	**<0.01**
	$EQ4_{Discomfort}$	-0.12	0.24	0.17	0.10	-0.09	0.38	0.03	0.85	0.21	0.07	**0.46**	**<0.01**
	$EQ5_{Anx/Dep}$	-0.10	0.34	0.10	0.35	-0.03	0.80	0.09	0.50	0.17	0.17	**0.72**	**<0.01**
	$EQVAS_{Health}$	**0.21**	**0.06**	-0.05	0.66	0.03	0.74	-0.09	0.49	-0.11	0.40	**-0.60**	**<0.01**
Risk	**STarT Back**	-0.17	0.10	0.17	0.09	-0.15	0.13	0.20	0.08	0.25	0.03	**0.69**	**<0.01**

ACR: American College of Rheumatology, **CMT:** Central Mechanisms Trait, **Cor:** Pearson or Spearman Correlation, **CPM**: Conditioned Pain Modulation, **EQ-5D-5L:** Quality of Life Instrument, **EQVAS:** Quality of Life Visual Analogue Scale, **FMSS:** Fibromyalgia Severity Scale, **FSS:** Fatigue Severity Scale, **FSVAS:** Fatigue Visual Analogue Scale, **HADS:** Hospital Anxiety & Depression Scale, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **NRS:** Pain Numerical Rating Scale, **PCS:** Pain Catastrophising Scale, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **RMDQ:** Roland-Morris Disability Questionnaire, **STarT-Back:** Stratification tool, **TS:** Temporal Summation, **TSK:** Tampa Scale of Kinesiophobia, **>9/24:** 24-site body manikin classification method

Values in **bold** indicate statistical significance (p<0.05), Yellow colour indicates correlation of r<0.50, Green colour indicates correlation of r>0.50

Correlation values between secondary psychological variables and SM/SC outcomes are presented in **Table 23**. Depression (HADS-D) demonstrated fair to good significant correlations with all SM/SC outcomes (r=-0.75 to 0.58) and HCUQ. PSEQ, HEIQ-PEL, HEIQ-CAA and HEIQ-ED were significantly correlated with all secondary variables at baseline (r=-0.61 to 0.67) with the rest of the SM/SC outcomes displayed significant correlations (r=-0.49 to 0.39) with most of the secondary outcomes.

Correlation values between each secondary variable aimed to be used in multivariable regression models are presented in **Table 24**. Age and sex demonstrated a significant correlation only with depression (r=-0.39, $p<0.01$) and programme allocation (r=0.23, $p<0.05$) respectively whereas the rest psychological variables demonstrated significant positive or negative correlations with each other (r=-0.65 to 0.58, $p<0.01$).

Table 23. Correlation matrix of SM/SC constructs with other outcomes at baseline.

	Sample Size n=97	Self-Management Domains (HEIQ)								Self-Care	
		HEIQ-HDB	**HEIQ-PEL**	**HEIQ-SMI**	**HEIQ-CAA**	**HEIQ-STA**	**HEIQ-SIS**	**HEIQ-HSN**	**HEIQ-ED**	**PSEQ**	**HCUQ**
		Cor	**Cor**	**Cor**	**Cor**	**Cor**	**Cor**	**Cor**	**Cor**	**Cor**	**Cor**
Pain	**NRS**	-0.18	**-0.25***	0.02	**-0.26***	-0.13	-0.16	-0.02	0.16	**-0.49*****	0.13
	PDETECT	-0.17	**-0.24****	0.08	**-0.28*****	-0.08	-0.09	0.03	**0.36*****	**-0.52*****	0.17
	Now	-0.19	**-0.32*****	0.03	**-0.21***	-0.08	-0.18	0.02	0.14	**-0.51*****	**0.24***
	Strongest	**-0.28*****	**-0.30*****	0.03	**-0.26*****	**-0.23***	-0.15	-0.06	**0.25****	**-0.41*****	0.09
	Average	**-0.23***	**-0.29*****	0.11	**-0.26*****	-0.16	-0.07	0.02	**0.23***	**-0.49*****	**0.21***
Negative Affect	**HADS Anx.**	-0.18	**-0.50*****	-0.14	**-0.47*****	**-0.31*****	**-0.32*****	**-0.20***	**0.61*****	**-0.52*****	0.18
	HADS Dep.	**-0.37*****	**-0.72*****	**-0.20***	**-0.70*****	**-0.47*****	**-0.49*****	**-0.22***	**0.58*****	**-0.75*****	0.12
	PCS	**-0.20***	**-0.40*****	0.02	**-0.37*****	-0.17	-0.07	-0.06	**0.70*****	**-0.54*****	0.17
	TSK	-0.11	**-0.20*****	0.07	**-0.23*****	-0.10	-0.05	-0.13	**0.57*****	**-0.31*****	0.07
Limiting Factors	**RMDQ**	**-0.37*****	**-0.46*****	-0.02	**-0.47*****	-0.10	-0.16	-0.14	**0.48*****	**-0.70*****	0.10
	FSS	**-0.27*****	**-0.27*****	-0.01	**-0.32*****	**-0.21*****	-0.11	-0.06	**0.35*****	**-0.43*****	-0.09
	FSVAS	0.08	**0.24*****	0.03	0.18	0.07	0.11	0.16	**-0.28*****	**0.40*****	**-0.22*****
	FMSS	**-0.21***	**-0.44*****	0.04	**-0.39*****	**-0.22***	**-0.24***	**-0.21*****	**0.48*****	**-0.51*****	0.12
Quality of Life	**EQ-5D-5L**	**0.31*****	**0.50*****	0.06	**0.52*****	0.18	**0.28*****	0.12	**-0.50*****	**0.72*****	-0.12
	EQ1$_{Mobility}$	**-0.31*****	**-0.25*****	0.03	**-0.30*****	-0.04	-0.08	0.04	0.18	**-0.51*****	-0.01
	EQ2$_{Self-care}$	**-0.29*****	**-0.41*****	-0.14	**-0.45*****	-0.16	-0.15	-0.07	**0.36*****	**-0.61*****	0.07
	EQ3$_{Activities}$	**-0.31*****	**-0.37*****	0.02	**-0.39*****	-0.19	-0.06	0.01	**0.24***	**-0.53*****	0.13
	EQ4$_{Discomfort}$	-0.15	**-0.25*****	-0.05	**-0.27*****	-0.07	-0.16	-0.06	**0.24***	**-0.57*****	0.09
	EQ5$_{Anx/Dep}$	**-0.24***	**-0.48*****	-0.01	**-0.49*****	**-0.29*****	**-0.33*****	-0.12	**0.65*****	**-0.52*****	0.15
	EQVAS$_{Health}$	**0.45*****	**0.47*****	**0.24***	**0.39*****	**0.39*****	**0.30*****	**0.35*****	**-0.55*****	**0.61*****	-0.14
Risk	**STarT-Back**	**-0.27*****	**-0.52*****	-0.16	**-0.51*****	**-0.35*****	**-0.28*****	**-0.22***	**0.67*****	**-0.58*****	0.10

Cor: Pearson or Spearman Correlation, **FMSS:** Fibromyalgia Severity Scale, **FSS:** Fatigue Severity Scale, **FSVAS:** Fatigue Severity Visual Analogue Scale, **EQ-5D-5L:** Quality of Life Instrument, **EQVAS:** Quality of Life Visual Analogue Scale, **HADS:** Hospital Anxiety & Depression Scale, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **NRS:** Pain Numerical Rating Scale, **PCS:** Pain Catastrophising Scale, **PSEQ:** Pain Self-Efficacy Questionnaire, **RMDQ:** Roland-Morris Disability Questionnaire, **STarT-Back:** Stratification tool, **TSK:** Tampa Scale of Kinesiophobia

Values in **bold** indicate statistical significance (p<0.05), Yellow colour indicates correlation of r<0.50, Green colour indicates correlation of r>0.50

Table 24. Bivariate correlations between each psychological variable aimed to be used as independent variables in multivariable regression models.

Independent Variables at Baseline	Independent Variables at Baseline								
	Age	**Sex**	**Pain**	**Disability**	**Depression**	**Catastrophisation**	**Fatigue**	**QoL**	**Programme**
Age	-								
Sex	-0.14	-							
Pain	0.02	-0.08	-						
Disability	-0.05	0.04	**0.48*** **	-					
Depression	**-0.39*** **	0.02	**0.37*** **	**0.49*** **	-				
Catastrophisation	-0.19	0.02	**0.31*** **	**0.52*** **	**0.58*** **	-			
Fatigue	-0.06	0.15	**0.28** **	**-0.64*** **	**0.35*** **	0.18	-		
Quality of Life	0.12	0.04	**-0.55*** **	**0.37*** **	**-0.65*** **	**-0.65*** **	**-0.35*** **	-	
Programme	-0.03	**0.23***	**0.41*** **	**0.58*** **	**0.50*** **	**0.47*** **	**0.31*** **	**-0.52*** **	-

QoL: Quality of Life

Values in **bold** indicate statistical significance (p<0.05)

Yellow colour indicates correlation of r<0.50

Green colour indicates correlation of r>0.50

5.4.6. Cross-sectional associations between CS indices and SM/SC outcomes adjusted for other factors

Multicollinearity testing yielded VIF values ranging from 1.19 to 1.97 for all independent variables indicating non-significant multicollinearity between them.

Details of the regression models between the different QST modalities and SM/SC outcomes are provided in **Table 25**. The association between TS and HEIQ-SMI as well as between CPM and HEIQ-PEL, HEIQ-CAA and HEIQ-HSN demonstrated in bivariate analyses remained significant in multivariable models which included age, sex, pain, depression, catastrophising and fatigue scores. TS demonstrated also a significant association with HEIQ-PEL that was not present in bivariate analyses. PPT, were not significantly associated ($p>0.05$) with any of the SM/SC outcomes at baseline.

Details of the regression models between the other CS indices and SM/SC outcomes are provided in **Table 26**. The ACR and >9/24 widespread pain classification methods were not significantly associated ($p>0.05$) with any of the SM/SC outcomes at baseline whereas the associations between the single Central Mechanisms trait and HEIQ-PEL, HEIQ-CAA, HEIQ-STA, HEIQ-ED and PSEQ demonstrated in bivariate analyses remained significant in multivariable models which included age, sex, pain, and quality of life scores.

Depression demonstrated significant associations across most models regardless of CS indices used as an independent variable. Catastrophisation, pain and fatigue were marginally associated significantly with some of the outcomes. In models where the Central Mechanisms trait was included as an independent variable, sex and quality of life yielded significant associations with HEIQ-SIS, HEIQ-HSN and HEIQ-ED, PSEQ respectively. HCUQ was universally associated only with age and sex.

Table 25. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct adjusted for age, sex, pain, depression, catastrophisation and fatigue.

	Variables used as adjustments	HEIQ-HDB β	HEIQ-PEL β	HEIQ-SMI β	HEIQ-CAA β	HEIQ-STA β	HEIQ-SIS β	HEIQ-HSN β	HEIQ-ED β	PSEQ β	HCUQ β
Quantitative Sensory Testing	PPT †	-0.02	-0.02	0.01	-0.01	-0.001	0.03	-0.12	0.03	-0.01	-0.09
	Age	-0.004	-0.001	-0.002	-0.01	-0.002	-0.02	-0.01	**-0.01***	-0.004	**-0.02****
	Sex	-0.23	**-0.03****	**-0.07***	-0.10	-0.01	**-1.26****	**-1.13****	0.07	-0.12	**-0.64****
	Pain	-0.01	0.001	0.002	0.01	0.001	-0.03	0.02	**-0.02***	**-0.02****	0.04
	Depression	**-0.09****	**-0.01*****	**-0.03****	**-0.15*****	**-0.05*****	**-0.50*****	**-0.19****	0.03	**-0.09*****	-0.06
	Catastrophisation	0.004	0.002	0.003	0.001	0.003	**0.05****	0.01	**0.03*****	-0.004	0.01
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.003	0.001	**0.002****	-0.001	-0.002
	TS †	0.12	**-1.02***	**-0.39***	-0.94	-0.003	-0.71	-2.35	0.10	0.11	-0.45
	Age	-0.004	-0.001	-0.002	-0.01	-0.001	-0.02	-0.01	**-0.01***	-0.004	**-0.02***
	Sex	-0.21	**-0.22***	**-0.07***	-0.07	-0.01	**-1.28*****	**-0.99****	0.05	-0.11	**-0.56****
	Pain	-0.006	0.002	0.003	0.01	0.001	-0.03	0.02	**-0.02***	**-0.02****	0.04
	Depression	**-0.08****	**-0.15*****	**-0.03*****	**-0.15*****	**-0.01*****	**-0.51*****	**-0.21****	0.03	**-0.09*****	-0.06
	Catastrophisation	0.005	0.004	**0.003***	0.004	0.001	**0.05****	0.02	**0.03*****	-0.005	0.02
	Fatigue	-0.001	0.001	0.001	-0.001	0.001	0.003	0.002	**0.002****	-0.001	-0.001
	CPM †	0.001	**0.001***	0.001	**0.002****	0.001	0.001	**0.004***	-0.001	0.001	0.002
	Age	-0.004	-0.001	-0.002	-0.01	-0.001	-0.02	-0.004	**-0.01***	-0.004	**-0.02***
	Sex	-0.21	**-0.002***	**-0.01***	-0.05	-0.06	**-1.27*****	**-0.95****	0.03	-0.10	**-0.53***
	Pain	-0.01	0.002	0.002	0.01	0.001	-0.03	0.01	-0.02	**-0.02****	0.03
	Depression	**-0.09****	**-0.01*****	**-0.02****	**-0.15*****	**-0.05*****	**-0.50*****	**-0.20****	0.03	**-0.09*****	-0.06
	Catastrophisation	0.005	0.002	0.002	0.003	0.004	**0.05****	0.02	**0.03*****	-0.004	0.02
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.003	0.002	**0.002****	-0.001	-0.001

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at baseline. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 22** and **Table 23**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.0 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 97 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Table 26. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct adjusted for age, sex, pain, depression, catastrophisation and fatigue.

	Variables used as adjustments	HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Body Manikin	ACR †	0.17	0.14	0.06	0.04	0.03	0.33	-0.20	0.08	0.05	0.10
	Age	-0.003	-0.001	-0.002	-0.01	-0.001	-0.02	-0.01	**-0.01***	-0.004	**-0.02***
	Sex	**-0.26***	**-0.28****	**-0.10****	-0.10	-0.10	**-1.39*****	**-0.98****	0.03	**-0.13***	**-0.59****
	Pain	-0.01	0.001	0.001	0.01	-0.002	-0.04	0.02	**-0.03***	**-0.02****	0.03
	Depression	**-0.08****	**-0.14*****	**-0.02****	**-0.15*****	**-0.05*****	**-0.50*****	**-0.20****	0.03	**-0.09*****	-0.06
	Catastrophisation	0.004	0.002	0.002	0.001	0.003	**0.05****	0.02	**0.03*****	-0.001	0.02
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.002	0.002	**0.002****	-0.001	-0.002
	>9/24 †	0.19	0.11	0.05	0.11	0.09	0.30	0.36	-0.06	0.07	0.18
	Age	-0.004	-0.001	-0.002	-0.01	-0.001	-0.02	-0.01	**-0.008***	-0.004	**-0.02****
	Sex	-0.26	**-0.26****	**-0.10***	-0.11	-0.10	**-1.36*****	**-1.12****	0.06	**-0.13***	**-0.60****
	Pain	-0.01	0.004	0.002	0.01	-0.001	-0.03	0.02	-0.02	**-0.02****	0.03
	Depression	**-0.09****	**-0.15*****	**-0.03****	**-0.15*****	**-0.05*****	**-0.51*****	**-0.21****	0.03	**-0.09*****	-0.06
	Catastrophisation	0.004	0.002	0.002	0.001	0.003	**0.05****	0.01	**0.03*****	-0.004	0.02
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.002	0.001	**0.001****	-0.001	-0.002
Trait Score	CMT †	-0.22	**-0.36****	0.0002	**-0.29***	**-0.15***	-0.12	0.09	**0.61*****	**-0.32*****	-0.13
	Pain	-0.007	0.005	0.002	0.01	0.002	-0.03	0.03	**-0.03***	-0.01	0.02
	Age	-0.001	0.004	-0.001	0.002	0.0001	0.01	0.01	-0.01	-0.001	**-0.02***
	Sex	-0.17	-0.13	-0.07	-0.01	-0.02	**-1.09***	**-0.99****	-0.04	-0.05	**-0.56***
	Quality of Life	0.23	0.46	0.06	**0.66**	0.10	1.49	1.24	**-0.58***	**0.67*****	-0.36

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at baseline. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 22** and **Table 23**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.0 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 97 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Exploratory secondary analyses of regression models assessing the relationship between CS indices and SM/SC outcomes where pain was replaced as an independent variable by disability, demonstrated negative associations of similar strength as those seen in primary analyses although, disability was significantly associated with self-management outcomes (HEIQ-HDB) that pain did not demonstrate any association with. Details of exploratory analyses are given in **Table 27** and **Table 28**.

The explanatory power of each regression model conducted for the purposes of primary and secondary analyses are given in **Appendix 8**. All models with SM/SC outcomes as dependent variables and CS indices as the primary independent variable demonstrated statistically significant ($p<0.05$) explanatory power (Adjusted R^2) in primary and secondary analyses ($R^2=0.07$ to 0.69) apart from HEIQ-SMI that, in models where the Central Mechanisms trait was the primary independent variable demonstrated non-statistically significant ($p>0.05$) explanatory power ($R^2=0.001$).

Table 27. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct adjusted for age, sex, disability, depression, catastrophisation and fatigue.

	Sample Size n=97	HEIQ-HDB β	HEIQ-PEL β	HEIQ-SMI β	HEIQ-CAA β	HEIQ-STA β	HEIQ-SIS β	HEIQ-HSN β	HEIQ-ED β	PSEQ β	HCUQ β
Quantitative Sensory Testing	PPT †	-0.04	-0.02	0.01	-0.002	-0.001	0.03	-0.12	0.03	-0.03	-0.08
	Age	-0.003	-0.001	-0.002	-0.001	-0.002	-0.02	-0.01	**-0.01***	-0.003	**-0.02****
	Sex	-0.23	**-0.25****	**-0.07***	-0.001	-0.08	**-1.25****	**-1.15****	0.09	-0.11	**-0.66****
	Disability	**-0.03***	**-0.02***	0.002	**-0.02***	0.01	-0.01	-0.02	0.01	**-0.03*****	0.02
	Depression	**-0.07***	**-0.13*****	**-0.03****	**-0.13*****	**-0.05*****	**-0.52*****	**-0.17***	0.02	**-0.08*****	-0.06
	Catastrophisation	0.009	0.005	0.002	0.01	0.002	**0.05****	0.02	**0.03*****	-0.002	0.01
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.002	0.002	**0.001***	-0.001	-0.001
	TS †	0.24	**-0.94***	**-0.40***	-0.08	-0.36	-0.81	-2.24	0.03	0.20	-0.49
	Age	-0.003	0.001	-0.002	-0.004	-0.002	-0.02	-0.004	**-0.01***	-0.003	**-0.02***
	Sex	-0.20	**-0.22***	**-0.07***	-0.07	-0.06	**-1.26*****	**-1.01****	0.06	-0.09	**-0.59****
	Disability	**-0.03***	-0.02	0.002	-0.02	0.01	0.01	-0.02	-0.006	**-0.03*****	0.02
	Depression	**-0.07***	**-0.14*****	**-0.03****	**-0.14*****	**-0.05*****	**-0.52*****	**-0.20***	0.02	**-0.08*****	-0.06
	Catastrophisation	0.01	0.01	**0.003***	0.007	0.003	**0.05****	0.02	**0.03*****	-0.001	0.01
	Fatigue	-0.001	0.001	0.001	-0.001	-0.001	0.002	0.002	**0.001***	-0.001	-0.001
	CPM †	-0.001	0.001	0.001	**0.002****	0.001	0.001	**0.004***	-0.001	0.001	0.002
	Age	-0.003	0.001	-0.002	-0.004	-0.002	-0.02	-0.004	**-0.009***	-0.003	**-0.02***
	Sex	-0.20	**-0.21***	**-0.08***	-0.05	-0.06	**-1.25****	**-0.96****	0.05	-0.09	**-0.55***
	Disability	**-0.03***	-0.01	0.002	-0.01	0.01	0.02	-0.003	0.003	**-0.03*****	0.03
	Depression	**-0.07***	**-0.14*****	**-0.03****	**-0.14*****	**-0.05*****	**-0.52*****	**-0.19****	0.02	**-0.08*****	-0.06
	Catastrophisation	0.009	0.005	0.002	0.005	0.002	**0.05****	0.02	**0.03*****	-0.001	0.01
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.002	0.002	**0.001***	-0.001	-0.001

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at baseline. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 22** and **Table 23**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.0 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 97 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Table 28. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct adjusted for age, sex, disability, depression, catastrophisation and fatigue.

	Sample Size n=97	HEIQ-HDB β	HEIQ-PEL β	HEIQ-SMI β	HEIQ-CAA β	HEIQ-STA β	HEIQ-SIS β	HEIQ-HSN β	HEIQ-ED β	PSEQ β	HCUQ β
Body Manikin	ACR †	0.18	0.14	0.06	0.06	0.03	0.28	-0.16	0.04	0.03	0.13
	Age	-0.002	0.001	-0.002	-0.004	-0.002	-0.02	-0.006	**-0.009***	-0.003	**-0.02***
	Sex	-0.25	**-0.28****	**-0.10****	-0.10	-0.08	**-1.35*****	**-1.00****	0.05	-0.10	**-0.63****
	Disability	**-0.03***	**-0.02***	0.002	**-0.02***	0.008	0.01	-0.02	0.006	**-0.03*****	0.02
	Depression	**-0.07***	**-0.13*****	**-0.02****	**-0.13*****	**-0.05*****	**-0.52*****	**-0.18***	0.02	**-0.08*****	-0.06
	Catastrophisation	0.01	0.004	0.002	0.005	0.002	**0.05****	0.02	**0.03*****	-0.001	0.01
	Fatigue	-0.001	0.001	0.001	-0.001	-0.001	0.002	0.002	**0.001***	-0.001	-0.002
	>9/24 †	0.23	0.14	0.05	0.16	0.08	0.26	0.41	-0.10	0.10	0.18
	Age	-0.003	-0.001	-0.002	-0.004	-0.002	-0.02	-0.005	**-0.009***	-0.003	**-0.02***
	Sex	**-0.24***	**-0.26****	**-0.09****	-0.12	-0.09	**-1.33*****	**-1.13*****	0.08	-0.11	**-0.63****
	Disability	**-0.03***	**-0.002***	0.001	**-0.02***	0.007	-0.007	-0.03	-0.008	**-0.03*****	0.02
	Depression	**-0.08****	**-0.14*****	**-0.03*****	**-0.14*****	**-0.05*****	**-0.52*****	**-0.19***	0.02	**-0.09*****	-0.06
	Catastrophisation	0.01	0.002	0.002	0.01	0.002	**0.05****	0.02	**0.03*****	-0.001	0.01
	Fatigue	-0.001	-0.001	0.001	-0.001	-0.001	0.002	0.001	**0.001***	-0.001	-0.002
Trait Score	CMT †	-0.13	**-0.29***	-0.001	-0.18	**-0.17****	-0.21	0.20	**0.51*****	**-0.26*****	-0.13
	Age	**-0.03***	-0.02	0.002	**-0.03***	0.01	0.01	-0.02	0.02	**-0.02*****	0.01
	Sex	-0.0001	0.01	-0.001	0.003	-0.0001	0.01	0.01	-0.01	-0.0003	**-0.02***
	Disability	-0.17	-0.14	-0.07	-0.02	-0.02	**-1.07***	**-1.01****	-0.01	-0.04	**-0.57***
	Quality of Life	-0.03	0.24	0.06	0.35	-0.03	1.79	0.90	-0.23	**0.50****	-0.40

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at baseline. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 22** and **Table 23**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.0 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 97 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

5.5. Discussion

This study showed that different indices of central sensitisation are associated with aspects of self-management in individuals with CLBP at a single timepoint. Depression, alongside other previously recognised indicators of poor treatment outcomes and potential drivers of central sensitisation such as catastrophising and fatigue, were also found to be associated with aspects of SM/SC in a single timepoint. Twenty-four sites on a body pain manikin appear to be the most appropriate classification method in CLBP as 9 or more painful sites can predict low pain thresholds at a distal site and therefore indicate increased pain sensitisation. Elements of depression, anxiety, neuropathic-like pain, catastrophising, fatigue, sleep disturbance, poor cognition and pain distribution can be considered a reliable index of CS when combined into a single factor as it was found to be associated with different SM/SC outcomes as well as with different CS indices.

Higher initial levels of pain sensitisation were associated with low scores in HEIQ-PEL, HEIQ-SMI and HEIQ-CAA both in bivariate correlation and multivariable regression analyses, indicating a link between CS and those specific aspects of self-management. The Central Mechanisms trait was associated with most SM/SC outcomes in unadjusted correlations and only with HEIQ-PEL, HEIQ-CAA, HEIQ-STA, HEIQ-ED and PSEQ in adjusted correlations. Depression was associated with all domains of SM/SC except HCUQ in unadjusted and adjusted correlation demonstrating a direct relationship between negative affect and SM/SC.

HEIQ-PEL measures the capacity of individuals to be active and engage in life-fulfilling activities with intent to promote change of lifestyle and improve their life circumstances, HEIQ-CAA measures the ability of individuals to adapt attitudes and approaches that minimise the effects of their condition without allowing it to control their life, and HEIQ-SMI encapsulates an individual's ability to monitor their condition physically or emotionally with demonstrated insight that leads to taking appropriate actions that promote self-management (Osborne et al., 2007). CS has been shown to be associated with poor health-related quality of life, greater disability and negative affect in individuals with CLBP (Smart et al., 2012a) whereas depression has been found to be associated with functional disability (Hung et al., 2015), failure to return to work (Parker et al., 2015) and limited treatment responses (Nicholas, 2007a).

Sex rather than age or body mass seems to be a factor influencing both SM/SC and CS indices at a single time-point as it yielded the most associations out of all

demographic variables in unadjusted and adjusted correlations. The findings indicate that females demonstrate lower pain pressure thresholds and their pain is more widely spreading across more areas. Also, females tend to demonstrate reduced engagement in life in a productive way, reduced self-monitoring insight, are more socially isolated, utilise more healthcare and do not navigate through healthcare services as effectively as their male counterparts. Female gender has been previously associated with reduced pain thresholds (Rolke et al., 2006a) as well as with poor SM/SC outcomes in individuals with CLBP (Ferrari et al., 2019, Koleck et al., 2006). The findings also suggest that older individuals possess the skills to more effectively engage in life-fulfilling activities but at the same time they tend to utilise healthcare more and cannot cope adequately with emotional distress. Similarly, increased body mass raises the possibility for pain to be distributed across larger anatomical sites and interferes with the ability of individuals to adapt health-directed behaviours such as exercise.

Significant associations between each of the CS indices are consistent with each measuring a related construct, which makes them valid approaches for the measurement and identification of CS. PPT was correlated with TS and the central mechanisms trait and TS correlated also with CPM. The >9/24 method correlated with PPT whereas both pain distribution approaches correlated significantly with each other and with the Central Mechanisms trait. Although the strength of the observed associations is often weak, this is consistent with CS being a complex and heterogeneous phenomenon, rather than a single and simple entity. Exploring associations between CS and self-management can be a challenging feat as self-management is also a notoriously difficult construct to measure and shares similar complexity and multidimensionality as CS. It encapsulates the ability of individuals to cope with their condition biologically, psychologically and socially (Barlow et al., 2002) and can be considerably variable in the way individuals (patients and clinicians alike) perceive it or attempt to enhance it (Nolte et al., 2013, Oliveira et al., 2012). Such findings further justify the implementation of different CS indices from studies that try to assess the effect of CS on psychophysical outcomes or the effect of specific treatments on CS.

Out of all three QST modalities, CPM and TS were found to be associated with self-management outcomes in unadjusted and adjusted correlations. CPM and TS are considered dynamic modalities, designed to measure more complex course of pain processing (Arendt-Nielsen and Yarnitsky, 2009). CPM specifically, tries to evoke cerebral processes implicated also in depressive disorders even without the presence

of obvious pathology (Arendt-Nielsen et al., 2018) whereas TS and negative affect have been frequently found to be influenced by one another (Gay et al., 2015, George et al., 2006). CS and depression might affect the ability of individuals to self-manage their condition in a similar manner although, their independent association with HEIQ-PEL, HEIQ-SMI and HEIQ-CAA in a single model suggests that they independently influence different aspects of the same self-management domain. Even though functional magnetic resonance imaging studies suggest that chronic pain and depression share similar brain activation pathways (Sheng et al., 2017, Han and Pae, 2015, Mutschler et al., 2012), CS might uniquely affect brain areas involved in pain processing. Centrally sensitised individuals have demonstrated altered function of brain areas (amygdala) that can control the ability of individuals to be more physically active (perform exercises) or engage positively in life and in daily activities (Simons et al., 2014, Hadjikhani et al., 2013, Kim et al., 2013). CS can also negatively affect areas responsible for the cognitive-evaluative dimensions of pain (prefrontal cortex) leading to inefficient pain-modulation, pain amplification, and adaptation of less constructive attitudes and approaches towards pain (pain anticipation, pain expectancy) (Atlas and Wager, 2012, Taylor et al., 2012, Kong et al., 2007). It is possible as well that the independent association of CS and depression with HEIQ-PEL, HEIQ-SMI and HEIQ-CAA are because those self-management domains are embodied by psychometric properties that involve physical and emotional responses influenced by both CS and depression. PPT as a static modality able to assess only a single sensation (pressure), might have not been able to fully capture such complex and subtle alterations in central pain processing, which could explain why no significant associations were demonstrated with any of the SM/SC domains.

The number of painful sites on a body manikin is a measure of widespread pain and ongoing research has added to its construct, content and face validity over the years (Wolfe et al., 2019, Wolfe et al., 2016, Wolfe et al., 2011). Since central sensitisation is considered a mechanism that drives widespread pain (Arendt-Nielsen et al., 2018) and facilitates the manifestation of pain at a site distant to the site of pathology (Woolf, 2011), a link between the number of painful sites and QST at a distant site seems straightforward. A recent study demonstrated that PPT at a distant site is associated with the number of painful sites as well as with the shaded distribution of the pain in people with knee pain (Akin-Akinyosoye et al., 2018). The results of the present study compliment further those findings as more than 9/24 painful sites optimally predicted low PPT at the forearm. That means that pain sensitisation can be identified with a combination of pain indices from PPT and a small number of painful sites on a body

manikin. These findings appear similar to the number of painful sites cut-off (>7/15) recommended by the ACR for identification of widespread pain (Wolfe et al., 2019). However, the ACR criteria, in combination with the number of painful sites, require the pain to manifest without a known cause in at least 4 out of 5 different regions, divided between left and right upper (1, 2) and lower limbs (3, 4) as well as neck, upper back and lower back (5). This consists a significant difference from >9/24 classification that simply takes into account the number of painful sites irrespective of the region. Another significant difference between the classification methods is that in >9/24 the number of painful sites reported could have been influenced by the prevalence of radiculopathy within the patient population. This brings into question how useful the ACR criteria can be for identification of widespread pain in a population with CLBP, a pathology that is frequently accompanied by elements of neuropathic-like pain. The significant unadjusted association between the >9/24 method and PSEQ or the Self-care dimension of the EQ-5D-5L suggests that a simple count of the number of painful sites is not enough to influence self-management outcomes and is potentially lacking the psychophysical components possessed by QST and the Central Mechanisms trait in capturing the influence of CS on the ability of individuals to self-manage their condition. Another explanation for no significant associations could be that any present relationship between >9/24 and SM/SC outcomes is mediated by the association of >9/24 with psychological variables such as depression.

In the current study, 8 previously identified and validated key traits in people with knee pain (Akin-Akinyosoye et al., 2018), loaded together onto a single Central Mechanisms construct considered to be reflecting central pain mechanisms in people with CLBP. Those traits were unique self-reported items representative of anxiety, depression, catastrophising, neuropathic-like pain, fatigue, sleep disturbance, pain distribution, and cognitive impact that are thought to be indicative of CS (Graven-Nielsen and Arendt-Nielsen, 2002).

As depression, catastrophising and fatigue are integral parts of the Central Mechanisms trait, they could not be used as independent variables in a single model with the derived CS index. Nevertheless, negative affect in the model can be implied as the items comprising the Central Mechanisms construct represent the emotional aspects of central sensitisation, which appears to be driven by overlapping mechanisms within the central nervous system (Akin-Akinyosoye et al., 2018). Insight about pain processing within the brain neuromatrix has highlighted how sensitisation to painful stimuli in areas other than the spinal cord alters the function of brain regions responsible for monitoring the emotional and cognitive-evaluative aspects of pain

leading to an amplified pain experience physically as well as emotionally (Nijs et al., 2015). The unique among CS indices unadjusted and adjusted association of the Central Mechanisms trait with HEIQ-STA suggests an implication of CS in the ability of individuals to effectively implement knowledge-based skills and techniques to manage condition-specific symptoms (Osborne et al., 2007) and further highlights the cognitive rather than just the physical or emotional aspects of CS. Since increased levels of depression, anxiety, catastrophising, fatigue and neuropathic-like pain were found to be independently associated with elevated emotional distress (HEIQ-ED) and reduced self-efficacy (PSEQ) in unadjusted correlations, an independent association of the Central Mechanisms trait with a domain that measures condition-related negative affect and a domain that captures the levels of confidence to perform physical, mental and social tasks, despite the pain, was anticipated.

Exploratory analyses featuring disability instead of pain in regression models provided no additional insight, as they demonstrated negative associations with SM/SC outcomes of similar strength to the one displayed by pain, without influencing the relationship of the other variables (age, sex, CS indices, depression, catastrophising and fatigue) with SM/SC outcomes. Such findings suggest that disability is not significantly different to pain in the way it influences SM/SC in individuals with CLBP or that its' relationship is influenced by low mood and fatigue. Past research suggests that CLBP-related disability appears to be predominantly driven by pain as well as the associated mood disturbance, fatigue, and decreased self-efficacy that it entails (Salvetti et al., 2012, Truchon, 2001). Nevertheless, disability demonstrated independent adjusted associations with HEIQ-HDB, HEIQ-CAA and PSEQ whereas pain was associated only with PSEQ in adjusted models. It is possible that RMDQ shares psychological constructs similar to those in HEIQ-HDB and HEIQ-CAA that reflect the ability of individuals to manage their condition within the paradigm of the specific domains whereas pain intensity was measured in a unidimensional numerical scale that cannot provide inference for effective SM/SC. Also, out of the two outcomes, disability correlated with PPT, TS and the Central Mechanisms trait compared to pain that correlated only with the single factor construct. Pain and disability seldomly correlate with CS indices such as QST suggesting that CS is not influencing how individuals report pain or disability (Hübscher et al., 2013).

This study is subject to several strengths and limitations. It is the first to explore associations between CS indices and SM/SC constructs with the aim to explain the apparent reduced long-term effectiveness of SM interventions in CLBP (Du et al., 2017). CS is an elusive concept with no universally accepted way to successfully

diagnose it in patients suffering from a chronic musculoskeletal condition (Girbés et al., 2013). For that purpose, different indices, proposed by literature, were implemented to increase the chances of CS identification in the study population. The Central Mechanisms trait developed for this study used self-reported items with demonstrated links with centrally augmented pain. However, the single construct was originally developed for a population suffering from knee pain that might display significant differences in pain processing and manifestation of CS than a population with CLBP. Similarly, self-management is a diverse concept with no widely accepted way of measurement. The instrument used for measuring self-management (HEIQ), even though it demonstrates good psychometric properties and encapsulates most self-management constructs, is a tool generic to populations with chronic pain and not specific to individuals with CLBP. For the purposes of confounding identification and appropriate analyses, a large variety of secondary outcomes were used that have been previously found to either influence SM/SC or be influenced by CS. Even though a great number of secondary endpoints were taken into account, the influence of CS indices on SM/SC might be mediated by other, unmeasured, socioeconomic confounders and therefore, findings should be interpreted with caution. Also, the experience of potential side-effects such as decreased reaction time, cloudy judgment, and drowsiness due to ongoing consumption of opioids or other strong medication was not recorded. Such side-effects could also influence the capacity of individuals to self-manage and therefore mediate the relationship of CS indices on SM/SC. The sample size, even though sufficient to power this study might be small to provide conclusive findings. Also, the study participants might not accurately reflect CLBP in the overall population with the condition as only individuals following a specific rehabilitation pathway were included. The item used to measure healthcare utilisation in this study was a custom tool that was adapted from an existed validated tool and underwent no previous reliability and validity testing. Therefore, any information provided by its use must be viewed with caution. Finally, even though the findings are promising and showcase that CS indices might influence the ability of individuals to self-manage their condition, the strength of demonstrated associations needs to be explored prospectively to examine the capacity of CS to predict SM/SC outcomes.

5.6. Conclusion

Levels of CS, measured through different indices, demonstrate significant associations with different domains of SM/SC indicating partial influence of central

pain mechanisms on the ability of individuals with CLBP to self-manage their condition. Such associations are suggestive of cognitive as well as physical and emotional aspects in CS. Depression appears to be the strongest indicator of poor SM/SC with levels of catastrophising and fatigue contributing to a lesser degree. The exploration of prospective associations between baseline CS indices and long-term SM/SC outcomes is needed to provide more information regarding the influence of CS on SM/SC.

6. PROSPECTIVE ASSOCIATIONS BETWEEN CS INDICES ANS SELF-MANAGEMENT/SELF-CARE OUTCOMES IN INDIVIDUALS WITH CLBP

6.1. Background

Prognosis refers to the scientific approach of longitudinally observing a population with a medical condition, gathering useful information at selected stages, and predicting the probable course and outcome of that condition over time (Hayden et al., 2008). Prognostic insight is used by clinicians for patient education purposes, to stratify target groups into appropriate treatment pathways or target interventions on specific factors thought to be influencing clinical outcomes (Croft et al., 2006). Prospective cohort studies are considered an appropriate observational study design to robustly answer scientific questions and provide reliable prognostic insight (Vandenbroucke, 2008). To gain prognostic inference, a multivariable approach is necessary in order to optimise or compare performance between predictive models (Moons et al., 2009).

Prognostic research is of particular importance in the field of LBP where identification of prognostic factors or individuals at high risk of developing CLBP could allow for more targeted, tailored to the individual treatment and prevention of chronicity, which could lead to enhanced self-management outcomes (LeResche et al., 2013). Nevertheless, CLBP is prevalent in a substantial proportion of people presenting with back pain and the fact it is not curable by current treatments leads to prioritisation of self-management interventions over other approaches (NICE, 2016). Self-management approaches for CLBP have demonstrated only moderate long-term effectiveness (Du et al., 2017), which highlights the need for more effective self-management approaches and identification of factors hindering the long-term ability of individuals with CLBP to manage their condition.

Central Sensitisation appears to be present in patient cohorts with CLBP in the form of reduced pain thresholds at sites distant to the site of pathology (Arendt-Nielsen et al., 2018) as well as inefficient descending inhibition and amplified temporal summation (McPhee et al., 2020) and is linked with worse prognosis in CLBP-related treatment outcomes. Early evidence of CS are predictive of worse musculoskeletal

outcomes (Georgopoulos et al., 2019) however, whether CS predicts or influences long-term SM/SC outcomes is yet unexplored.

6.2. Aims and Objectives

6.2.1. Aims

The main aim of this chapter was to examine whether CS predicts or may contribute to 3-months follow-up self-management outcomes.

6.2.2. Objectives

- To test whether the associations identified in the previous chapter between baseline CS indices and SM/SC outcomes are also present between baseline CS indices and follow-up SM/SC outcomes, after study participants have undertaken an intervention programme that aimed to improve such outcomes.
- To determine whether different baseline CS indices associated specifically with all SM/SC measures at 3 months.
- To explore changes SM/SC outcomes between baseline and follow-up.
- To identify predictors of poor SM/SC other than CS.

6.3. Methods

The full methodological details of this study are given in the Methods chapter (**METHODS**) and only a brief outline is presented in this chapter whenever necessary.

Analysis of prospectively collected baseline and follow-up data was used to assess the association of baseline CS indices with follow-up (3-months) SM/SC outcomes in individuals with CLBP and changes in measured outcomes between timepoints.

6.3.1. Study participants

Participants who enrolled in the study on day 1 of their participation in a group intervention programme (baseline) were invited to undertake a clinical examination on the last day (approximately 3 months after day 1) of their intervention (follow-up), featuring QST assessment and completion of the questionnaire booklet. At follow-up, QST assessment was offered to all participants. In cases where participants did not

complete the programme or were unable to attend for QST, were given the option to return the questionnaire booklet via mail. All participants were eligible for a follow-up examination and no minimum amount of attended intervention sessions was required. Details on follow-up assessments and procedures to minimise loss to follow-up are given in **2.8.2.1**.

6.3.2. Clinical assessment and application methods

All QST applications and used self-reported outcome measures were identical to those used at baseline and are described in detail in sections **2.6.1**, **2.6.2** and **2.6.3** of the Methods chapter.

6.3.3. Data analysis and analytical procedures

Data normality determination, data transformation, determination of difference between groups and timepoints, calculations of the MCID, QST calculation approaches, correlation and regression modelling (association between baseline CS indices and SM/SC outcomes at 3-months follow-up) that were implemented for the purposes of prospective analyses were identical with those used in cross-sectional analyses in **Chapter 6** and are given in detail in the sections **2.9.2.1**, **2.9.2.2, 2.9.2.3**, **2.10.3.1**, **2.10.3.2**, **2.10.3.3**, **2.10.3.4.1**, **2.10.3.4.2** and **2.10.3.4.3** of the Methods chapter.

6.4. Results

6.4.1. Data management and transformation

Details about data distribution as well as the results of the log-transformation process can be found in **Appendix 9**. Not all primary and secondary outcomes demonstrated normal distribution upon testing. TS, HEIQ-SMI, HEIQ-STA, HEIQ-HSN, PSEQ, HCUQ, NRS, HADS-Dep., FSS, EQ-5D-5L and STarT-Back demonstrated distributions significantly different to normal after logarithmic transformation.

6.4.2. Demographics data and clinical characteristics

Baseline and follow-up population demographic data and clinical features for QST and psychological variables are given in **Table 29** and **Table 30**.

Table 29. Overall patient demographic characteristics, self-management/self-care measures and pain sensitivity factors at baseline and 3 months follow-up along with the timepoint change of each factor and indicators of its' statistical as well as clinical significance.

Variables (Unit, Value, Range, %)	**Baseline** Mean (± SD) Median (IQ Range)	**3 Months** Mean (± SD) Median (IQ Range)	**Timepoint Change** Mean (± SD) Median (IQ Range)	**Change Significance** WSRT (p-value)	**MCID** SD †
No. Participants	97	87	87		
Age (y)	56 (±13)	57 (±13)	57 (±13)		
BMI	29.4 (25.7 to 34.6)	29.4 (26.0 to 34.5)	29.4 (26.0 to 34.5)		
Female (%)	71%	67%	67%		
Setting					
Hospital	92	82	82		
Community	5	5	5		
Primary Outcomes (SM/SC) †					
Health Education Impact Questionnaire Domains					
Health Directed Behaviour (1-4)	2.5 (2.3 to 3.0)	3.0 (2.8 to 3.5)	+0.3 (0.0 to 0.8)	**72 (<0.01)**	**+0.3**
Positive Engagement in Life (1-4)	2.6 (2.2 to 3.0)	3.0 (2.6 to 3.2)	+0.2 (0.0 to 0.6)	**432 (<0.01)**	**+0.2**
Self-monitoring & Insight (1-4)	3.0 (2.8 to 3.2)	3.0 (3.0 to 3.3)	+0.2 (0.0 to 0.3)	**620 (<0.01)**	**+0.2**
Constructive Attitudes & Approaches (1-4)	2.8 (2.2 to 3.0)	3.0 (2.6 to 3.4)	+0.0 (0.0 to 0.4)	**535 (<0.01)**	+0.3
Skill & Technique Acquisition (1-4)	2.8 (2.5 to 3.0)	3.0 (2.8 to 3.0)	+0.2 (0.0 to 0.5)	**639 (<0.01)**	+0.3
Social Integration and Support (1-4)	2.8 (2.4 to 3.0)	3.0 (2.6 to 3.0)	+0.0 (-0.2 to 0.4)	**779 (0.04)**	+0.3
Health Services Navigation (1-4)	3.0 (2.6 to 3.2)	3.0 (2.8 to 3.4)	+0.2 (-0.1 to 0.4)	**666 (0.01)**	+0.3
Emotional Distress (1-4)	2.8 (2.3 to 3.2)	2.5 (2.0 to 3.0)	-0.2 (-0.7 to 0.0)	**2173 (<0.01)**	-0.3
Self-Care †					
Pain Self-Efficacy Questionnaire (0-60)	27 (20 to 41)	35 (28 to 47)	+5 (-2 to 11)	**791 (<0.01)**	+7
Health Care Utilisation Questionnaire (Units)	4 (2 to 6)	9 (6 to 12) ¥	+4 (2 to 8)	**454 (<0.01)**	**+2**
QST †, ₸, $					
Pressure Pain Detection Threshold (kPa)	205.8 (148.2 to 297.6)	182.9 (126.0 to 244.5)	-29.7 (-60.8 to 22.3)	**1316 (0.02)**	-41.1
Temporal Summation (0-10)	1.0 (0.4 to 2.8)	1.8 (0.7 to 3.3)	+0.08 (-0.19 to 0.89)	611 (0.06)	+0.7
Conditioned Pain Modulation (kPa)	59.1 (5.6 to 99.3)	41.3 (4.2 to 90.9)	+1.3 (-85.1 to 58.0)	1089 (0.43)	+54.9

BMI: Body Mass Index, **IQ:** Interquartile, **kPa:** Kilopascals, **MCID:** Minimum Clinically Important Difference, **QST:** Quantitative Sensory Testing, **SD:** Standard Deviation, **SM/SC:** Self-management/Self-care, **WSRT:** Wilcoxon Signed Rank Test (Paired)

† MCID determination was based on one-half the SD calculated for each outcome at baseline.

¥ The number of visitations due to participation in treatment programmes is included in calculations of health-care utilisation

₸ Calculations are based on 62 participants who underwent QST at follow-up.

$ Negative change indicates sensitivity increase in pressure pain detection threshold, positive change indicates sensitivity increase in temporal summation and positive change indicates sensitivity decrease in conditioned pain modulation.

Values in **bold** indicate statistical (p<0.05) and/or clinical significance (timepoint change > ½ baseline SD)

Table 30. Overall patient secondary outcome measures at baseline and 3 months follow-up along with the timepoint change of each factor and indicators of its' statistical as well as clinical significance.

Variables (Unit, Value, Range, %)	**Baseline** Mean (± SD) Median (IQ Range)	**3 Months** Mean (± SD) Median (IQ Range)	**Timepoint Change** Mean (± SD) Median (IQ Range)	**Change Significance** WSRT (p-value)	**MCID** SD †
Secondary Outcomes †					
Pain Numerical Rating Scale (0-10)	6 (5 to 7)	5 (4 to 7)	-1 (-2 to 0)	**1765 (<0.01)**	**-1**
PainDETECT (0-38)	17 (12 to 24)	16 (11 to 22)	-1 (-6 to 2)	**2137 (<0.01)**	-4
Now (0-10)	6 (4 to 7)	5 (3 to 7)	-1 (-2 to 1)	**1794 (<0.01)**	**-1**
Strongest (0-10)	8 (8 to 9)	8 (7 to 9)	0 (-1 to 0)	**991 (0.01)**	-1
Average (0-10)	6 (6 to 7)	6 (5 to 7)	-1 (-2 to 0)	**1471 (0.01)**	**-1**
Hospital *Anxiety* Scale (0-21)	9 (6 to 13)	9 (5 to 12)	-1 (-3 to 1)	**1961 (0.01)**	-3
Hospital *Depression* Scale (0-21)	9 (5 to 12)	7 (3 to 10)	-1 (-3 to 0)	**2062 (<0.01)**	-2
Pain Catastrophising Scale (0-52)	22 (11 to 31)	13 (5 to 23)	-4 (-9 to 0)	**2507 (<0.01)**	-7
Tampa Scale of Kinesiophobia (17-68)	38 (33 to 43)	33 (28 to 38)	-4 (-7 to 0)	**2765 (<0.01)**	**-4**
Roland-Morris Disability Questionnaire (0-24)	13 (9 to 18)	10 (5 to 14)	-3 (-6 to 0)	**2492 (<0.01)**	**-3**
Fatigue Severity Scale (7-63)	42 (29 to 52)	36 (28 to 46)	-4 (-1 to 2)	**2196 (0.01)**	-7
Fatigue-Visual Analogue Scale (0-10)	5 (3 to 6)	5 (4 to 7)	0 (-1 to 2)	1007 (0.22)	+1
Fibromyalgia Severity Scale (0-31)	13 (8 to 18)	12 (7 to 15)	-1 (-2 to 0)	**1868 (<0.01)**	-3
EQ-5D-5L (Index)	0.45 (0.23 to 0.64)	0.53 (0.34 to 0.69)	+0.02 (-0.04 to 0.13)	**872 (<0.01)**	+0.12
EQ-Visual Analogue Scale (0-100)	50 (40 to 65)	60 (40 to 75)	+7.1 (±17.0)	**716 (<0.01)**	+10
STarT-Back (0-9)	6 (4 to 7)	4 (2 to 6)	-1 (-2 to 0)	**1866 (<0.01)**	**-1**

EQ-5D-5L: Quality of Life Instrument, **IQ:** Interquartile, **MCID:** Minimum Clinically Important Difference, **SD:** Standard Deviation, **STarT-Back:** Stratification tool, **WSRT:** Wilcoxon Signed Rank Test (Paired)

† MCID determination was based on one-half the SD calculated for each outcome at baseline.

Values in **bold** indicate statistical (p<0.05) and/or clinical significance (timepoint change > ½ baseline SD)

Baseline and follow-up data separated according to intervention programme are given in **Table 31**, **Table 32**, **Table 33** and **Table 34**. Out of the 97 (71% females, mean age 56±13 years) participants at baseline, 87 (67% females, mean age 57±13 years) provided SM/SC outcomes data at 3-months follow-up and 10 (50% females, mean age 49±16 years) dropped out of the study without giving a reason. Out of the 87 individuals providing follow-up data, 62 (50% females, mean age 57±13 years) participated in QST assessments and 25 (66% females, mean age 54±14 years) opted to return only a completed questionnaire booklet. No significant differences ($p>0.05$) were observed between the participants who provided QST data at follow-up and those who did not, as revealed by unpaired Wilcoxon signed-rank tests for all QST modalities.

Table 31. Patient demographic characteristics, self-management/self-care measures, and pain sensitivity factors based on participation in the physiotherapy-led pain management program at baseline and 3 months follow-up along with the timepoint change of each factor and indicators of its' statistical as well as clinical significance.

Physiotherapy-Led Intervention Programme (PT)					
Variables (Unit, Value, Range, %)	**Baseline** Mean (± SD), Median (IQ Range)	**3 Months** Mean (± SD), Median (IQ Range)	**Timepoint Change** Mean (± SD), Median (IQ Range)	**Change Significance** WSRT (p-value)	**MCID** SD †
No. Participants	42	39	39		
Age (y)	57 (±13)	58 (±13)	58 (±13)		
BMI	28.7 (25.4 to 33.4)	28.7 (25.5 to 33.5)	28.7 (25.5 to 33.5)		
Female (%)	52%	51%	51%		
Setting					
Hospital	38	35	35		
Community	4	4	4		
Primary Outcome †					
Health Education Impact Questionnaire Domains					
Health Directed Behaviour (1-4)	2.8 (2.4 to 3.0)	3.0 (3.0 to 3.4)	+0.2 (0.0 to 0.8)	**20 (<0.01)**	+0.3
Positive Engagement in Life (1-4)	2.9 (2.6 to 3.2)	3.0 (2.9 to 3.4)	+0.2 (0.0 to 0.3)	**134 (0.04)**	**+0.2**
Self-monitoring & Insight (1-4)	3.0 (2.8 to 3.2)	3.0 (3.0 to 3.3)	+0.0 (-0.1 to 0.2)	156 (0.29)	+0.2
Constructive Attitudes & Approaches (1-4)	3.0 (2.8 to 3.4)	3.2 (3.0 to 3.4)	+0.0 (-0.2 to 0.2)	129 (0.37)	+0.3
Skill & Technique Acquisition (1-4)	2.8 (2.5 to 3.0)	3.0 (2.8 to 3.0)	+0.0 (-0.2 to 0.4)	208 (0.43)	+0.3
Social Integration and Support (1-4)	3.0 (2.6 to 3.2)	3.0 (3.0 to 3.1)	+0.0 (-0.1 to 0.3)	115 (0.12)	+0.2
Health Services Navigation (1-4)	3.0 (2.8 to 3.4)	3.0 (2.9 to 3.4)	+0.2 (0.0 to 0.2)	140 (0.06)	+0.3
Emotional Distress (1-4)	2.7 (2.0 to 3.4)	2.2 (1.8 to 3.4)	-0.2 (-0.7 to 0.0)	**403 (0.01)**	-0.3
Self-Care †					
Pain Self-Efficacy Questionnaire (0-60)	41 (32 to 48)	45 (38 to 52)	+2 (-2 to 9)	**234 (0.03)**	+5
Health Care Utilisation Questionnaire (Units)	3 (2 to 6)	6 (5 to 7) ¥	+3 (0 to 4)	189 (0.06)	**+3**
QST †, ₸, $					
Pain Pressure Detection Threshold (kPa)	240.4 (175.3 to 293.9)	222.5 (155.6 to 328.7)	-8.0 (-60.6 to 44.8)	324 (0.66)	-43.7
Temporal Summation (0-10)	0.8 (0.4 to 2.4)	1.4 (0.7 to 2.7)	+0.2 (-0.1 to 1.5)	150 (0.06)	+0.6
Conditioned Pain Modulation (kPa)	66.3 (22.5 to 110.3)	52.4 (2.9 to 97.1)	-25.2 (-95.8 to 51.4)	387 (0.13)	-53.6

BMI: Body Mass Index, **IQ:** Interquartile, **kPa:** Kilopascals, **MCID:** Minimum Clinically Important Difference, **PT:** Physiotherapy Intervention Programme, **QST:** Quantitative Sensory Testing, **SD:** Standard Deviation, **WSRT:** Wilcoxon Signed Rank Test (Paired)

† MCID determination was based on one-half the SD calculated for each outcome at baseline.

¥ The number of visitations due to participation in treatment programmes is included in calculations of health-care utilisation. ₸ Calculations are based on 34 participants of the PT-L programme who underwent QST at follow-up. $ Negative change indicates sensitivity increase in pressure pain detection threshold, positive change indicates sensitivity increase in temporal summation and positive change indicates sensitivity decrease in conditioned pain modulation.

Values in **bold** indicate statistical ($p<0.05$) and/or clinical significance (timepoint change > ½ baseline SD)

Table 32. Patient secondary psychological measures based on participation in the physiotherapy-led pain management program at baseline and 3 months follow-up along with the timepoint change of each factor and indicators of its' statistical as well as clinical significance.

Physiotherapy-Led Intervention Programme (PT)					
Variables (Unit, Value, Range, %)	**Baseline** Mean (± SD), Median (IQ Range)	**3 Months** Mean (± SD), Median (IQ Range)	**Timepoint Change** Mean (± SD), Median (IQ Range)	**Change Significance** WSRT (p-value)	**MCID** SD †
Secondary Outcomes †					
Pain Numerical Rating Scale (0-10)	5 (4 to 6)	4 (3 to 6)	-1 (-2 to 1)	**416 (0.04)**	**-1**
PainDETECT (0-38)	15 (11 to 19)	12 (8 to 16)	-2 (-6 to 1)	**515 (<0.01)**	-3
Now (0-10)	5 (3 to 6)	4 (3 to 6)	-1 (-2 to 1)	405 (0.26)	**-1**
Strongest (0-10)	8 (7 to 9)	8 (6 to 8)	0 (-1 to 0)	175 (0.12)	-1
Average (0-10)	6 (5 to 7)	6 (5 to 7)	0 (-2 to 1)	290 (0.24)	-1
Hospital *Anxiety* Scale (0-21)	7 (5 to 9)	5 (3 to 8)	-1 (-3 to 0)	**416 (0.01)**	-2
Hospital *Depression* Scale (0-21)	5 (3 to 9)	4 (2 to 7)	-1 (-3 to 0)	**398 (<0.01)**	-2
Pain Catastrophising Scale (0-52)	13 (7 to 21)	6 (5 to 13)	-3 (-5 to 0)	**376 (0.01)**	-5
Tampa Scale of Kinesiophobia (17-68)	36 (31 to 42)	31 (28 to 35)	-4 (-8 to 0)	**537 (<0.01)**	**-4**
Roland-Morris Disability Questionnaire (0-24)	10 (5 to 12)	6 (3 to 10)	-2 (-5 to 0)	**487 (<0.01)**	**-2**
Fatigue Severity Scale (7-63)	33 (26 to 45)	30 (24 to 38)	-4 (-11 to 3)	469 (0.08)	-7
Fatigue-Visual Analogue Scale (0-10)	5 (3 to 7)	6 (4 to 8)	0 (-1 to 3)	143 (0.06)	+1
Fibromyalgia Severity Scale (0-31)	9 (6 to 14)	9 (5 to 12)	-1 (-3 to 0)	**338 (0.03)**	-2
EQ-5D-5L (Index)	0.61 (0.46 to 0.70)	0.65 (0.50 to 0.72)	0.0 (-0.05 to 0.08)	217 (0.38)	+0.09
EQ-Visual Analogue Scale (0-100)	61 (50 to 70)	75 (53 to 82)	+5 (-5 to 15)	**196 (0.03)**	+9
STarT-Back (0-9)	4 (3 to 6)	3 (2 to 4)	-2 (-5 to 0)	**428 (0.01)**	**-1**

BMI: Body Mass Index, **EQ-5D-5L:** Quality of Life Instrument, **IQ:** Interquartile, **kPa:** Kilopascals, **MCID:** Minimum Clinically Important Difference, **PT:** Physiotherapy Intervention Programme, **QST:** Quantitative Sensory Testing, **SD:** Standard Deviation, **STarT-Back:** Stratification toot, **WSRT:** Wilcoxon Signed Rank Test (Paired)

† MCID determination was based on one-half the SD calculated for each outcome at baseline.
Values in **bold** indicate statistical (p<0.05) and/or clinical significance (timepoint change > ½ baseline SD)

Table 33. Patient demographic characteristics, self-management/self-care measures, and pain sensitivity factors based on participation in the multidisciplinary team-led pain management program at baseline and 3 months follow-up along with the timepoint change of each factor and indicators of its' statistical as well as clinical significance.

Multidisciplinary-led Intervention Programme (MDT)					
Variables (Unit, Value, Range, %)	**Baseline** Mean (± SD), Median (IQ Range)	**3 Months** Mean (± SD), Median (IQ Range)	**Timepoint Change** Mean (± SD), Median (IQ Range)	**Change Significance** WSRT (p-value)	**MCID** SD †
No. Participants	55	48	48		
Age (y)	55 (±14)	56 (±13)	56 (±13)		
BMI	31.6 (26.1 to 36.0)	31.6 (26.8 to 36.3)	31.6 (26.8 to 36.3)		
Female (%)	75%	79%	79%		
Setting					
Hospital	54	47	47		
Community	1	1	1		
Primary Outcome †					
Health Education Impact Questionnaire Domains					
Health Directed Behaviour (1-4)	2.5 (2.0 to 2.9)	3.0 (2.8 to 3.5)	+0.5 (0.0 to 1.0)	**18 (<0.01)**	**+0.3**
Positive Engagement in Life (1-4)	2.4 (2.2 to 2.7)	2.8 (2.4 to 3.0)	+0.2 (0.0 to 0.6)	**86 (<0.01)**	**+0.2**
Self-monitoring & Insight (1-4)	3.0 (2.8 to 3.1)	3.0 (3.0 to 3.3)	+0.2 (0.0 to 0.4)	**148 (<0.01)**	**+0.2**
Constructive Attitudes & Approaches (1-4)	2.6 (2.2 to 2.8)	2.8 (2.4 to 3.0)	+0.2 (0.0 to 0.6)	**122 (<0.01)**	+0.3
Skill & Technique Acquisition (1-4)	2.8 (2.5 to 3.0)	3.0 (2.8 to 3.0)	+0.2 (0.0 to 0.7)	**111 (<0.01)**	+0.3
Social Integration and Support (1-4)	2.6 (2.3 to 3.0)	2.8 (2.2 to 3.0)	+0.0 (-0.2 to 0.4)	299 (0.14)	+0.3
Health Services Navigation (1-4)	3.0 (2.6 to 3.0)	3.0 (2.6 to 3.1)	+0.0 (-0.2 to 0.4)	**206 (0.05)**	+0.3
Emotional Distress (1-4)	3.0 (2.7 to 3.0)	2.7 (2.3 to 3.1)	-0.3 (-0.7 to 0.0)	**713 (<0.01)**	**-0.3**
Self-Care †					
Pain Self-Efficacy Questionnaire (0-60)	21 (16 to 27)	30 (22 to 37)	+8 (0 to 11)	**182 (<0.01)**	**+5**
Health Care Utilisation Questionnaire (Units)	4 (2 to 6)	11 (10 to 13) ¥	+8 (4 to 10)	**647 (<0.01)**	**+2**
QST †, ₸, $					
Pain Pressure Detection Threshold (kPa)	170.5 (110.4 to 249.8)	155.6 (120.7 to 204.8)	-39.4 (-60.7 to -5.4)	**347 (<0.01)**	**-34.6**
Temporal Summation (0-10)	1.1 (0.5 to 3.1)	2.3 (0.7 to 3.7)	+0.03 (-0.2 to 0.7)	166 (0.59)	+0.8
Conditioned Pain Modulation (kPa)	46.7 (4.2 to 88.3)	32.4 (11.6 to 79.0)	+19.2 (-54.0 to 60.5)	168 (0.44)	+56.0

BMI: Body Mass Index, **IQ:** Interquartile, **kPa:** Kilopascals, **MCID:** Minimum Clinically Important Difference, **MDT:** Multidisciplinary Intervention Programme, **QST:** Quantitative Sensory Testing, **SD:** Standard Deviation, **WSRT:** Wilcoxon Signed Rank Test (Paired)

† MCID determination was based on one-half the SD calculated for each outcome at baseline.

¥ The number of visitations due to participation in treatment programmes is included in calculations of health-care utilisation. ₸ Calculations are based on 28 participants of the MDT-L programme who underwent QST at follow-up. $ Negative change indicates sensitivity increase in pressure pain detection threshold, positive change indicates sensitivity increase in temporal summation and positive change indicates sensitivity decrease in conditioned pain modulation. Values in **bold** indicate statistical (p<0.05) and/or clinical significance (timepoint change > ½ baseline SD)

Table 34. Patient secondary psychological measures, and pain sensitivity factors based on participation in the multidisciplinary team-led pain management program at baseline and 3 months follow-up along with the timepoint change of each factor and indicators of its' statistical as well as clinical significance.

Multidisciplinary-led Intervention Programme (MDT)					
Variables (Unit, Value, Range, %)	**Baseline** Mean (± SD), Median (IQ Range)	**3 Months** Mean (± SD), Median (IQ Range)	**Timepoint Change** Mean (± SD), Median (IQ Range)	**Change Significance** WSRT (p-value)	**MCID** SD †
Secondary Outcomes †					
Pain Numerical Rating Scale (0-10)	7 (6 to 7)	5 (5 to 7)	-1 (-2 to 0)	**479 (0.01)**	**-1**
PainDETECT (0-38)	21 (16 to 25)	21 (16 to 26)	-1 (-4 to 2)	561 (0.29)	-4
Now (0-10)	6 (6 to 7)	5 (4 to 7)	-1 (-2 to 0)	**502 (<0.01)**	**-1**
Strongest (0-10)	9 (8 to 9)	8 (7 to 9)	0 (-1 to 0)	**345 (0.02)**	-1
Average (0-10)	7 (6 to 8)	6 (5 to 7)	-1 (-1 to 0)	**474 (0.01)**	**-1**
Hospital *Anxiety* Scale (0-21)	12 (8 to 15)	11 (8 to 15)	-1 (-3 to 2)	595 (0.14)	-2
Hospital *Depression* Scale (0-21)	12 (8 to 13)	10 (6 to 12)	-2 (-4 to 0)	**673 (<0.01)**	**-2**
Pain Catastrophising Scale (0-52)	26 (18 to 35)	18 (9 to 30)	-6 (-13 to 1)	**660 (<0.01)**	**-6**
Tampa Scale of Kinesiophobia (17-68)	39 (35 to 44)	36 (30 to 40)	-4 (-6 to -1)	**877 (<0.01)**	**-4**
Roland-Morris Disability Questionnaire (0-24)	16 (6 to 20)	13 (5 to 18)	-3 (-6 to 0)	**796 (<0.01)**	**-2**
Fatigue Severity Scale (7-63)	46 (39 to 53)	41 (33 to 48)	-4 (-13 to 1)	**646 (0.04)**	-6
Fatigue-Visual Analogue Scale (0-10)	5 (3 to 6)	5 (4 to 6)	0 (-1 to 1)	399 (0.90)	+1
Fibromyalgia Severity Scale (0-31)	16 (12 to 20)	14 (10 to 20)	-1 (-5 to 1)	**641 (0.04)**	-3
EQ-5D-5L (Index)	0.32 (0.14 to 0.53)	0.42 (0.26 to 0.59)	+0.10 (0.0 to 0.20)	**231 (<0.01)**	+0.12
EQ-Visual Analogue Scale (0-100)	40 (35 to 50)	50 (40 to 65)	+5 (0 to 11)	**169 (<0.01)**	+8
STarT-Back (0-9)	6 (5 to 8)	5 (3 to 7)	-1 (-2 to 0)	**513 (<0.01)**	**-1**

BMI: Body Mass Index, **EQ-5D-5L:** Quality of Life Instrument, **IQ:** Interquartile, **kPa:** Kilopascals, **MCID:** Minimum Clinically Important Difference, **MDT:** Multidisciplinary Intervention Programme, **QST:** Quantitative Sensory Testing, **SD:** Standard Deviation, **STarT-Back:** Stratification toot, **WSRT:** Wilcoxon Signed Rank Test (Paired)

† MCID determination was based on one-half the SD calculated for each outcome at baseline.
Values in **bold** indicate statistical significance (p<0.05)

When the population was separated according to allocated group intervention programme, patients under the MDT programme participated in a median of 9/10 (IQR: 8 to 10) intervention sessions and participants following the PT programme participated in a median of 5/5 (IQR: 4 to 5) intervention sessions. Overall, participants demonstrated improved ability to self-manage in all heiQ domains (±0.2 to ±0.3/4.0) as well as reduced pain (-1/10), depression (-1/21), anxiety (-1/21) and catastrophising (-4/52) levels compared to baseline. Participants in the PT programme showed reduced levels of pain (-1/10), depression (-1/21), anxiety (-1/21) and catastrophising (-3/52) as well as increased capacity to self-manage in almost all heiQ domains (0.0 to ±0.2/4.0) compared to baseline. Participants in the MDT programme demonstrated reduced levels of pain (-1/10), depression (-2/21), anxiety (-1/21) and catastrophising (-6/52) and improved self-management capacity on almost all heiQ domains (0.0 to ±0.5/4.0) compared to baseline. The trends the SM/SC outcomes followed from baseline to follow-up for the overall as well as per programme participation population are illustrated in **Figure 24**.

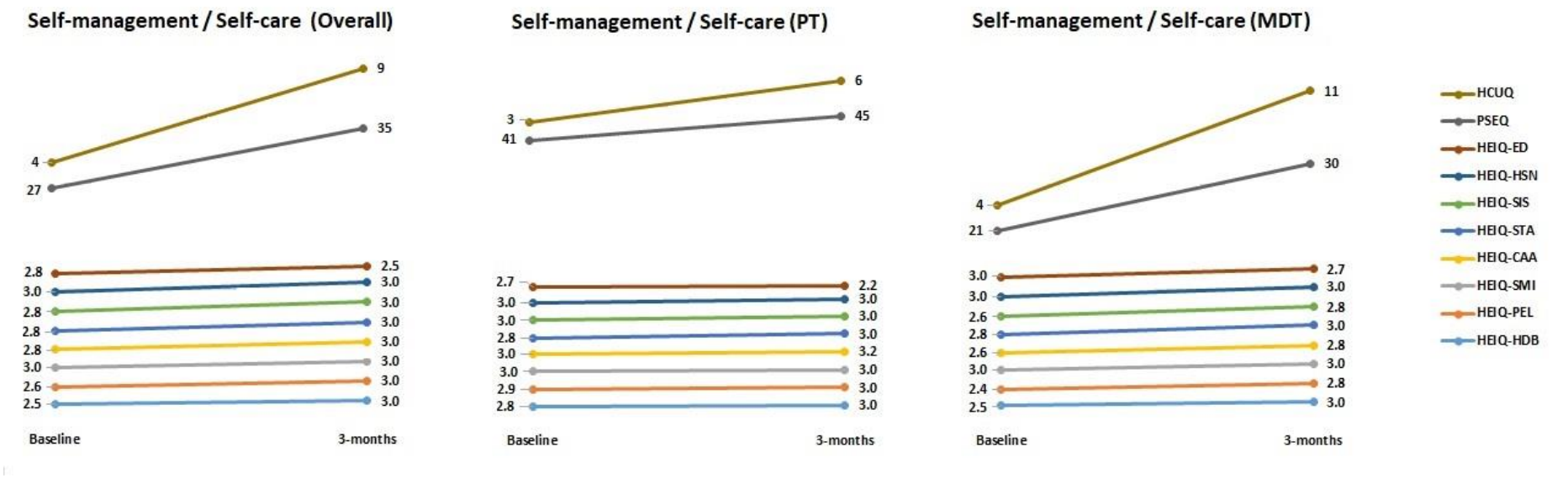

Figure 24. Self-management/Self-care outcomes trends depicting increase from baseline to 3-months follow-up in the entire population as well as across populations separated according to group intervention programme.

HCUQ: Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **MDT:** Multidisciplinary-lead Group Intervention Programme, **PPT:** Pressure Pain Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **PT:** Physiotherapy-lead group intervention programme, **TS:** Temporal Summation

Paired Wilcoxon Signed Rank tests between the study population at baseline and at 3-months follow-up as a whole revealed that changes in self-management outcomes as well as in secondary endpoints were statistically significant ($p<0.05$) apart from changes in the fatigue severity visual analogue scale ($p=0.22$). Despite the statistically significant difference between baseline and follow-up measurements, out of all primary and secondary outcomes, positive engagement in life, health-directed behaviour, self-monitoring and insight, health-care utilisation, pain, kinesiophobia and the STarT-Back tool demonstrated clinically meaningful differences (**Table 29** and **Table 30**). Statistical and clinical significance values are given separately for each intervention programme in **Table 31**, **Table 32**, **Table 33** and **Table 34**.

Median values of QST at follow-up (n=62) showed an overall increase in pain sensitivity, illustrated by a decreased median PPT value of 182.9 (IQR: 126.0 to 244.5, $p=0.02$). Increases in median TS value of 1.8 (IQR: 126.0 to 244.5, $p=0.06$) and median CPM value of 1.3 (-85.1 to 58.0, $p=0.43$) compared to baseline did not reach statistical significance. The trends QST followed from baseline to follow-up for the overall as well as per programme participation population are illustrated in **Figure 25**. SD calculations revealed that increase in pain sensitivity as demonstrated by all QST modalities in the overall population could be the result of measurement error. When SD calculations were conducted specifically for group populations, the pain sensitivity increase captured by PPT in the MDT-L group (-39.4, IQR: -60.7 to -5.4, $p<0.01$), was higher than the calculated SD (-34.6) for that group and therefore clinically significant (**Table 33**).

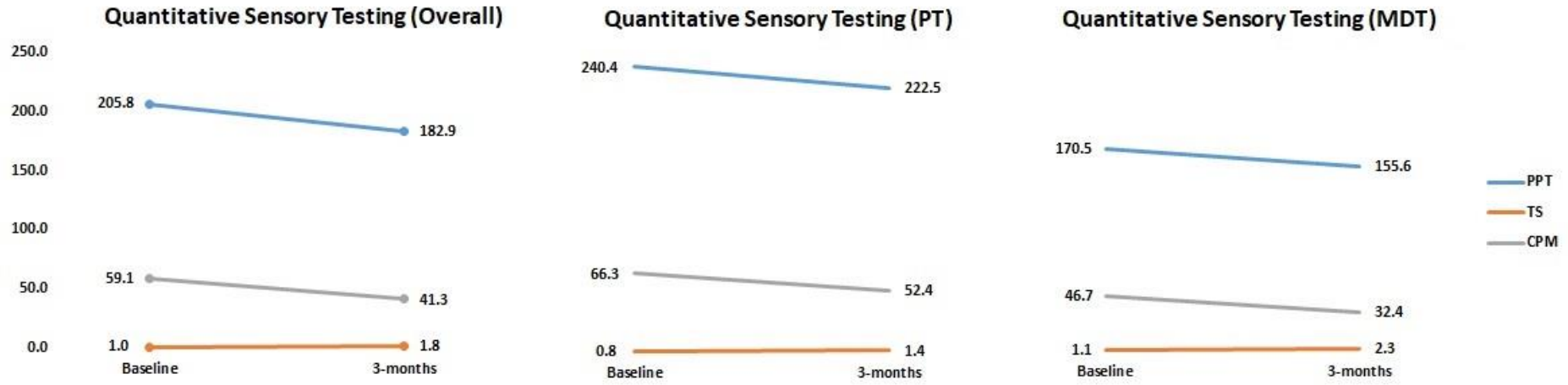

Figure 25. Quantitative Sensory Testing Trends depicting increased sensitivity (decrease in PPT and CPM values, increase in TS values) from baseline to follow-up in the entire population as well as across populations separated according to group intervention programme.

CPM: Conditioned Pain Modulation, **MDT:** Multidisciplinary-lead Group Intervention Programme, **PPT:** Pressure Pain Detection Threshold, **PT:** Physiotherapy-lead group intervention programme, **TS:** Temporal Summation

6.4.3. Prospective unadjusted associations of baseline CS indices or demographic variables with follow-up SM/SC outcomes

Correlation values between baseline age, sex and body mass index and SM/SC outcomes at follow-up are presented in **Table 35**. None of the baseline demographic variables demonstrated statistically significant correlation ($p>0.05$) with any of the self-management outcomes at 3-months follow up except for BMI, which was significantly correlated with follow-up PSEQ ($r=-0.24$, $p<0.04$).

Correlation coefficient values of each pairing of CS indices at baseline with SM/SC outcomes at follow-up are presented in **Table 36**. Overall, from the QST modalities, baseline PPT demonstrated significant positive correlation with follow-up HEIQ-PEL ($r=0.23$, $p=0.03$), HEIQ-SIS ($r=0.28$, $p<0.01$) and negative correlation with HEIQ-ED ($r=-0.21$, $p=0.05$). Baseline TS did not correlate significantly ($r=-0.11$ to 0.21, $p>0.05$) with SM/SC outcomes at follow-up and baseline CPM demonstrated significant positive correlation with HEIQ-PEL ($r=0.31$, $p=0.01$) and negative correlation with HEIQ-ED ($r=-0.29$, $p=0.01$). From the body manikin indices, >9/24 classification method displayed a statistically significant negative correlation with HEIQ-PEL ($r=-0.32$, $p<0.01$) and a positive correlation with HCUQ ($r=0.35$, $p<0.01$) whereas the ACR classification method displayed a significant negative correlation only with HEIQ-PEL ($r=-0.22$, $p=0.05$). The Central Mechanisms trait displayed statistically significant negative correlations with HEIQ-HDB ($r=-0.25$, $p<0.03$), HEIQ-PEL ($r=-0.54$, <0.01), HEIQ-CAA ($r=-0.51$, $p<0.01$), HEIQ-SIS ($r=-0.37$, $p<0.01$), and PSEQ ($r=-0.56$, <0.01) as well as positive correlations with HEIQ-ED ($r=0.54$, <0.01) and HCUQ ($r=0.46$, $p<0.01$).

Table 35. Correlation of baseline demographic variables with self-management/self-care constructs at 3-months follow-up.

	Sample Size n=87	Age		Sex		BMI	
		r	**p-value**	**r**	**p-value**	**r**	**p-value**
Self-management/Self-care at 3-months follow-up	**HEIQ_HDB**	-0.05	0.66	-0.08	0.49	-0.03	0.79
	HEIQ-PEL	0.15	0.20	-0.14	0.22	-0.12	0.32
	HEIQ-SMI	-0.13	0.25	-0.10	0.43	0.08	0.51
	HEIQ-CAA	0.16	0.17	-0.04	0.76	-0.22	0.06
	HEIQ-STA	-0.18	0.13	0.03	0.83	0.07	0.55
	HEIQ-SIS	0.15	0.20	-0.17	0.16	-0.04	0.73
	HEIQ-HSN	0.10	0.38	-0.19	0.10	0.06	0.60
	HEIQ-ED	-0.13	0.27	0.11	0.34	0.09	0.46
	PSEQ	0.14	0.25	-0.08	0.49	**-0.24**	**0.04**
	HCUQ	-0.03	0.78	0.19	0.10	0.13	0.28

BMI: Body Mass Index, **Cor:** Pearson or Spearman Correlation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PSEQ:** Pain Self-Efficacy Questionnaire

Values in **bold** indicate statistical significance ($p<0.05$), Yellow colour indicates correlation of $r<0.50$

Table 36. Longitudinal unadjusted correlations between CS-related indices with SM/SC outcomes at baseline and at 3-months follow-up.

	Sample Size	PPT		TS		CPM		ACR		>9/24		CMT	
	n=87	Cor	p-value	Cor	p-value	Cor	p-value	Cor	p-value	Cor	p-value	Cor	p-value
Self-management / Self-care	**HEIQ-HDB**	0.12	0.27	0.09	0.46	0.21	0.07	-0.16	0.16	-0.16	0.17	**-0.25**	**0.03**
	HEIQ-PEL	**0.23**	**0.03**	-0.02	0.87	**0.31**	**0.01**	**-0.22**	**0.05**	**-0.32**	**<0.01**	**-0.54**	**<0.01**
	HEIQ-SMI	0.19	0.08	-0.05	0.70	0.09	0.45	0.16	0.16	-0.04	0.76	0.02	0.82
	HEIQ-CAA	0.11	0.29	-0.11	0.37	0.13	0.28	-0.10	0.38	-0.20	0.08	**-0.51**	**<0.01**
	HEIQ-STA	0.12	0.25	0.02	0.87	0.17	0.14	0.09	0.44	-0.04	0.76	-0.05	0.67
	HEIQ-SIS	**0.28**	**<0.01**	-0.07	0.60	0.11	0.34	-0.05	0.65	-0.16	0.17	**-0.37**	**<0.01**
	HEIQ-HSN	0.09	0.39	0.02	0.87	0.15	0.20	-0.10	0.38	-0.18	0.12	-0.21	0.06
	HEIQ-ED	**-0.21**	**0.05**	0.21	0.07	**-0.29**	**0.01**	0.17	0.16	0.13	0.24	**0.54**	**<0.01**
	PSEQ	0.18	0.09	0.05	0.70	0.12	0.29	-0.01	0.94	-0.19	0.10	**-0.56**	**<0.01**
	HCUQ	-0.13	0.24	-0.10	0.44	0.04	0.73	0.17	0.14	**0.35**	**<0.01**	**0.46**	**<0.01**
Pain	**NRS**	-0.18	0.10	0.08	0.46	-0.10	0.39	0.19	0.08	0.20	0.07	**0.34**	**<0.01**
	PDETECT	-0.19	0.09	0.02	0.84	0.01	0.98	**0.27**	**0.01**	**0.27**	**0.01**	**0.57**	**<0.01**
	Now	-0.15	0.19	0.02	0.89	-0.17	0.13	0.13	0.25	0.16	0.14	**0.31**	**<0.01**
	Strongest	-0.05	0.67	0.17	0.10	-0.15	0.17	**0.27**	**0.01**	**0.26**	**0.02**	**0.34**	**<0.01**
	Average	-0.14	0.19	0.05	0.62	**-0.21**	**0.05**	**0.23**	**0.04**	0.19	0.08	**0.37**	**<0.01**
Negative Affect	**HADS Anx.**	-0.18	0.11	0.02	0.88	-0.04	0.71	0.12	0.29	0.15	0.18	**0.66**	**<0.01**
	HADS Dep.	-0.20	0.07	0.06	0.57	-0.06	0.59	0.15	0.18	**0.24**	**0.03**	**0.70**	**<0.01**
	PCS	**-0.30**	**0.01**	0.16	0.14	-0.17	0.13	0.18	0.10	0.14	0.19	**0.52**	**<0.01**
	TSK	-0.15	0.19	0.06	0.61	-0.14	0.22	-0.06	0.60	0.01	0.94	**0.29**	**0.01**
Limiting Factors	**RMDQ**	-0.17	0.12	0.03	0.75	-0.17	0.12	0.12	0.26	**0.27**	**0.01**	**0.44**	**<0.01**
	FSS	-0.06	0.60	0.01	0.97	0.01	0.91	0.12	0.28	0.16	0.16	**0.39**	**<0.01**
	FSVAS	-0.06	0.60	-0.04	0.71	-0.12	0.28	-0.10	0.38	-0.12	0.28	**0.39**	**<0.01**
	FMSS	**-0.23**	**0.03**	0.04	0.73	-0.05	0.68	**0.37**	**<0.01**	**0.37**	**<0.01**	**0.63**	**<0.01**
Quality of Life	**EQ-5D-5L**	0.17	0.12	0.03	0.79	0.16	0.16	-0.17	0.12	**-0.26**	**0.01**	**-0.52**	**<0.01**
	$EQ1_{Mobility}$	-0.14	0.24	0.07	0.55	-0.08	0.49	0.09	0.39	**0.25**	**0.02**	**0.26**	**0.02**
	$EQ2_{Self\text{-}care}$	-0.10	0.34	0.01	0.96	-0.14	0.22	0.09	0.42	**0.31**	**<0.01**	**0.45**	**<0.01**
	$EQ3_{Activities}$	-0.10	0.34	-0.01	0.90	-0.08	0.47	0.17	0.12	**0.27**	**0.01**	**0.39**	**<0.01**
	$EQ4_{Discomfort}$	-0.14	0.24	-0.03	0.81	-0.14	0.21	0.14	0.21	0.20	0.07	**0.30**	**0.01**
	$EQ5_{Anx/Dep}$	**-0.25**	**0.03**	0.09	0.38	-0.09	0.41	0.13	0.26	0.16	0.14	**0.63**	**<0.01**
	$EQVAS_{Health}$	0.21	0.06	0.02	0.85	0.16	0.15	-0.16	0.15	-0.13	0.24	**-0.46**	**<0.01**
Risk	**STarT Back**	-0.14	0.21	0.04	0.72	-0.13	0.24	**0.27**	**0.01**	**0.25**	**0.02**	**0.51**	**<0.01**

ACR: American College of Rheumatology, **CMT:** Central Mechanisms Trait, **Cor:** Pearson or Spearman Correlation, **CPM**: Conditioned Pain Modulation, **CS:** Central Sensitisation, **EQ-5D-5L:** Quality of Life Instrument, **EQVAS:** Quality of Life Visual Analogue Scale, **FMSS:** Fibromyalgia Severity Scale, **FSS:** Fatigue Severity Scale, **FSVAS:** Fatigue Visual Analogue Scale, **HADS:** Hospital Anxiety & Depression Scale, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **NRS:** Pain Numerical Rating Scale, **PCS:** Pain Catastrophising Scale, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **RMDQ:** Roland-Morris Disability Questionnaire, **SM/SC:** Self-management/Self-care, **STarT-Back:** Stratification tool, **TS:** Temporal Summation, **TSK:** Tampa Scale of Kinesiophobia, **>9/24:** 24-site body manikin classification method

Values in **bold** indicate statistical significance (p<0.05), Yellow colour indicates correlation of r<0.50, Green colour indicates correlation of r>0.50

Correlation values between baseline CS indices and pain, disability and neuropathic pain as well as other secondary outcomes at follow-up are also presented in **Table 36**. PPT demonstrated significant negative correlations with PCS (r=-0.30, p=0.01), FMSS (r=-0.23, p=0.03) and the emotional distress dimension (EQ5) of the EQ-5D-5L tool (r=-0.25, p=0.03). TS did not correlate significantly (r=-0.04 to 0.17, p>0.05) with any of the secondary outcomes and CPM demonstrated a significant negative correlation only with average NRS pain intensity (r=-0.21, p=0.05). The ACR and >9/24 classification methods correlated significantly with neuropathic-like pain (r=0.27, p<0.01) and the FMSS (r=0.37, p<0.01). Both pain distribution classification methods correlated significantly with the STarT-Back (ACR: r=0.27, p=0.01; >9/29: r=0.25, p=0.02) whereas >9/24 alone displayed significant positive correlations with disability (r=0.27, p=0.01) and depression (r=0.24, p=0.03) as well as a negative correlation with EQ-5D-5L (r=-0.26, r=0.01). The Central Mechanisms trait demonstrated significant correlations (p<0.05) with each secondary outcome (r=-0.52 to 0.70). The FMSS significantly correlated (p<0.05) with all CS indices (r=-0.26 to 0.80) except TS and CPM.

Correlation values between baseline secondary psychological variables and 3-months SM/SC outcomes are presented in **Table 37**. Pain severity (NRS), neuropathic-like pain (painDETECT), negative affect (HADS-A, HADS-D, PCS and TSK) and limited factors (RMDQ, FSS and FMSS) demonstrated significant correlations with all SM/SC outcomes (r=-0.73 to 0.77) except HEIQ-SMI and HEIQ-STA. EQ-5D-5L and STarT Back were significantly correlated (p<0.05) with all follow-up SM/SC outcomes (r=-0.67 to 0.72).

Correlation values between baseline and follow-up SM/SC outcomes as well as between each secondary variable aimed to be used in multivariable regression models are presented in **Appendix 10**. All baseline SM/SC constructs were found to be significantly correlated with their follow-up counterparts (r=0.34 to 0.72, p<0.01) apart from HCUQ (r=0.16, p>0.05).

Table 37. Longitudinal unadjusted correlations between other baseline characteristics and SM/SC constructs at 3 months follow-up.

Sample Size n=87		Self-Management Domains (HEIQ)								Self-Care	
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		Cor	Cor	Cor	Cor	Cor	Cor	Cor	Cor	Cor	Cor
Pain	**NRS**	**-0.34*** **	**-0.44*** **	-0.11	**-0.33*** **	-0,20	**-0,22***	**-0.23***	**0.51*** **	**-0.50*** **	**0.31*** **
	PDETECT	**-0.30** **	**-0.40*** **	0.05	**-0.44*** **	-0.05	**-0.34*** **	-0.20	**0.48*** **	**-0.50*** **	**0.46*** **
	Now	**-0.29** **	**-0.45*** **	-0.14	**-0.32*** **	-0.19	-0.17	**-0.29** **	**0.52*** **	**-0.47*** **	**0.32*** **
	Strongest	**-0.30** **	**-0.39*** **	-0.02	**-0.38*** **	-0.17	**-0.22***	-0.17	**0.44*** **	**-0.36*** **	**0.27*** **
	Average	**-0.38*** **	**-0.46*** **	-0.11	**-0.37*** **	-0.15	**-0.24***	-0.21	**0.52*** **	**-0.40*** **	**0.25***
Negative Affect	**HADS Anx.**	-0.21	**-0.53*** **	-0.12	**-0.61*** **	-0.09	**-0.43*** **	**-0.31*** **	**0.69*** **	**-0.61*** **	**0.34*** **
	HADS Dep.	**-0.25***	**-0.65*** **	-0.05	**-0.73*** **	-0.14	**-0.40*** **	**-0.27** **	**0.72*** **	**-0.73*** **	**0.46*** **
	PCS	**-0.33*** **	**-0.60*** **	**-0.24***	**-0.63*** **	**-0.26***	**-0.43*** **	**-0.29** **	**0.77*** **	**-0.60*** **	**0.24***
	TSK	**-0.26***	**-0.42*** **	-0.18	**-0.53*** **	-0.20	**-0.26***	-0.21	**0.59*** **	**-0.53*** **	**0.23***
Limiting Factors	**RMDQ**	**-0.40*** **	**-0.52*** **	-0.12	**-0.59*** **	-0.18	**-0.25***	-0.19	**0.65*** **	**-0.71*** **	**0.49*** **
	FSS	-0.20	**-0.32*** **	0.18	**-0.31*** **	0.04	-0.20	-0.15	**0.36*** **	**-0.38*** **	**0.42*** **
	FSVAS	**0.29** **	**0.37*** **	-0.15	**0.45*** **	0.01	0.19	0.07	**-0.46*** **	**0.44*** **	**-0.36*** **
	FMSS	**-0.30** **	**-0.45*** **	-0.06	**-0.31*** **	-0.12	**-0.28** **	**-0.26***	**0.50*** **	**-0.51*** **	**0.44*** **
Quality of Life	**EQ-5D-5L**	**0.43*** **	**0.67*** **	**0.26***	**0.66*** **	**0.34*** **	**0.43*** **	**0.41*** **	**-0.67*** **	**0.69*** **	**-0.47*** **
	EQ1$_{Mobility}$	**-0.38*** **	**-0.44*** **	**-0.29** **	**-0.50*** **	**-0.42*** **	**-0.32*** **	**-0.28** **	**0.45*** **	**-0.58*** **	**0.37*** **
	EQ2$_{Self-care}$	-0.21	**-0.55*** **	**-0.25***	**-0.58*** **	**-0.25***	**-0.38*** **	**-0.46*** **	**0.49*** **	**-0.53*** **	**0.41*** **
	EQ3$_{Activities}$	**-0.40*** **	**-0.55*** **	-0.18	**-0.65*** **	**-0.38*** **	**-0.33*** **	**-0.25***	**0.51*** **	**-0.67*** **	**0.39*** **
	EQ4$_{Discomfort}$	**-0.37*** **	**-0.41*** **	**-0.22***	**-0.42*** **	**-0.29** **	**-0.24***	**-0.29** **	**0.48*** **	**-0.52*** **	**0.36*** **
	EQ5$_{Anx/Dep}$	**-0.26*** **	**-0.58*** **	-0.12	**-0.61*** **	-0.12	**-0.42*** **	**-0.25***	**0.69*** **	**-0.59*** **	**0.33*** **
	EQVAS$_{Health}$	**0.48*** **	**0.52*** **	0.12	**0.54*** **	**0.24***	**0.36***	**0.35*** **	**-0.60*** **	**0.57*** **	**-0.35*** **
Risk	**STarT-Back**	**-0.47*** **	**-0.59*** **	**-0.22***	**-0.60*** **	**-0.27***	**-0.39*** **	**-0.28** **	**0.72*** **	**-0.63*** **	**0.30** **

Cor: Pearson or Spearman Correlation, **FMSS:** Fibromyalgia Severity Scale, **FSS:** Fatigue Severity Scale, **FSVAS:** Fatigue Severity Visual Analogue Scale, **EQ-5D-5L:** Quality of Life Instrument, **EQVAS:** Quality of Life Visual Analogue Scale, **HADS:** Hospital Anxiety & Depression Scale, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **NRS:** Pain Numerical Rating Scale, **PCS:** Pain Catastrophising Scale, **PSEQ:** Pain Self-Efficacy Questionnaire, **RMDQ:** Roland-Morris Disability Questionnaire, **STarT-Back:** Stratification tool, **TSK:** Tampa Scale of Kinesiophobia * ≤0.05, ** <0.01, ***<0.001. Values in **bold** indicate statistical significance (p<0.05).
Yellow colour indicates correlation of r<0.50, Green colour indicates correlation of r>0.50

6.4.4. Prospective associations of baseline CS indices with follow-up SM/SC outcomes adjusted for other factors

Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating no significant multicollinearity between all independent variables. All examined models demonstrated normal distribution (Shapiro-Wilk $p>0.05$) of their residuals allowing for implementation of linear regression modelling.

Details of the multivariable regression models between the different QST modalities and SM/SC outcomes adjusted for baseline age, sex, pain, depression, catastrophising and fatigue scores are provided in **Table 38**. The association between PPT and HEIQ-SIS as well as between CPM and HEIQ-PEL or CPM and HEIQ-ED demonstrated in bivariate analyses remained significant in models where those associations were explored when baseline age, sex, pain, depression, catastrophising and fatigue were also taken into account. Associations of PPT with HEIQ-PEL and HEIQ-ED demonstrated in bivariate models lost statistical significance in multivariable models adjusted for the same factors whereas TS displayed a significant association with HEIQ-ED in a multivariable model that was not evident in bivariate analyses.

Details of the regression models between the other CS indices and SM/SC outcomes, adjusted for baseline age, sex, pain, depression, catastrophising and fatigue scores together within each model are provided in **Table 39**. The ACR and >9/24 widespread pain classification methods did not maintain their significant bivariate associations with HEIQ-PEL or HCUQ in multivariable models whereas the single Central Mechanisms trait, in multivariable models which all included age, sex, pain, and quality of life scores, retained its significant association only with HEIQ-CAA, HEIQ-ED and PSEQ.

Table 38. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Quantitative Sensory Testing	PPT †	0.03	0.001	0.08	0.03	0.08	**0.81***	0.007	-0.08	1.44	-0.003
	Age	-0.003	-0.001	-0.01	-0.004	**-0.03****	0.01	0.001	0.003	-0.07	0.04
	Sex	-0.08	-0.001	-0.13	0.04	0.06	-0.16	-0.06	0.04	-0.47	1.31
	Pain	-0.03	-0.001	0.02	-0.002	-0.001	-0.13	-0.001	0.02	-0.56	**0.27***
	Depression	-0.03	**-0.001****	-0.05	**-0.11*****	**-0.28*****	**-0.50****	-0.01	**0.10****	**-2.48*****	**0.53****
	Catastrophisation	0.002	-0.001	-0.02	-0.02	**0.13***	0.10	-0.003	**0.07***	-0.53	0.08
	Fatigue	-0.001	-0.001	**0.002***	-0.001	0.004	0.004	0.001	-0.001	0.01	0.003
	TS †	1.31	-0.001	-0.77	-1.11	-0.24	-4.59	0.04	**2.20***	14.04	-10.79
	Age	-0.01	-0.001	-0.009	-0.003	**-0.04****	-0.001	0.001	0.003	-0.11	0.05
	Sex	-0.12	-0.001	-0.18	0.04	0.004	-0.71	-0.07	0.07	-1.79	**1.48***
	Pain	-0.03	-0.001	0.02	-0.001	0.001	-0.12	-0.001	0.02	-0.56	**0.28****
	Depression	-0.03	**-0.001****	-0.06	**-0.11*****	**-0.28*****	**-0.52****	-0.01	**0.11****	**-2.38*****	**0.47***
	Catastrophisation	-0.008	-0.001	-0.02	-0.02	0.12	0.05	-0.004	**0.07***	-0.75	0.14
	Fatigue	-0.001	-0.001	**0.002***	-0.001	0.003	0.004	0.001	-0.001	0.01	0.004
	CPM †	0.001	**0.001***	0.001	0.001	0.002	0.001	0.001	**-0.002***	0.01	0.01
	Age	-0.004	-0.001	-0.01	-0.004	**-0.04****	-0.004	0.001	0.004	-0.09	0.05
	Sex	-0.08	-0.001	-0.19	0.03	0.03	-0.76	-0.06	0.07	-1.31	1.44
	Pain	-0.03	-0.001	0.02	-0.002	-0.001	-0.12	-0.001	0.02	-0.55	**0.26***
	Depression	-0.03	**-0.001****	-0.05	**-0.11*****	**-0.27*****	**-0.49****	-0.01	**0.09***	**-2.43*****	**0.55****
	Catastrophisation	-0.002	-0.001	-0.03	-0.03	0.12	0.03	-0.004	**0.08****	-0.69	0.07
	Fatigue	0.001	-0.001	**0.003***	-0.001	**0.004***	0.004	0.001	-0.001	0.01	0.003

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Table 39. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Body Manikin	ACR †	-0.09	-0.001	0.19	0.05	0.36	0.74	0.001	0.03	4.62	-0.45
	Age	-0.004	-0.001	-0.01	-0.004	**-0.04****	-0.004	0.001	0.004	-0.09	0.04
	Sex	-0.08	-0.001	-0.25	0.01	-0.10	-0.98	-0.07	0.10	-2.88	1.44
	Pain	-0.03	-0.001	0.02	-0.002	-0.006	-0.14	-0.001	0.02	-0.64	**0.28****
	Depression	-0.03	**-0.001****	-0.05	**-0.11*****	**-0.28*****	**-0.51****	-0.02	**0.09****	**-2.55*****	**0.54****
	Catastrophisation	-0.001	-0.001	-0.02	-0.03	0.12	0.03	-0.004	**0.08***	-0.66	0.08
	Fatigue	0.001	-0.001	**0.002***	-0.001	0.003	0.003	0.001	-0.001	0.004	0.003
	>9/24 †	-0.04	-0.001	-0.03	0.02	0.17	0.32	-0.02	-0.13	0.41	0.79
	Age	-0.004	-0.001	-0.01	-0.004	**-0.04****	-0.01	0.001	0.005	-0.10	0.04
	Sex	-0.10	-0.001	-0.19	0.02	-0.04	-0.85	-0.06	0.13	-1.67	1.13
	Pain	-0.03	-0.001	0.02	-0.002	-0.001	-0.13	-0.001	0.02	-0.55	**0.26***
	Depression	-0.03	**-0.001****	-0.05	**-0.11*****	**-0.28*****	**-0.51****	-0.01	**0.10****	**-2.49*****	**0.49***
	Catastrophisation	-0.001	-0.001	-0.02	-0.03	0.12	0.03	-0.004	**0.08***	-0.66	0.08
	Fatigue	-0.001	-0.001	**0.003***	-0.001	0.003	0.003	0.001	-0.001	0.008	0.002
Trait Score	CMT †	-0.10	-0.002	0.16	**-0.32***	0.30	-0.93	0.04	**0.35***	**-6.61***	1.15
	Age	-0.003	0.001	-0.004	-0.001	-0.02	0.02	0.001	0.001	0.001	0.02
	Sex	-0.08	-0.001	-0.17	0.09	0.004	-0.50	-0.07	0.10	0.39	1.07
	Pain	-0.02	-0.001	0.03	0.01	0.02	-0.10	0.004	-0.002	-0.07	0.20
	Quality of Life	0.28	**0.006****	0.74	**0.68***	**1.57***	1.30	**0.30****	**-0.92***	**18.31****	-3.91

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a QST modality and an SM/SC outcome of β<0.50, Green colour indicates the association between a QST modality and an SM/SC outcome of β>0.50

Associations of non-CS variables with SM/SC in multivariable models are also given in **Table 38** and **Table 39**. Baseline depression demonstrated significant independent associations with follow-up SM/SC outcomes across most models regardless of CS indices used as an independent variable. Catastrophisation was significantly associated with HEIQ-ED across models and with HEIQ-STA when PPT was also an independent variable. Fatigue demonstrated significant positive associations only with HEIQ-SMI and pain significant negative associations only with HCUQ. In models where the Central Mechanisms trait was included as an independent variable, quality of life yielded significant associations with HEIQ-PEL, HEIQ-CAA, HEIQ-STA, HEIQ-HSN, HEIQ-ED and PSEQ respectively. HCUQ was universally associated only with pain and depression. Sex did not demonstrate significant independent associations with any SM/SC outcome whereas age was found to be significantly associated only with HEIQ-STA in all regression models and across CS indices.

Exploratory secondary analyses of regression models assessing the relationship between baseline CS indices and follow-up SM/SC outcomes adjusted for the baseline score of that SM/SC outcome, demonstrated similar significant associations of PPT, TS and the Central Mechanisms trait with HEIQ-SIS, HEIQ-ED and HEIQ-CAA respectively as in primary regression analyses (**Table 40** and **Table 41**). Adjusting for baseline SM/SC scores revealed a significant association of the Central Mechanisms trait with HEIQ-STA, which did not reach statistical significance in bivariate analysis. Associations of CPM with HEIQ-PEL or HEIQ-ED lost statistical significance after adjustment for baseline SC/SM scores. In such analyses, depression was significantly associated only with HEIQ-STA, HEIQ-ED and HCUQ while fatigue revealed a significant positive association with HEIQ-STA not present before.

Table 40. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue as well as the baseline SM/SC score of each dependent SM/SC variable.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Quantitative Sensory Testing	PPT †	0.05	0.001	0.07	0.03	0.13	**0.70***	0.02	-0.09	1.45	0.07
	Baseline SM/SC	**0.43*****	**0.003****	**1.29*****	**0.33****	**1.43*****	**0.69*****	**0.06*****	**0.47****	**0.15*****	0.51
	Age	-0.003	-0.001	-0.01	-0.001	**-0.03****	0.02	0.001	0.008	-0.001	0.06
	Sex	-0.04	0.001	-0.02	0.07	0.22	0.63	0.02	0.02	1.22	**1.65***
	Pain	-0.03	-0.001	0.02	-0.004	0.002	-0.10	-0.003	0.03	-0.09	**0.25***
	Depression	-0.01	-0.001	-0.02	-0.05	**-0.18***	-0.07	-0.001	**0.09****	-0.54	**0.58****
	Catastrophisation	0.004	-0.001	-0.02	-0.03	0.09	-0.11	-0.007	-0.004	0.06	0.03
	Fatigue	0.001	-0.001	**0.003***	-0.001	**0.004***	0.001	0.001	-0.001	0.03	0.004
	TS †	1.30	0.004	0.29	-0.61	0.53	-2.36	0.32	2.05*	8.48	-10.70
	Baseline SM/SC	**0.43*****	**0.003****	**1.33*****	**0.31****	**1.40****	**0.71*****	**0.06*****	**0.44****	**15.31*****	0.50
	Age	-0.004	-0.001	-0.007	-0.002	**-0.03****	0.007	0.001	0.008	-0.03	**0.07***
	Sex	-0.01	-0.001	-0.07	0.05	0.11	0.15	-0.003	0.06	-0.02	**1.75***
	Pain	-0.03	-0.001	0.01	-0.002	0.004	-0.09	-0.003	0.03	-0.09	**0.26***
	Depression	0.007	-0.001	-0.01	-0.06	**-0.17***	-0.07	0.001	**0.10****	-0.48	**0.51***
	Catastrophisation	-0.02	-0.001	-0.04	-0.03	0.07	-0.18	-0.01	-0.004	-0.13	0.08
	Fatigue	0.001	-0.001	**0.002***	-0.001	**0.004***	0.001	0.001	-0.001	0.03	0.01
	CPM †	0.001	0.001	0.001	-0.001	0.002	0.001	0.001	-0.001	-0.002	0.006
	Baseline SM/SC	**0.42*****	**0.002****	**1.31*****	**0.33****	**1.40****	**0.72*****	**0.06*****	**0.42****	**15.47*****	0.46
	Age	-0.003	-0.001	-0.01	-0.002	**-0.03****	0.005	0.001	0.008	-0.02	0.06
	Sex	0.03	-0.001	-0.06	0.04	0.16	0.12	0.001	0.07	0.09	**1.68***
	Pain	-0.03	-0.001	0.01	-0.003	0.003	-0.09	-0.002	0.03	-0.07	**0.25***
	Depression	0.003	-0.001	-0.02	-0.05	**-0.17***	-0.05	-0.001	**0.08***	-0.52	**0.59****
	Catastrophisation	-0.01	-0.001	-0.04	-0.04	0.07	-0.19	-0.009	0.01	-0.07	0.02
	Fatigue	0.001	0.001	**0.002***	-0.001	**0.004***	0.001	0.001	-0.001	0.03	0.004

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (baseline scores of each SM/SC outcome, age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Table 41. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue as well as the baseline SM/SC score of each dependent SM/SC variable.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Body Manikin	ACR †	-0.16	-0.001	0.10	0.02	0.29	0.52	0.01	-0.02	3.80	-0.52
	Baseline SM/SC	**0.44*** **	**0.003** **	**1.27*** **	**0.32** **	**1.35** **	**0.71*** **	**0.06*** **	**0.46** **	**15.04*** **	0.52
	Age	-0.03	-0.001	-0.007	-0.002	**-0.03** **	0.005	0.001	0.009	-0.02	0.06
	Sex	0.06	-0.001	-0.10	0.04	0.03	-0.04	-0.004	0.10	-1.0	**1.74***
	Pain	-0.02	-0.001	0.01	-0.004	-0.002	-0.10	-0.003	0.03	-0.16	**0.26***
	Depression	0.003	-0.001	-0.02	-0.05	**-0.18***	-0.06	-0.002	**0.09***	-0.63	**0.59** **
	Catastrophisation	-0.008	-0.001	-0.04	-0.04	0.08	-0.18	-0.009	0.006	-0.09	0.02
	Fatigue	0.001	-0.001	**0.002***	-0.001	0.003	-0.001	0.001	-0.001	0.02	0.004
	>9/24 †	-0.14	-0.001	-0.11	-0.03	0.006	-0.005	-0.06	-0.11	-0.20	0.71
	Baseline SM/SC	**0.44*** **	**0.003** **	**1.35*** **	**0.33** **	**1.39** **	**0.72*** **	**0.06*** **	**0.45** **	**15.42*** **	0.48
	Age	-0.003	-0.001	-0.006	-0.002	**-0.03** **	0.005	0.001	0.01	-0.02	0.05
	Sex	0.04	0.001	-0.04	0.05	0.12	0.12	0.02	0.12	0.16	1.41
	Pain	-0.03	-0.001	0.02	-0.003	0.004	-0.09	-0.002	0.03	-0.07	**0.24***
	Depression	0.009	-0.001	-0.009	-0.05	**-0.18***	-0.05	0.002	**0.09** **	-0.51	**0.54***
	Catastrophisation	-0.009	-0.001	-0.04	-0.04	0.08	-0.19	-0.01	0.006	-0.08	0.03
	Fatigue	0.001	0.001	**0.002***	-0.001	**0.004***	0.001	0.001	-0.001	0.03	0.003
Trait Score	CMT †	-0.008	-0.001	0.14	**-0.24***	**0.65***	-0.94	0.03	0.09	-0.76	1.20
	Baseline SM/SC	**0.40*** **	**0.003*** **	**1.34*** **	**0.36*** **	1.88	**0.72*** **	**0.06*** **	**0.42** **	**14.01*** **	0.30
	Age	-0.003	-0.001	-0.003	-0.001	-0.02	0.002	0.001	0.004	0.01	0.03
	Sex	0.02	-0.001	-0.04	0.11	0.11	0.34	-0.003	0.04	0.78	1.24
	Pain	-0.02	-0.001	0.03	0.01	0.03	-0.09	0.002	0.02	0.06	0.19
	Quality of Life	0.20	**0.005***	0.63	0.42	**1.34** **	-0.33	**0.20***	-0.63	4.45	-3.72

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (baseline scores of each SM/SC outcome, age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Details of exploratory regression models where pain was replaced as an independent variable by disability are given in **Table 42** and **Table 43** and details where models featuring disability as well as baseline scores of SM/SC outcomes are given in **Appendix 11** and **Appendix 12**. Regression multivariable models featuring disability instead of pain as an independent variable demonstrated negative associations of similar strength as those seen in primary analyses although, disability was significantly associated with most SM/SC outcomes where pain was not (HEIQ-HDB, HEIQ-PEL, HEIQ-CAA, HEIQ-ED and PSEQ). The inclusion of disability as an independent variable revealed the significant association of the ACR classification method with PSEQ. Addition of baseline scores of each examined dependent variable as independent variables in models already featuring disability did not alter the significant relationship of CS indices with SM/SC outcomes wherever present. Disability retained its significant associations with HEIQ-PEL, HEIQ-CAA, HEIQ-ED and HCUQ.

Table 42. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, disability, depression, catastrophisation and fatigue.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Quantitative Sensory Testing	PPT †	0.01	0.001	0.09	0.01	0.08	**0.79***	0.01	-0.06	0.95	0.16
	Age	-0.004	-0.001	-0.01	-0.002	**-0.03****	0.001	0.001	0.002	-0.05	0.04
	Sex	-0.05	-0.001	-0.15	0.06	0.06	-0.10	-0.06	0.02	0.28	1.04
	Disability	**-0.03***	**-0.001***	0.02	**-0.03****	-0.001	0.01	-0.002	**0.03***	**-0.90*****	**0.26*****
	Depression	-0.03	**-0.001****	-0.05	**-0.09*****	**-0.28*****	**-0.55****	-0.01	**0.09***	**-2.26*****	**0.50***
	Catastrophisation	0.02	0.001	-0.02	-0.001	0.13	0.05	-0.002	0.05	0.07	-0.06
	Fatigue	0.001	0.001	**0.003***	0.001	0.004	0.003	0.001	-0.001	0.02	0.001
	TS †	1.39	0.001	-0.80	-0.92	-0.22	-5.03	0.05	**2.04***	18.27	**-11.58***
	Age	-0.004	-0.001	-0.01	-0.002	**-0.04****	-0.01	0.001	0.002	-0.08	0.05
	Sex	-0.08	-0.001	-0.21	0.06	0.004	-0.63	-0.06	0.04	-0.69	1.08
	Disability	**-0.03***	**-0.001***	0.02	**-0.03****	-0.002	0.005	-0.002	**0.03***	**-0.93*****	**0.27*****
	Depression	-0.02	**-0.001****	-0.06	**-0.10*****	**-0.27*****	**-0.57****	-0.01	**0.10****	**-2.12****	**0.43***
	Catastrophisation	0.01	0.001	-0.02	-0.004	0.12	0.02	-0.002	0.04	-0.09	-0.02
	Fatigue	-0.001	0.001	**0.003***	0.001	0.004	0.003	0.001	-0.001	0.02	0.002
	CPM †	0.001	0.001	0.001	0.001	0.002	0.001	0.001	-0.001	-0.001	**0.01***
	Age	-0.004	-0.001	-0.01	-0.002	**-0.04****	-0.01	0.001	0.003	-0.06	0.04
	Sex	-0.05	-0.001	-0.21	0.05	0.03	-0.68	-0.06	0.04	-0.46	1.11
	Disability	-0.02	**-0.001***	0.02	**-0.03***	0.004	0.001	-0.001	**0.03***	**-0.92*****	**0.30*****
	Depression	-0.03	**-0.001****	-0.05	**-0.09*****	**-0.27*****	**-0.54****	-0.01	**0.08***	**-2.24*****	**0.52****
	Catastrophisation	0.01	-0.001	-0.04	-0.002	0.11	-0.01	-0.004	0.06	0.008	-0.14
	Fatigue	0.001	0.001	**0.003***	0.001	**0.004***	0.003	0.001	-0.001	0.02	0.002

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (age, sex, disability, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Table 43. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, disability, depression, catastrophisation and fatigue.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Body Manikin	ACR †	-0.10	-0.001	0.19	0.07	0.36	0.59	0.002	0.02	**4.70***	-0.36
	Age	-0.004	-0.002	-0.01	-0.002	**-0.04****	-0.01	0.001	0.003	-0.06	0.04
	Sex	-0.04	-0.001	-0.27	0.03	-0.09	-0.85	-0.06	0.06	-1.67	1.01
	Disability	**-0.03***	**-0.001***	0.01	**-0.03****	-0.01	-0.01	-0.002	**0.04****	**-0.95*****	**0.26*****
	Depression	-0.03	**-0.001****	-0.05	**-0.10*****	**-0.28*****	**-0.55****	-0.01	**0.08***	**-2.33*****	**0.51****
	Catastrophisation	0.02	0.001	-0.03	-0.001	0.12	-0.004	-0.002	0.05	0.01	-0.08
	Fatigue	0.001	0.001	**0.002***	0.001	0.003	0.002	0.001	-0.001	0.01	0.001
	>9/24 †	-0.01	-0.001	-0.05	0.07	0.18	0.23	-0.02	-0.18	1.71	0.54
	Age	-0.004	-0.001	-0.01	-0.003	**-0.04****	-0.01	0.001	0.004	-0.07	0.04
	Sex	-0.06	-0.001	-0.21	0.03	-0.04	-0.74	-0.06	0.10	-0.80	0.81
	Disability	**-0.03***	**-0.001***	0.02	**-0.03****	-0.01	-0.01	-0.002	**0.04****	**-0.94*****	**0.24****
	Depression	-0.03	**-0.001****	-0.05	**-0.10*****	**-0.28*****	**-0.55****	-0.01	**0.09****	**-2.33*****	**0.48***
	Catastrophisation	0.02	-0.001	-0.03	0.001	0.13	-0.002	-0.003	0.05	0.03	-0.07
	Fatigue	0.001	0.001	**0.003***	0.001	0.004	0.003	0.001	-0.001	0.02	0.001
Trait Score	CMT †	-0.07	0.001	0.14	-0.23	0.29	-1.12	0.04	0.25	-4.64	0.74
	Age	-0.003	0.001	-0.004	0.002	-0.02	0.01	0.002	-0.001	0.03	0.02
	Sex	-0.05	-0.001	-0.021	0.09	-0.02	-0.42	-0.07	-0.003	0.74	0.81
	Disability	-0.02	-0.001	0.02	**-0.03***	0.01	0.02	0.001	**0.03***	**-0.69****	**0.22****
	Quality of Life	0.23	**0.005***	0.14	0.33	**1.58***	2.15	**0.28***	-0.60	12.01	-3.01

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (age, sex, disability, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Secondary analyses also showed that intervention programme participation demonstrated significant positive associations with follow-up HEIQ-STA and HCUQ and a negative association with follow-up PSEQ. Nevertheless, inclusion of programme participation as a variable did not appear to affect the previously established significant associations of CS indices with SM/SC outcomes (**Appendix 13** and **Appendix 14**).

The explanatory power of each regression model conducted for the purposes of primary and secondary analyses are given in **Appendix 15**. HEIQ-HDB, HEIQ-SMI, HEIQ-STA and HEIQ-HSN demonstrated non-significant ($p>0.05$) explanatory power (Adjusted R^2) in primary analyses but statistically significant in exploratory analyses when the baseline scores of each SM/SC outcome was added in the model ($R^2=0.01$ to 0.35). HEIQ-PEL, HEIQ-CAA, HEIQ-SIS, HEIQ-ED, PSEQ and HCUQ demonstrated statistically significant ($p<0.05$) explanatory power in primary and secondary analyses ($R^2=0.11$ to 0.54).

6.5. Discussion

This study extended the findings from cross-sectional analyses by showing that several baseline indices of central sensitisation were associated with aspects of self-management at 3-months. Baseline depression scores and, to a lesser degree, catastrophising and fatigue, were also found to predict aspects of SM/SC at 3-months follow-up. The number of painful sites on a body manikin, regardless of classification method, appears to be associated with worse SM/SC outcomes in individuals with CLBP undergoing CBT-based group interventions.

Higher initial levels of pain sensitisation were prospectively associated with low scores in HEIQ-PEL, HEIQ-SIS and HEIQ-ED at 3-months both in bivariate correlation and multivariable regression analyses, indicating that CS can predict some aspects of long-term self-management. The Central Mechanisms trait was also associated with most SM/SC outcomes in unadjusted correlations but associations only with HEIQ-CAA, HEIQ-ED and PSEQ remained significant in adjusted models. Apart from TS, all other CS indices seem to be able to predict HEIQ-PEL in unadjusted analyses. The >9/24 classification method appears to be more useful than the ACR method in predicting SM/SC and other outcomes in individuals with CLBP, although both methods appear to lose their predictive capacity once other factors are taken into account.

Baseline depression was longitudinally associated with most domains of SM/SC except HEIQ-SMI and HEIQ-STA in unadjusted as well as HEIQ-HDB and HEIQ-HSN in adjusted correlations, indicating that negative affect is predictive of most but not all aspects of long-term SM/SC in individuals with CLBP. Catastrophising was found to be associated with all SM/SC outcomes in unadjusted correlations however, in adjusted correlations, it was significantly associated only with HEIQ-STA and HEIQ-ED suggesting that maladaptive beliefs can predict only some of the SM/SC outcomes when other factors are taken into account. Increased baseline fatigue was found to be the sole predictor of increased long-term HEIQ-SMI in adjusted analyses despite its negative association with HEIQ-HDB, HEIQ-PEL, HEIQ-CAA, HEIQ-ED, PSEQ and HCUQ in unadjusted correlations.

Analysis of prospectively collected longitudinal data highlighted associations with different constructs to those observed in baseline cross-sectional analyses. HEIQ-SIS measures the impact social engagement and interaction with others can have on individuals as well as the level of confidence people have to seek support from others or from community-based organisations, HEIQ-STA measures the ability of individuals to apply knowledge-based skills and techniques in order to manage more effectively condition-specific symptoms or health-related issues, and HEIQ-ED encapsulates the negative attitude individuals can have towards their condition or their life, characterised by distress, frustration, anger, anxiety, and depression (Osborne et al., 2007). Independent associations of the different baseline CS indices with these 3 outcomes at 3-months, both in models adjusted or not adjusted for baseline scores, implies that CS might be also implicated in the development of long-term self-management as their association is maintained despite initial levels of SM/SC. None of the CS indices measured for the purposes of this study (by contrast disability) seemed to be able to predict or explain poor HEIQ-HSN (confidence people communicate or share information with their health-care providers) or poor HEIQ-HDB (willingness to apply changes to behaviour or diet and explore relaxation as well as disease prevention pathways) (Osborne et al., 2007) suggesting that different aspects of SM/SC are influenced by different biopsychosocial factors.

Demographics such as sex and body mass did not significantly predict or associate with how well or poorly individuals with CLBP will self-manage 3-months after commencing a group intervention programme. Sex was associated with SM/SC in cross sectional analysis, but, in contrast, did not predict SM/SC outcomes at 3 months. No demographic differences were observed with individuals who did not participate in the intervention, which suggests that factors associated with SM/SC in

CLBP populations might therefore be different from those that influence changes during participation in an intervention programme. Age was only associated with HEIQ-STA at 3-months, suggesting that older individuals might be less able to incorporate knowledge-based skills and techniques in order to manage their CLBP more effectively. Older age has been linked with poor performance on tasks that require cognitive control and processing or manipulation of information (Roberts and Allen, 2016).

The findings from prospective analyses further compliment the implementation of different CS indices to explore long-term associations with a multidimensional construct such as self-management as distinct indices were again associated with distinct SM/SC outcomes. Only PPT and CPM were found to be associated with self-management outcomes in unadjusted correlations whereas all three modalities were associated with SM/SC outcomes in adjusted correlations. This suggests that pain mechanisms might be affected by different factors that could all contribute or mediate pain processing and, subsequently, how individuals respond to a highly subjective experience such as pain. Therefore, responses to TS might have been moderated by other factors (e.g. depression and catastrophising), which could explain why a significant prospective association of TS with SM/SC outcomes was demonstrated when all factors were taken into consideration. QST can be affected by the presence of co-morbidities and they need to be accounted for in order to extract accurate information from this type of psychophysical testing (Uddin and MacDermid, 2016). The findings demonstrated associations of QST modalities operating under different theoretical principles ('static' and 'dynamic') with the same SM/SC outcomes that were not evident cross-sectionally (e.g. PPT and CPM with HEIQ-PEL). This suggests that CS, even when manifested "statically" (increased sensitivity to pressure at a distant site) as opposed to just "dynamically" (inefficient wind-up and/or descending inhibition) (Arendt-Nielsen and Yarnitsky, 2009), can predict poor self-management at 3-months.

PPT was associated with positive engagement in life, social integration and support and emotional distress in unadjusted correlations. However, in multivariable models featuring additional independent factors (baseline age, sex, pain, depression, catastrophising and fatigue), it was shown that reduced PPT can predict only poor HEIQ-SIS and not HEIQ-PEL or HEIQ-ED. When baseline levels of HEIQ-SIS are considered, it is shown that CS might be influencing the development of ineffective social integration as well as lack of confidence to seek support from other individuals or organisations. Individuals with CS have demonstrated altered function of the

anterior cingulate cortex, an area essential for affective-emotional aspects of pain, including empathy and social exclusion (Nijs et al., 2015). Neuroimaging studies have shown that brain areas activated during distress caused by social exclusion are the same with those activated during physical pain (Yanagisawa et al., 2011, Eisenberger et al., 2003), resulting to the experience of 'social' pain by excluded individuals (Panksepp, 2003). Those findings suggest that CS could be implicated in the development of generalised feelings of isolation and social exclusion or ill-conceived beliefs about lack of understanding by peers, family members or health-care providers, which could all be reflected in a tool that has been designed to measure such feelings and thoughts (HEIQ-SIS).

Social integration and support was among the SM/SC outcomes that did not show a clinically meaningful change from baseline, which for the distinct HEIQ domains is taken to be of at least half the SD of baseline measurements (Maunsell et al., 2014). PSEQ also did not demonstrate a clinically significant change in the overall population (increase of 5 units) when significance was calculated based on SDs, however, recent CLBP-related literature indicates an increase of 5.5 units as clinically meaningful change (Chiarotto et al., 2016), which highlights that the intervention only marginally did not reach meaningful increase of PSEQ. In terms of pain sensitivity, all QST modalities numerically worsened at 3-months follow-up with PPT appearing to be the most sensitive to change by showing statistically significant change from baseline in the overall population and statistically as well as clinically significant change in the population undertaking the multidisciplinary programme. Nevertheless, there are no standardised cut-off values indicating clinically meaningful changes in relation to pain sensitivity testing although, a gain of 10 kiloPascals using pressure algometry in individuals with knee osteoarthritis is suggestive of meaningful benefits from joint replacement treatment (Graven-Nielsen et al., 2012).

Such evidence hint that current strategies fail to significantly reduce pain sensitivity and are less likely to meaningfully improve SM/SC in general and facilitate social integration in particular. This comes into contrast with the operational paradigm of pain management programs that were developed to reduce pain (Reid et al., 2008) and subsequently CS, which was hypothesised in this study to hinder SM/SC outcomes. Beyond SM/SC, this study demonstrated that, despite the increase in CS levels from baseline to follow-up, programmes managed to promote statistically and clinically significant, albeit marginal, reduction in pain (-1/10) and disability (-3/24). Disability change is consistent with the amount of change considered meaningful in CLBP-related literature (reduction of 2-3 units in RMDQ) (Bombardier et al., 2001) but

change in pain did not reach the cut-off change of at least 30% (reduction of 2-3 units in NRS) (Farrar et al., 2001). Similarly, no clinically significant changes were observed in depression, anxiety, catastrophising and fatigue in the overall population as the outcomes did not reach a reduction of at least half the baseline SD nor a reduction of 1.7 units in the subscales of HADS, 5.1 units in PCS and -15.3 in FSS that have been set in literature to be the minimum clinically meaningful change in those constructs (Lemay et al., 2019, Pettersson et al., 2015, Darnall et al., 2014).

For the purposes of this study, individuals were recruited from group intervention programmes undertaken predominantly within a hospital setting. Traveling to and from the hospital might have increased the physical effort or psychological distress due to commuting needs, which could have contributed to increased CS levels. Programmes operating within a community rather than a hospital-based setting are thought to perform better in facilitating enhanced self-management outcomes (Du et al., 2017, Du et al., 2011). Also, exercise-related activities featured in the programmes might have caused initial short-term pain increase, which could have contributed to unsuccessful reduction of CS or, perhaps, high levels of CS at the start of the program might have limited the benefits patients could receive from paced exercises. High levels of CS have been found to hinder response to exercise regimes in individuals suffering from knee osteoarthritis (O'Leary et al., 2018). Successful decrease of CS by future intervention programmes of similar format could promote further, clinically significant reduction of pain, disability, negative affect, maladaptive beliefs and fatigue, which could allow better SM/SC outcomes in the long-term.

Conditioned Pain Modulation was found to be associated with self-management outcomes in unadjusted and adjusted correlations whereas the role of TS as a predictor of self-management was revealed only in adjusted analyses and only with HEIQ-ED. However, CPM lost its association with follow-up HEIQ-PEL when baseline scores of HEIQ-PEL were taken into account. Nevertheless, associations of CPM with such a construct could indicate that CS negatively affects brain areas (amygdala) that facilitate the development of chronic pain (Simons et al., 2014), cultivate negative treatment expectations (Schmid et al., 2013) and are responsible for the creation of protective pain memories upon certain movements (e.g. exercises) (Zusman, 2004). Such alterations of cerebral function due to increased CS could reinforce the experience of pain and, therefore, limit the ability of individuals to engage with health-promoting activities (HEIQ-PEL). Both CPM and TS are dynamic modalities, designed to measure more complex course of pain processing (Arendt-Nielsen and Yarnitsky, 2009) and both have been previously found to be prospectively associated with

emotional distress linked to a musculoskeletal disorder (Gay et al., 2015, George et al., 2006). This extends the cross-sectional findings and highlights the capacity of dynamic modalities to predict prolonged emotional distress (Georgopoulos et al., 2019). Such findings also emphasise that the experience of pain and the experience of distress, associated with depressive disorders, share similar cerebral processes that can be revealed by dynamic QST modalities such as TS and CPM (Arendt-Nielsen et al., 2018). TS remained significantly associated with long-term emotional distress even when their relationship was adjusted for baseline scores of HEIQ-ED revealing that central pain mechanisms are implicated in the development of prolonged feelings of frustration, anger, anxiety and depression.

The ACR and >9/24 classification methods demonstrated a significant association with HEIQ-PEL in unadjusted analyses showing that the number of painful sites alone can indicate who is more willing to undergo lifestyle changes and adapt life-altering activities. The >9/24 method was correlated with worse health-care utilisation indicating that people with more painful sites seek more care. Prospective analyses compliment further the cross-sectional findings that promoted >9/24 as a more useful to the ACR tool in relation to patients with CLBP. Prospective unadjusted correlation showed that >9/24 could predict long-term depression, pain-related disability and quality of life that the ACR method could not. However, both methods displayed no association with long-term SM/SC outcomes when other factors were accounted for, limiting their predictive utility over other indices. As in cross-sectional analyses, it is possible that the relationship between >9/24 and SM/SC outcomes is mediated by psychological variables such as depression (**Table 22**), meaning that more widely distributed pain (>9/24) is linked with more depression, which is also linked independently with poorer outcomes in most SM/SC domains and, therefore, influencing the association between pain distribution and SM/SC.

The Central Mechanisms trait predicted most future SM/SC outcomes in unadjusted correlations and specifically with HEIQ-CAA, HEIQ-ED and PSEQ when other variables were included in multivariable models. The trait retained its independent association with long-term HEIQ-STA in exploratory models adjusted for baseline HEIQ-STA, complimenting further the findings from cross-sectional analyses, indicating that central pain mechanisms, as an index of CS, are associated with SM/SC outcomes that possibly require adequate cognitive function for their achievement. Chronic pain appears to be associated with cognitive impairment as neural brain systems involved in cognition have been found to be closely linked with pain processing (Moriarty et al., 2011, Villemure and Bushnell, 2009). It is thought that

the presence of chronic pain creates a persistent influx of nociceptive input that competes with other sensory inputs, hindering the division of limited neurophysiological resources in discrete brain regions, causing neuroplastic changes (neural rewiring or reorganisation), and leading, ultimately, to cognitive impairment (Hart et al., 2000, Eccleston and Crombez, 1999).

The influence of CS on the prefrontal cortex (area responsible for the cognitive-evaluative dimensions of pain) might be responsible for poor monitoring of conflicting thoughts towards pain or every-day matters (Verdejo-García et al., 2009). Individuals with CLBP and increased pain sensitivity have been found to display reduced cognitive performance in decision-making tasks (Apkarian et al., 2004) as well as impaired memory and reduced language and communication skills (Jorge et al., 2009). Depression could mediate the relationship between central pain mechanisms and poor cognitive function as it is also cross-sectionally and prospectively associated with HEIQ-STA in unadjusted as well as adjusted correlations but, since it is one of the components used for the derivation of the Central Mechanisms trait, it was not included as an independent factor in regression models featuring that specific CS index. Nevertheless, cognitive dysfunction alone might put individuals into a negative mindset, which could interfere with adaptation of constructive attitudes and approaches towards pain (HEIQ-CAA). Persistent amplified pain has been also linked with inadequate executive function (planning and organisational skills, attentional control, goal-directed behaviour and action-taking) (Solberg Nes et al., 2009, Veldhuijzen et al., 2006, Bosma and Kessels, 2002), which could be a driver of sustained cognitive impairment, manifested in people with CLBP as reduced capacity to use skills and techniques to manage their condition (HEIQ-STA).

The possibility of depression and CS negatively affecting the ability of individuals to effectively self-manage was highlighted in the previous chapter as well as their independent association with some SM/SC outcomes. Depression was also found to be a predictor of almost all SM/SC outcomes at 3-months, a relationship that was also found when other factors such as CS indices, age, sex, catastrophising and fatigue were considered. The predictive capacity of depression for low back pain outcomes is well-demonstrated (George and Beneciuk, 2015, Glombiewski et al., 2010, Arpino et al., 2004, Leino and Magni, 1993). Exploratory analyses considering baseline scores of SM/SC outcomes in multivariable models that included other factors revealed significant associations of baseline depression with follow-up HEIQ-STA, HEIQ-ED and HCUQ. This suggests that depression might be responsible for poor long-term SM/SC outcomes although, the disappearance of significant associations

after adjusting for other factors indicates that some of those factors (e.g. pain) might be moderators rather than confounders in the relationship of depression with SM/SC outcomes.

Emotional distress is thought to be also implicated in the activation of brain centres that drive cognitive-evaluative (prefrontal cortex), emotional (amygdala) and sensory-discriminative (insula) functions (Nijs et al., 2015) and therefore involved in reduced capacity to acquire skills and techniques or to manage distress. Maladaptive thinking such as catastrophising shares similar mechanisms in the brain with negative affect, and that similarity in processing as well as their strong bivariate correlation might be behind the similar unadjusted and adjusted associations found between those two constructs (catastrophising and depression) with SM/SC outcomes. Depression and catastrophisation scores were two of the factors that determined participation to the PT or the MDT group intervention programmes during clinical examination of patients, which explains why individuals with increased depression scores are prone to increased health-care utilisation. Exploratory analyses of regression models adjusted for programme participation did not alter any of the aforementioned associations but showed that group allocation could predict higher HEIQ-STA scores, more health-care utilisation and worse self-efficacy, all of which were shown in analyses to be characteristics of participants to the MDT intervention at 3-months (maximum of 10 treatment sessions, statistically significant increase of HEIQ-STA and lower PSEQ than participants of PT intervention).

As in the previous chapter, exploratory analyses featuring disability instead of pain in regression models did not alter the overlying associations between CS indices and SM/SC outcomes. Even though baseline pain as well as baseline disability predicted the same (and with similar strength) SM/SC outcomes, disability was shown again to be a predictor of most SM/SC outcomes in adjusted analyses whereas pain was not significantly associated with any SM/SC constructs. Such findings further give credence to the point made in the previous chapter regarding the apparent superiority of disability over pain as a predictor of SM/SC outcomes. Disability has been found to be highly correlated with self-efficacy in previous research (Denison et al., 2004). Such an association has been attributed to the influence enactive mastery experiences (“psychological states through which learners organise their own set of beliefs regarding ability from a variety of sources”) hold over self-efficacy, which are thought to be strongly linked with the activities people are capable of doing (disability) as well as with their confidence in performing the necessary behaviours that such activities require (self-efficacy) (Bandura et al., 1999). An alternative explanation for

the demonstrated association could be the overlapping content between the measurement tools used for the three distinct constructs (RMDQ, PSEQ and HEIQ). The three tools require from patients to rate how pain interferes with their ability to perform certain activities (RMDQ), the confidence in performing them (PSEQ) as well as their ability to demonstrate the behaviours such activities require (HEIQ), which individuals might not easily perceive as distinct concepts (Denison et al., 2007). There might also be an underlying neurophysiological explanation as it has been shown that in pain conditions with an easily defined area of pathology such as CLBP, it is perceived disability rather than perceived pain that is linked with parasympathetic activity (Cacioppo et al., 1995). Homeostasis between the sympathetic and parasympathetic systems can be disturbed by psychological triggers, suggesting that it is not the pain but rather how people interpret the painful stimuli (disability) that influences psychophysical experiences and subsequent responses (Gockel et al., 2008). Overall, findings from exploratory models featuring disability instead of pain as a predictor of SM/SC outcomes suggest that perceived disability can indicate and perhaps influence the development of poor self-management at 3-months, highlighting further that it is probably not the pain severity but rather the inability to perform certain activities because of the pain that drives self-management. Future studies need to further examine this theory and validate the potentially causal relationship between disability and poor self-management.

Longitudinal studies are considered superior to those of cross-sectional design as they measure differences of the same individuals across time and therefore any change is less likely to be the result of chance or inter-population cultural and biopsychosocial differences (Shadish et al., 2002). This means that longitudinal studies are ideal to explore multiple exposures and multiple outcomes in a single population as well as to identify predictors, confounders and mediators of a given outcome (Euser et al., 2009). While longitudinal cohort studies address some of the limitations of cross-sectional study design, prospective analysis exploring the association between CS indices and SM/SC outcomes is also subject to several limitations. Different indices were used to detect and measure CS with some of them developed in populations suffering from conditions other than CLBP, which means that the predictive capacity of some of those instruments needs to be viewed with caution when interpreting the results. Self-management is also a diverse concept with no specific tool for its measurement in populations with CLBP. The self-management tool used for the purposes of this study, even though reliable and valid to be used across chronic pathologies, might omit constructs more specific to the self-

management of CLBP. Similarly, no healthcare utilisation measurement tool specific to CLBP exists, which constitutes the tool used in this study subject to biases that could have been addressed by reliability and validity analyses. The small loss to follow up (10.3%) was a strength of the current study, but, even so, the number of individuals with follow-up data (n=87) limits the study power in undertaking such complex regression modelling and therefore, the results need to be viewed with caution as the exploratory models might be subject to over-adjustments. Data on social factors previously shown to be indicative of SM/SC outcomes (Ferrari et al., 2019, Koleck et al., 2006) such as working, marital and family status, job satisfaction, education levels and financial factors such as personal and family income were not collected, which might have inadvertently limited the identification of additional predictors or confounders of poor SM/SC in CLBP. No record of medication side-effects, such as decreased reaction time, cloudy judgment, and drowsiness was kept and, therefore, it is unclear to what degree they have influenced the capacity of individuals to self-manage. Even though analysis was undertaken to account for programme participation, the presence of 'channelling' bias between the two intervention pathways is possible, as programme allocation was conducted by clinicians based on satisfaction of a-priori clinical criteria.

6.6. Conclusion

Different CS indices, collected early in rehabilitation pathways, can predict different aspects of SM/SC, and therefore different SM/SC outcomes, indicating that pain mechanisms such as CS could be used as a component of future clinical prediction tools for self-management. Pain mechanisms appear to be also implicated in the development rather than just the prediction of poor self-management, suggesting that appropriate early reduction of pain sensitisation could allow for better SM/SC outcomes. CS is imposing a cognitive as well as physical and emotional burden on individuals with CLBP, which highlights the need for improved, more diverse treatments that will attempt to address cognitive impairment alongside physical dysfunction and emotional distress. Depression appears to be a strong predictor of future self-management and it appears to be involved in the course of self-management. Higher levels of baseline disability more strongly predict poor self-management than does pain and it appears to hinder the achievement of enhanced self-management outcomes in the long-term. The validity of this finding needs to be further examined in other populations with CLBP. Future research needs to focus on replicating such findings in other chronic pain populations in order to make these

findings generalisable across conditions. Similarly, the ability of CS to predict SM/SC outcomes needs to be ascertained on larger populations suffering from CLBP, that are perhaps under different treatment pathways, in order to be able to generalise across individuals with the same pathology.

7. DISCUSSION

7.1. Overview

This project was driven by the hypothesis that prolonged and disproportionate centrally-driven pain is implicated in the ability of individuals with CLBP to self-manage their condition. The studies encapsulating this endeavour collectively showed that central sensitisation, a clinical phenomenon consistent with altered pain mechanisms, appears to be involved in complex biopsychosocial behaviours such as self-management. The findings indicated that individuals with CS demonstrate poor long-term outcomes in some of the aspects that consist self-management and there are hints that CS might be one of the drivers behind those poor outcomes.

Quantitative Sensory Testing has been proposed as a reliable and valid collection of discrete clinical tests for the successful detection and measurement of CS, particularly in relation to conditions within the musculoskeletal spectrum (Fernandes et al., 2019, Kennedy et al., 2016, Suokas et al., 2012). QST in general, and distinct modalities in particular (PPT, TS, CPM), were found, through a systematic literature review, to predict long-term musculoskeletal outcomes (pain, disability and negative affect) but no available evidence indicated its capacity to predict self-management and/or self-care outcomes.

Exploration of the reliability of a combined protocol featuring both static and dynamic modalities, such as PPT (static), TS (dynamic) and CPM (dynamic), designed to target different pain mechanisms established that QST (particularly PPT and TS) is a reliable tool across raters and study populations. It was found that such a protocol is also valid as discrete modalities and other CS or pain indices are associated with each other, highlighting its sensitivity to assess a multidimensional and complex phenomenon such as pain hypersensitivity.

Central sensitisation, as assessed primarily by QST and on a secondary basis by other indices at baseline was found to demonstrate significant associations with SM/SC outcomes both at a single timepoint as well as prospectively. Such findings provided significant insight on how pain mechanisms might be interfering with self-management and can drive the development of poor long-term outcomes alongside other factors such as depression, catastrophising and disability.

7.2. Novel findings and implications

This study is the first to adapt the novel approach that CS is a multidimensional clinical phenomenon involving biological as well as emotional, social and cognitive aspects. The findings also indicated that prolonged depression alters the function of brain areas related with cognitive control, rumination and somatic/visceral function (Downar and Daskalakis, 2013). Given that depression and chronic pain may share similar brain activation pathways (Sheng et al., 2017), the possibility of CS interfering also with those regions seems likely to be high.

Self-management, as a multidimensional concept, is considered notoriously difficult to measure effectively (Nolte et al., 2013). It entails, among others, the ability to acquire and use skills, solve problems, make meaningful decisions, set achievable goals, form functional relationships, adopt a healthy lifestyle and follow wellbeing guidelines (Lorig and Holman, 2003). The ability to appropriately develop and implement such skills in the long-term management of a chronic condition, most probably, requires adequate physical, emotional, social, and cognitive function. These are also the health domains that are thought to be considerably affected by CS (Bourke et al., 2015).

This project has demonstrated that QST can predict long-term musculoskeletal outcomes (literature review and observational study) and that an elaborate QST protocol, consisted by a combination of ‘static’ and ‘dynamic’ modalities can be reliably used to detect CS (reliability study). Through this project, CS was also shown to predict and possibly promote, poor self-management outcomes related with emotional, cognitive, social and physical dysfunction (observational study). A Central Mechanisms construct, developed by the combination of different self-reported traits shown to be reflecting different aspects of CS (Akin-Akinyosoye et al., 2018) and to predict musculoskeletal outcomes (Akin-Akinyosoye et al., 2020) was also found to predict discrete poor self-management outcomes that were different to those predicted by QST. This simply highlights firstly, the multidimensionality of CS, which can have its different aspects measured by different tools and secondly, that early detection of CS can indicate how effectively individuals will manage their CLBP as well as that it might be implicated in the development of poor self-management from the moment it becomes prevalent.

The findings of this project indicate that CS could interfere with cerebral functions through a complex and potentially lengthy interaction with brain regions that might be

associated with those functions (brain pain neuromatrix). The results suggest that, once detected, CS can provide predictive inference about future SM/SC outcomes. CS seems to alter the function of regions responsible for; impeding individuals from being physically active (amygdala: association between CPM and HEIQ-PEL), misinterpreting the experience of pain and adapting less constructive attitudes towards it (prefrontal cortex: association between the Central Mechanisms trait and HEIQ-CAA), developing aberrant emotional responses (insula: association of TS, CPM and the Central Mechanisms trait with HEIQ-ED), performing poorly on cognitive tasks like acquiring and adequately using skills and techniques for self-management purposes (prefrontal cortex: associations between the Central Mechanisms trait and HEIQ-STA) and experiencing social pain or difficulty to socially integrate and seek support (anterior cingulate cortex: association between PPT and HEIQ-SIS). The findings further suggested that CS, if remained unaltered, perpetuates poor self-management, possibly, through the continuous negative interaction with those brain regions, developing potentially a de-facto causal relationship with poor SM/SC outcomes. The complex interplay ('vicious circle') between CS, emotional distress and increased pain or disability is well established in international research (Arendt-Nielsen et al., 2018, Edwards et al., 2016b, Woolf, 2011) but how CS appears to be at the centre of this circle and how it predicts as well as potentially hinders self-management is the novelty of this study and is illustrated in the model presented in **Figure 26**.

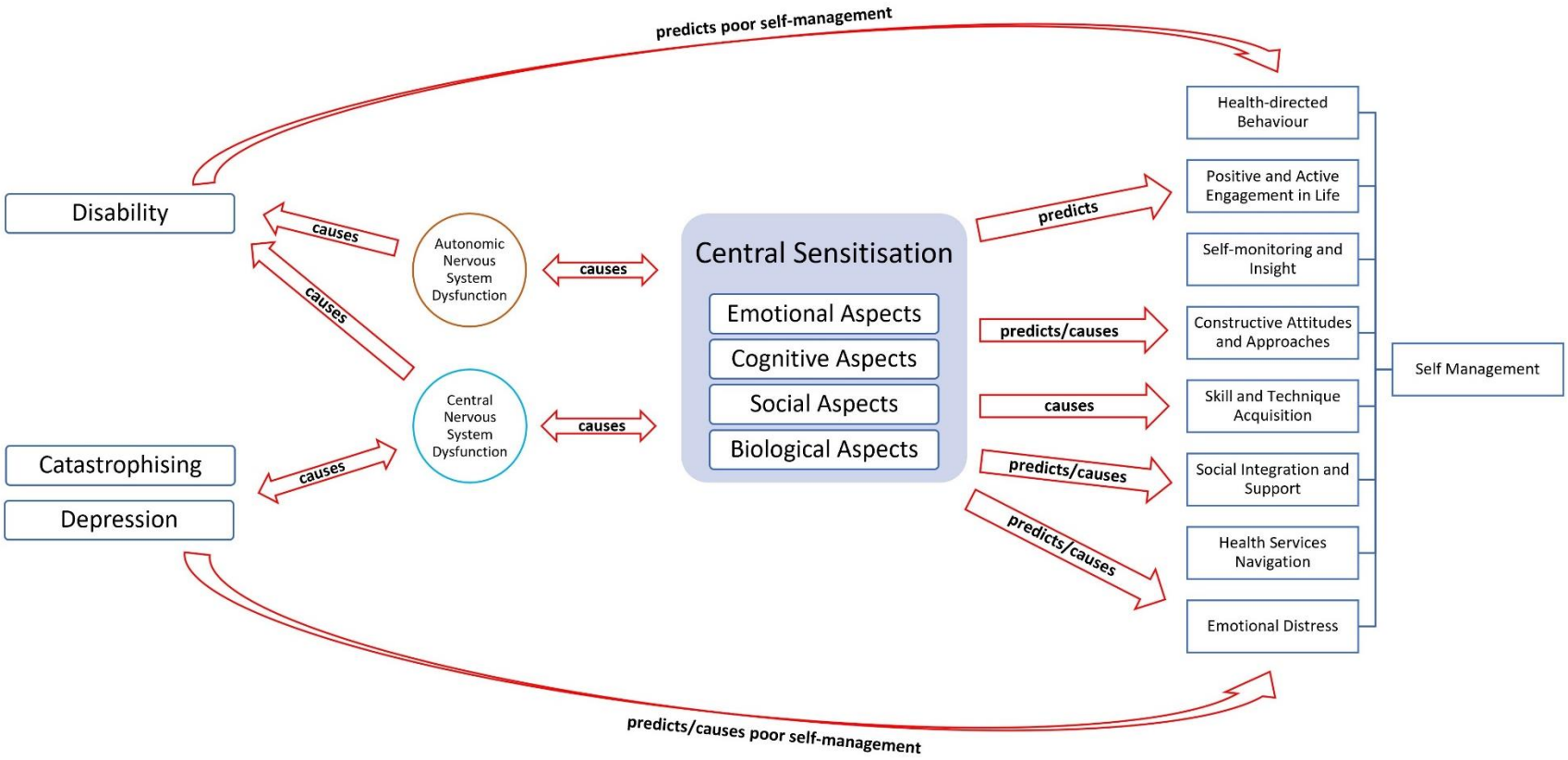

Figure 26. Model demonstrating how central sensitisation might independently predict or facilitate poor self-management and how it might be contributing to the development of additional co-morbidities that could also serve as predictors or facilitators of poor self-management.

The clinical implications of the findings and the relationships illustrated suggest that early detection of CS could facilitate the identification or formation of relevant patient phenotypes. Successful patient stratification could, in turn, streamline clinical decision-making towards treatments (pharmacotherapy with e.g. Serotonin-noradrenaline reuptake inhibitors or Calcium-channel alpha(2)delta ligands), or combination of individualised non-pharmacological approaches (e.g. therapeutic exercise with CBT or pain neuroscience education) that could specifically aim to reduce CS. Potentially, identified individuals with evidence of pain sensitivity could follow a slightly different rehabilitation pathway with more passive treatment (CBT, education) at the early stages of their treatment. People with centrally-driven pain hypersensitivity often need to have their pain levels substantially reduced before being exposed to active interventions such as exercise or manual therapy (Akinci et al., 2016). CS detection should be part of clinical assessment and the contribution of pain mechanisms in the overall ability of individuals to self-manage should be taken into consideration, possibly on an individual level, when deciding the best treatment pathway.

This study built upon existing research and presented means for identifying whether central mechanisms (8 discrete traits) contribute to the experience of pain. Such methods can be used as supplementary, cost-effective tools alongside the use of elaborate equipment (QST). In practical terms, when an individual clinically reports sensitivity to cold/hot in painful areas, panic attacks, tiredness, sleep disturbance, inability to enjoy simple things, concentration problems, irrational focus on the levels of pain and widespread pain in distant areas should make clinicians aware that this individual is probably displaying features of centrally driven pain.

The use of self-reported outcome measures for clinical evaluation is equally relevant to the use of QST and, despite their perceived cost-effectiveness and ease of use, do not preclude additional benefit from the utilisation of QST. Most equipment needed for the application of the protocol featured in this study are either already part of a clinic's inventory (blood pressure cuff) or can be replaced with low-cost alternatives (analogue algometer, monofilaments). QST application within a clinical setting could be used for confirmations of central pain mechanisms involvement made via self-reported characteristics. This study demonstrated that all different CS indices are necessary to evaluate the entire spectrum of centrally-driven pain sensitivity and explore the different ways it could affect an individual's pain experience and ability to self-manage. CS testing in clinic should inform decision-making on unidisciplinary or multidisciplinary management, medication, exercise prescription, psychological

support and referral of the given individual to the most appropriate health-care professional.

7.3. Future research

Prognostic research is of particular importance in the field of LBP where identification of prognostic factors or individuals at high risk of developing CLBP could allow for more targeted, tailored to the individual treatment, and prevention of chronicity (LeResche et al., 2013). The methodological approach of this study was along those lines as it aimed to provide the first stage of prognostic research for the development of a clinical prediction rule (CPR) determining the relationship between CS and self-management. CPRs are the combination of various factors that have been proven to have predictive capacity of a given outcome and are used to facilitate decision making (Reilly and Evans, 2006), predict outcome patterns (Randolph et al., 1998) and assist the early identification of high risk patients prone to adverse treatment outcomes (Hartling et al., 2002). This study has shown that CS, depression, disability and catastrophising predict different aspects of self-management. A robust CPR, in order to be clinically meaningful, needs to feature at least three prognostic factors (Laupacis et al., 1997) that have been identified through a rigorous multivariable approach (Moons et al., 2009).

It would be of significant value to explore whether CS in individuals suffering from Osteoarthritis, Whiplash Associated Disorders and Fibromyalgia hinders self-management in similar ways and magnitude. Future studies need to establish whether the same factors are predicting self-management within a different rehabilitation paradigm, in a different socioeconomic and cultural background or when a different musculoskeletal pathology is prevalent. That would internally and externally validate such a predictive tool, which is considered the next stage of CPR development (McGinn et al., 2008). The validation of such a rule in different contexts could trigger subsequent research that would explore the clinical impact of the rule and establish its sensitivity and specificity to predict poor SM/SC outcomes in relation to CLBP (Haskins et al., 2012).

Quantitative Sensory Testing was shown, alongside other factors used for the same reason, to have the capacity to assist in the prediction of outcomes other than pain and disability. Future studies need to further explore its internal and external validity by reporting associations with all outcomes believed to be contributing to CS development and perpetuation. Similarly, forthcoming research needs to establish

and reflect on associations between different modalities or between different CS indices. The implementation of a diverse protocol featuring static and dynamic modalities is endorsed by the findings of this study as it became apparent that all modalities are necessary to capture the different aspects of CS as each QST approach aims to assess a distinct pain processing mechanism.

This study used several indices other than QST for the identification of CS. Such indices have been previously developed for populations with knee pain or fibromyalgia. The findings suggest that certain indices might not be appropriate for a population with a clearly defined pain localisation such as CLBP or that their results might be influenced by associated symptoms not linked with CS. The finding that the Central Mechanisms trait, that has been previously found to predict musculoskeletal outcomes in knee pain, was also found to predict self-management in CLBP is promising. Nevertheless, future research should further explore its validity on CLBP as the 8 distinct traits were identified in a population with a different pathology. Identification of self-reported traits indicating central nervous system involvement (if not the same with 8 traits previously identified), their association with QST and their capacity to predict condition-specific clinical outcomes needs to be undertaken again to see whether distinct traits are underlying features of CS in CLBP.

Finally, future research needs to be conducted in different population and rehabilitation contexts to further explore and validate findings suggesting that increased levels of CS, depression, catastrophising and disability are contributing to the development of poor self-management in the long-term. Future studies need to take into account socioeconomic factors such as income, education, employment and/or family status that perhaps offer significant biopsychosocial diversity to the examined population that could influence the performance of the factors shown in this project to predict SM/SC. Multiple longitudinal observations of a much larger population are needed to allow analyses that will produce self-management and CS trajectories through time, which could further facilitate mechanistic research looking at existing mediators, moderators and confounders. Such studies, if replicate the findings presented in this body of work could provide consistency and promote the planning of innovative randomised controlled trials (RCTs) using cluster-randomisation that could prove or disprove any causal relationship between CS and poor SM/SC outcomes.

7.4. General limitations

The project is subject to several limitations. The predictive capacity of QST was established across a diverse cohort of musculoskeletal disorders. Even though sensitivity analyses of a CLBP sub-group did not suggest a different conclusion, the amount of CLBP-related studies might have been small and prone to publication bias. Dynamic modalities were found to perform better in the prediction of musculoskeletal outcomes. However, CPM demonstrated poor reliability, which is consistent with existing literature and therefore, any inference from the application of this QST modality needs to be viewed with caution. Distinct CS indices, other than QST, were implemented for the purposes of this study. None of these indices were originally developed to measure or identify CS on populations with CLBP and their reliability as well as internal and external validity has not been formally established in relation to that pathology. Even though cognitive capacity was a prerequisite for undertaking the CBT-based intervention as well as participating in the observational study, cognitive testing during the intervention was not conducted to assess whether altered cognitive capacity (medically-induced or otherwise) could be a predictor of poor self-management. Causal inference within this study has been speculated as a result of exploratory analyses and readers are advised to take this under consideration when interpreting the results.

7.5. Conclusion

In summary, the key findings from the work featured within this thesis are:

1. Quantitative Sensory Testing generally and dynamic modalities specifically can predict follow-up pain, disability and negative affect in musculoskeletal disorders, which indicated that early signs of CS can provide information about poor future musculoskeletal outcomes and hinted that it could also predict outcomes of a more multidimensional nature such as self-management.
2. A QST protocol combining static and dynamic modalities is a reliable tool in the detection of CS in individuals with CLBP and it demonstrates good internal validity as modalities are associated with each other in unique patterns.
3. A total number of 9 or more painful sites on a body manikin divided in 24 body areas appears to be indicative of high levels of CS in a population

suffering from CLBP and is associated with musculoskeletal outcomes and other CS indices at a single timepoint.

4. Eight distinct items representing somatic, cognitive and psychological self-report traits contribute to the development of a single underlying Central Mechanisms construct that is associated with other CS indices, musculoskeletal and self-management outcomes at a single timepoint.
5. Central Sensitisation, measured with QST and a single Central Mechanisms construct can predict distinct long-term self-management outcomes highlighting that CS is a complex multidimensional phenomenon with physical, psychological and cognitive aspects.
6. High levels of depression, catastrophising and disability are also predictors of poor self-management outcomes in the long-term.
7. Central Sensitisation, depression and disability appear to be directly and independently influencing self-management across time leading to poor long-term self-management outcomes.

REFERENCES

Aasvang, E. K., Gmaehle, E., Hansen, J. B., Gmaehle, B., Forman, J. L., Schwarz, J., Bittner, R. & Kehlet, H. 2010. Predictive risk factors for persistent postherniotomy pain. *Anesthesiology: The Journal of the American Society of Anesthesiologists,* 112**,** 957-969.

Abramson, L. Y., Garber, J. & Seligman, M. E. 1980. Learned helplessness in humans: An attributional analysis. *Human helplessness: Theory and applications,* 3**,** 34.

Abrishami, A., Chan, J., Chung, F. & Wong, J. 2011. Preoperative Pain Sensitivity and Its Correlation with Postoperative Pain and Analgesic ConsumptionA Qualitative Systematic Review. *The Journal of the American Society of Anesthesiologists,* 114**,** 445-457.

Adnan, R., Van Oosterwijck, J., Cagnie, B., Dhondt, E., Schouppe, S., Van Akeleyen, J., Logghe, T. & Danneels, L. 2017. Determining predictive outcome factors for a multimodal treatment program in low back pain patients: A retrospective cohort study. *Journal of manipulative and physiological therapeutics,* 40**,** 659-667.

Akin-Akinyosoye, K., Frowd, N., Marshall, L., Stocks, J., Fernandes, G. S., Valdes, A., Mcwilliams, D. F., Zhang, W., Doherty, M., Ferguson, E. & Walsh, D. A. 2018. Traits associated with central pain augmentation in the Knee Pain In the Community (KPIC) cohort. *Pain,* 159**,** 1035.

Akin-Akinyosoye, K., Sarmanova, A., Fernandes, G., Frowd, N., Swaithes, L., Stocks, J., Valdes, A., Mcwilliams, D., Zhang, W. & Doherty, M. 2019. Baseline self-report 'central mechanisms' trait predicts persistent knee pain in the Knee Pain in the Community (KPIC) cohort. *Osteoarthritis and Cartilage*.

Akin-Akinyosoye, K., Sarmanova, A., Fernandes, G., Frowd, N., Swaithes, L., Stocks, J., Valdes, A., Mcwilliams, D. F., Zhang, W. & Doherty, M. 2020. Baseline self-report 'central mechanisms' trait predicts persistent knee pain in the Knee Pain in the Community (KPIC) cohort. *Osteoarthritis and Cartilage,* 28**,** 173-181.

Akinci, A., Al Shaker, M., Chang, M., Cheung, C., Danilov, A., José Dueñas, H., Kim, Y. C., Guillen, R., Tassanawipas, W. & Treuer, T. 2016. Predictive factors and clinical biomarkers for treatment in patients with chronic pain caused by osteoarthritis with a central sensitisation component. *International journal of clinical practice,* 70**,** 31-44.

Al-Harthy, M., Ohrbach, R., Michelotti, A. & List, T. 2016. The effect of culture on pain sensitivity. *Journal of oral rehabilitation,* 43**,** 81-88.

Alkan, H., Ardic, F., Erdogan, C., Sahin, F., Sarsan, A. & Findikoglu, G. 2013. Turkish version of the painDETECT questionnaire in the assessment of neuropathic pain: a validity and reliability study. *Pain Medicine,* 14**,** 1933-1943.

Allegri, M., Montella, S., Salici, F., Valente, A., Marchesini, M., Compagnone, C., Baciarello, M., Manferdini, M. E. & Fanelli, G. 2016. Mechanisms of low back pain: a guide for diagnosis and therapy. *F1000Research,* 5.

Altman, D. G. 2001. Systematic reviews in health care: Systematic reviews of evaluations of prognostic variables. *BMJ: British Medical Journal,* 323**,** 224.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education & Joint Committee on Standards for Educational and Psychological Testing 1999. *Standards for educational and psychological testing*, Amer Educational Research Assn.

Andersen, T. E. 2012. *The development of chronic pain–a diathesis-stress perspective on whiplash associated disorders and other pain conditions.* Syddansk Universitet. Det Sundhedsvidenskabelige Fakultet.

Andersen, T. E., Karstoft, K. I., Brink, O. & Elklit, A. 2016. Pain-catastrophizing and fear-avoidance beliefs as mediators between post-traumatic stress symptoms and pain following whiplash injury–A prospective cohort study. *European Journal of Pain,* 20**,** 1241-1252.

Antoine-Mahut, D. & Gaukroger, S. 2017. *Descartes' Treatise on Man and its Reception*, Springer.

Antonovsky, A. 1979. *Health, Stress, and Coping: New Perspectives on Mental and Physical Well-being,* San Francisco, Jossey-Bass.

Apkarian, A. V., Bushnell, M. C., Treede, R. D. & Zubieta, J. K. 2005. Human brain mechanisms of pain perception and regulation in health and disease. *European journal of pain,* 9**,** 463-463.

Apkarian, A. V., Sosa, Y., Krauss, B. R., Thomas, P. S., Fredrickson, B. E., Levy, R. E., Harden, R. N. & Chialvo, D. R. 2004. Chronic pain patients are impaired on an emotional decision-making task. *Pain,* 108**,** 129-136.

Arendt-Nielsen, L., Egsgaard, L. L. & Petersen, K. K. 2016. Evidence for a central mode of action for etoricoxib (COX-2 inhibitor) in patients with painful knee osteoarthritis. *Pain,* 157**,** 1634-1644.

Arendt-Nielsen, L., Fernández-De-Las-Peñas, C. & Graven-Nielsen, T. 2011. Basic aspects of musculoskeletal pain: from acute to chronic pain. *Journal of Manual & Manipulative Therapy,* 19**,** 186-193.

Arendt-Nielsen, L. & Yarnitsky, D. 2009. Experimental and clinical applications of quantitative sensory testing applied to skin, muscles and viscera. *The Journal of Pain,* 10**,** 556-572.

Arendt-Nielsen, L., Egsgaard, L. L., Petersen, K. K., Eskehave, T. N., Graven-Nielsen, T., Hoeck, H. & Simonsen, O. 2015. A mechanism-based pain sensitivity index to characterize knee osteoarthritis patients with different disease stages and pain levels. *European Journal of Pain,* 19**,** 1406-1417.

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**,** 216-241.

Arpino, L., Iavarone, A., Parlato, C. & Moraci, A. 2004. Prognostic role of depression after lumbar disc surgery. *Neurological Sciences,* 25**,** 145-147.

Asghari, A. & Nicholas, M. K. 2001. Pain self-efficacy beliefs and pain behaviour. A prospective study. *Pain,* 94**,** 85-100.

Asghari, A. & Nicholas, M. K. 2009. An investigation of pain self-efficacy beliefs in Iranian chronic pain patients: a preliminary validation of a translated English-language scale. *Pain medicine,* 10**,** 619-632.

Atlas, L. Y. & Wager, T. D. 2012. How expectations shape pain. *Neuroscience letters,* 520**,** 140-148.

Backonja, M. M., Attal, N., Baron, R., Bouhassira, D., Drangholt, M., Dyck, P. J., Edwards, R. R., Freeman, R., Gracely, R. & Haanpaa, M. H. 2013. Value of quantitative sensory testing in neurological and pain disorders: NeuPSIG consensus. *Pain,* 154**,** 1807-1819.

Balaguier, R., Madeleine, P. & Vuillerme, N. 2016. Is one trial sufficient to obtain excellent pressure pain threshold reliability in the low back of asymptomatic individuals? A test-retest study. *PloS one,* 11**,** e0160866.

Baliki, M. N., Chialvo, D. R., Geha, P. Y., Levy, R. M., Harden, R. N., Parrish, T. B. & Apkarian, A. V. 2006. Chronic pain and the emotional brain: specific brain activity associated with spontaneous fluctuations of intensity of chronic back pain. *Journal of Neuroscience,* 26**,** 12165-12173.

Bandura, A. 1986. *Social foundations of thought and action: A social cognitive theory*, Prentice-Hall, Inc.

Bandura, A. 2001. Social cognitive theory: An agentic perspective. *Annual review of psychology,* 52**,** 1-26.

Bandura, A., Freeman, W. & Lightsey, R. 1999. Self-efficacy: The exercise of control. Springer.
Banerjee, A., Hendrick, P., Bhattacharjee, P. & Blake, H. 2018. A systematic review of outcome measures utilised to assess self-management in clinical trials in patients with chronic pain. *Patient education and counseling,* 101**,** 767-778.
Banic, B., Petersen-Felix, S., Andersen, O. K., Radanov, B. P., Villiger, P., Arendt-Nielsen, L. & Curatolo, M. 2004. Evidence for spinal cord hypersensitivity in chronic pain after whiplash injury and in fibromyalgia. *Pain,* 107**,** 7-15.
Bar Ziv, Y., Shemesh, S., Agar, G., Benedict, S., Heller, S. & Kosashvili, Y. 2016. The Sphygmomanometer Pain Test: A Simple Method for Identifying Patients at Risk of Excessive Pain after Total knee Arthroplasty. *The Journal of arthroplasty,* 31**,** 798-801.
Barlow, J., Wright, C., Sheasby, J., Turner, A. & Hainsworth, J. 2002. Self-management approaches for people with chronic conditions: a review. *Patient education and counseling,* 48**,** 177-187.
Bartlett, M. S. 1947. The use of transformations. *Biometrics,* 3**,** 39-52.
Beecham, J. & Knapp, M. 2001. Costing psychiatric interventions. *Measuring mental health needs,* 2**,** 200-224.
Bergman, S. 2007. Management of musculoskeletal pain. *Best Practice & Research Clinical Rheumatology,* 21**,** 153-166.
Bipm, I., Ifcc, I. & Iso, I. 2008. IUPAP, and OIML, 2008,"Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement," Joint Committee for Guides in Metrology. Technical Report No. JCGM 100.
Blackburn, J., Qureshi, A., Amirfeyz, R. & Bannister, G. 2012. Does preoperative anxiety and depression predict satisfaction after total knee replacement? *The Knee,* 19**,** 522-524.
Bland, J. M. & Altman, D. G. 1999. Measuring agreement in method comparison studies. *Statistical methods in medical research,* 8**,** 135-160.
Bodenheimer, T. 2005. Helping patients improve their health-related behaviors: what system changes do we need? *Disease Management,* 8**,** 319-330.
Bodenheimer, T., Lorig, K., Holman, H. & Grumbach, K. 2002. Patient self-management of chronic disease in primary care. *Jama,* 288**,** 2469-2475.
Boger, E., Ellis, J., Latter, S., Foster, C., Kennedy, A., Jones, F., Fenerty, V., Kellar, I. & Demain, S. 2015. Self-management and self-management support outcomes: a systematic review and mixed research synthesis of stakeholder views. *PloS one,* 10**,** e0130990.
Boger, E. J., Demain, S. & Latter, S. 2013. Self-management: a systematic review of outcome measures adopted in self-management interventions for stroke. *Disability and rehabilitation,* 35**,** 1415-1428.
Boldy, D. & Silfo, E. 2006. Chronic disease self-management by people from lower socio-economic backgrounds: action planning and impact. *Journal of Integrated Care,* 14**,** 19.
Bombardier, C., Hayden, J. & Beaton, D. E. 2001. Minimal clinically important difference. Low back pain: outcome measures. *The Journal of Rheumatology,* 28**,** 431-438.
Borkan, J., Van Tulder, M., Reis, S., Schoene, M. L., Croft, P. & Hermoni, D. 2002. Advances in the field of low back pain in primary care: a report from the fourth international forum. *Spine,* 27**,** E128-E132.
Borrell-Carrió, F., Suchman, A. L. & Epstein, R. M. 2004. The biopsychosocial model 25 years later: principles, practice, and scientific inquiry. *The Annals of Family Medicine,* 2**,** 576-582.
Bosma, F. K. & Kessels, R. P. 2002. Cognitive impairments, psychological dysfunction, and coping styles in patients with chronic whiplash syndrome. *Cognitive and Behavioral Neurology,* 15**,** 56-65.

Bosma, R. L., Mojarad, E. A., Leung, L., Pukall, C., Staud, R. & Stroman, P. W. 2016. F MRI of spinal and supra-spinal correlates of temporal pain summation in fibromyalgia patients. *Human brain mapping,* 37**,** 1349-1360.

Bourke, J. H., Langford, R. M. & White, P. D. 2015. The common link between functional somatic syndromes may be central sensitisation. *Journal of psychosomatic research,* 78**,** 228-236.

Braveman, P. & Gruskin, S. 2003. Defining equity in health. *Journal of Epidemiology & Community Health,* 57**,** 254-258.

Bridges, S. 2012. Chronic Pain. *Health Survey for England 2011.* UK: Health and Social Care Information Centre.

Brotman, D. J., Walker, E., Lauer, M. S. & O'brien, R. G. 2005. In search of fewer independent risk factors. *Archives of internal medicine,* 165**,** 138-145.

Brown, D., Mulvey, M., Cordingley, L., Rashid, A., Horan, M., Pendleton, N., Duncan, R. & Mcbeth, J. 2016. The relationship between psychological distress and multiple tender points across the adult lifespan. *Archives of gerontology and geriatrics,* 63**,** 102-107.

Browne, M. W. & Cudeck, R. 1992. Alternative ways of assessing model fit. *Sociological methods & research,* 21**,** 230-258.

Bruyère, O., Demoulin, M., Beaudart, C., Hill, J. C., Maquet, D., Genevay, S., Mahieu, G., Reginster, J.-Y., Crielaard, J.-M. & Demoulin, C. 2014. Validity and reliability of the French version of the STarT Back screening tool for patients with low back pain. *Spine,* 39**,** E123-E128.

Buchgreitz, L., Lyngberg, A., Bendtsen, L. & Jensen, R. 2008. Increased pain sensitivity is not a risk factor but a consequence of frequent headache: a population-based follow-up study. *PAIN®,* 137**,** 623-630.

Buragadda, S., Aleisa, E. S. & Melam, G. R. 2018. Fear avoidance beliefs and disability among women with low back pain. *Neuropsychiatry (London),* 8**,** 80-86.

Burke, A. L., Mathias, J. L. & Denson, L. A. 2015. Psychological functioning of people living with chronic pain: A meta-analytic review. *British Journal of Clinical Psychology,* 54**,** 345-360.

Burton, A. K., Balagué, F., Cardon, G., Eriksen, H., Henrotin, Y., Lahad, A., Leclerc, A., Muller, G. & Van Der Beek, A. 2004. European guidelines for prevention in low back pain.

Bushnell, M. C., Čeko, M. & Low, L. A. 2013. Cognitive and emotional control of pain and its disruption in chronic pain. *Nature Reviews Neuroscience,* 14**,** 502.

Byford, S., Leese, M., Knapp, M., Seivewright, H., Cameron, S., Jones, V., Davidson, K. & Tyrer, P. 2007. Comparison of alternative methods of collection of service use data for the economic evaluation of health care interventions. *Health economics,* 16**,** 531-536.

Cacioppo, J. T., Malarkey, W. B., Kiecolt-Glaser, J. K., Uchino, B. N., Sgoutas-Emch, S. A., Sheridan, J. F., Berntson, G. G. & Glaser, R. 1995. Heterogeneity in neuroendocrine and immune responses to brief psychologicalstressors as a function of autonomic cardiac activation. *Psychosomatic medicine,* 57**,** 154-164.

Campbell, C. M., Buenaver, L. F., Finan, P., Bounds, S. C., Redding, M., Mccauley, L., Robinson, M., Edwards, R. R. & Smith, M. T. 2015. Sleep, pain catastrophizing, and central sensitization in knee osteoarthritis patients with and without insomnia. *Arthritis care & research,* 67**,** 1387-1396.

Campbell, C. M. & Edwards, R. R. 2012. Ethnic differences in pain and pain management. *Pain management,* 2**,** 219-230.

Campbell, J. N. & Meyer, R. A. 2006. Mechanisms of neuropathic pain. *Neuron,* 52**,** 77-92.

Cao, H. & Zhang, Y.-Q. 2008. Spinal glial activation contributes to pathological pain states. *Neuroscience & Biobehavioral Reviews,* 32**,** 972-983.

Cappelleri, J. C., Bienen, E. J., Koduru, V. & Sadosky, A. 2014. Measurement properties of painDETECT by average pain severity. *ClinicoEconomics and outcomes research: CEOR,* 6**,** 497.

Cathcart, S. & Pritchard, D. 2006. Reliability of pain threshold measurement in young adults. *The journal of headache and pain,* 7**,** 21.

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

Chacur, M., Lambertz, D., Hoheisel, U. & Mense, S. 2009. Role of spinal microglia in myositis-induced central sensitisation: an immunohistochemical and behavioural study in rats. *European Journal of Pain,* 13**,** 915-923.

Chaparro, L. E., Furlan, A. D., Deshpande, A., Mailis-Gagnon, A., Atlas, S. & Turk, D. C. 2014. Opioids compared with placebo or other treatments for chronic low back pain: an update of the Cochrane Review. *Spine,* 39**,** 556-563.

Chapman, C. R. Neuromatrix theory: do we need it? Pain Forum, 1996. Churchill Livingstone, 139-142.

Chapman, C. R., Donaldson, G. W., Davis, J. J. & Bradshaw, D. H. 2011. Improving individual measurement of postoperative pain: the pain trajectory. *The journal of Pain,* 12**,** 257-262.

Cheng, S. K. & Leung, F. 2000. Catastrophizing, locus of control, pain, and disability in Chinese chronic low back pain patients. *Psychology and Health,* 15**,** 721-730.

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**,** 760-766.

Chiarotto, A., Vanti, C., Cedraschi, C., Ferrari, S., Ostelo, R. W. & Pillastrini, P. 2016. Responsiveness and minimal important change of the pain self-efficacy questionnaire and short forms in patients with chronic low back pain. *The Journal of Pain,* 17**,** 707-718.

Chou, R., Loeser, J. D., Owens, D. K., Rosenquist, R. W., Atlas, S. J., Baisden, J., Carragee, E. J., Grabois, M., Murphy, D. R. & Resnick, D. K. 2009. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. *Spine,* 34**,** 1066-1077.

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**,** 28-34.

Clark, N. M., Becker, M. H., Janz, N. K., Lorig, K., Rakowski, W. & Anderson, L. 1991. Self-management of chronic disease by older adults a review and questions for research. *Journal of Aging and Health,* 3**,** 3-27.

Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. 2013. *Applied multiple regression/correlation analysis for the behavioral sciences*, Routledge.

Coombes, B. K., Bisset, L. & Vicenzino, B. 2015. Cold hyperalgesia associated with poorer prognosis in lateral epicondylalgia: a 1-year prognostic study of physical and psychological factors. *The Clinical journal of pain,* 31**,** 30-5.

Cooper, K., Smith, B. H. & Hancock, E. 2009. Patients' perceptions of self-management of chronic low back pain: evidence for enhancing patient education and support. *Physiotherapy,* 95**,** 43-50.

Copay, A. G., Subach, B. R., Glassman, S. D., Polly Jr, D. W. & Schuler, T. C. 2007. Understanding the minimum clinically important difference: a review of concepts and methods. *The Spine Journal,* 7**,** 541-546.

Corey, D. M., Dunlap, W. P. & Burke, M. J. 1998. Averaging correlations: Expected values and bias in combined Pearson rs and Fisher's z transformations. *The Journal of general psychology,* 125**,** 245-261.

Coronado, R., Bialosky, J., Bishop, M., Riley, J., Robinson, M., Michener, L. & George, S. 2015a. The comparative effects of spinal and peripheral thrust manipulation and exercise on pain sensitivity and the relation to clinical outcome: a mechanistic trial using a shoulder pain model. *The Journal of orthopaedic and sports physical therapy* [Online], 45. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/569/CN-01132569/frame.html.

Coronado, R. A., George, S. Z., Devin, C. J., Wegener, S. T. & Archer, K. R. 2015b. Pain Sensitivity and Pain Catastrophizing Are Associated With Persistent Pain and Disability After Lumbar Spine Surgery. *Archives of physical medicine and rehabilitation,* 96**,** 1763-70.

Corrêa, J. B., Costa, L. O. P., De Oliveira, N. T. B., 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. *Experimental brain research,* 233**,** 2391-2399.

Costa-Black, K. M., Loisel, P., Anema, J. R. & Pransky, G. 2010. Back pain and work. *Best Practice & Research Clinical Rheumatology,* 24**,** 227-240.

Courtney, C. A., Kavchak, A. E., Lowry, C. D. & O'hearn, M. A. 2010. Interpreting joint pain: quantitative sensory testing in musculoskeletal management. *journal of orthopaedic & sports physical therapy,* 40**,** 818-825.

Craig, K. D. 2009. The social communication model of pain. *Canadian Psychology/Psychologie canadienne,* 50**,** 22.

Croft, P., Burt, J., Schollum, J., Thomas, E., Macfarlane, G. & Silman, A. 1996. More pain, more tender points: is fibromyalgia just one end of a continuous spectrum? *Annals of the rheumatic diseases,* 55**,** 482-485.

Croft, P. R., Dunn, K. M. & Raspe, H. 2006. Course and prognosis of back pain in primary care: the epidemiological perspective. *Pain,* 122**,** 1-3.

Cruz-Almeida, Y. & Fillingim, R. B. 2014. Can quantitative sensory testing move us closer to mechanism-based pain management? *Pain medicine,* 15**,** 61-72.

Daenen, L., Nijs, J., Raadsen, B., Roussel, N., Cras, P. & Dankaerts, W. 2013a. Cervical motor dysfunction and its predictive value for long-term recovery in patients with acute whiplash-associated disorders: a systematic review. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine,* 45**,** 113-122.

Daenen, L., Nijs, J., Roussel, N., Wouters, K., Van Loo, M. & Cras, P. 2013b. Dysfunctional pain inhibition in patients with chronic whiplash-associated disorders: an experimental study. *Clinical rheumatology,* 32**,** 23-31.

Dagenais, S., Caro, J. & Haldeman, S. 2008. A systematic review of low back pain cost of illness studies in the United States and internationally. *The spine journal,* 8**,** 8-20.

Damush, T., Kroenke, K., Bair, M., Wu, J., Tu, W., Krebs, E. & Poleshuck, E. 2016. Pain self-management training increases self-efficacy, self-management behaviours and pain and depression outcomes. *European Journal of Pain*.

Darnall, B. D., Sturgeon, J. A., Kao, M.-C., Hah, J. M. & Mackey, S. C. 2014. From catastrophizing to recovery: a pilot study of a single-session treatment for pain catastrophizing. *Journal of pain research,* 7**,** 219.

Davies, C. C. & Nitz, A. J. 2009. Psychometric properties of the Roland-Morris Disability Questionnaire compared to the Oswestry Disability Index: a systematic review. *Physical therapy reviews,* 14**,** 399-408.

Davis, A., Chinn, D. J. & Sharma, S. 2013. Prediction of post-operative pain following arthroscopic subacromial decompression surgery: an observational study. *F1000Research,* 2.

Dazinger, S., Corcoran, M., Cabal, A., Dazinger, S. & Heflin, C. 2000. Barriers to the employment of welfare recipients. *Prosperity for all?: The economic boom and African Americans***,** 245.

Delitto, A., Erhard, R. E. & Bowling, R. W. 1995. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. *Physical therapy,* 75**,** 470-485.

Delitto, A., George, S. Z., Van Dillen, L. R., Whitman, J. M., Sowa, G., Shekelle, P., Denninger, T. R. & Godges, J. J. 2012. Low back pain. *J Orthop Sports Phys Ther,* 42**,** A1-57.

Denison, E., Asenlof, P. & Lindberg, P. 2004. Self-efficacy, fear avoidance, and pain intensity as predictors of disability in subacute and chronic musculoskeletal pain patients in primary health care. *Pain,* 111**,** 245-52.

Denison, E., Åsenlöf, P., Sandborgh, M. & Lindberg, P. 2007. Musculoskeletal pain in primary health care: subgroups based on pain intensity, disability, self-efficacy, and fear-avoidance variables. *The Journal of Pain,* 8**,** 67-74.

Devinsky, O., Morrell, M. J. & Vogt, B. A. 1995. Contributions of anterior cingulate cortex to behaviour. *Brain,* 118**,** 279-306.

Deyo, R. A., Mirza, S. K., Turner, J. A. & Martin, B. I. 2009. Overtreating chronic back pain: time to back off? *The Journal of the American Board of Family Medicine,* 22**,** 62-68.

Di Iorio, A., Abate, M., Guralnik, J. M., Bandinelli, S., Cecchi, F., Cherubini, A., Corsonello, A., Foschini, N., Guglielmi, M. & Lauretani, F. 2007. From chronic low back pain to disability, a multifactorial mediated pathway: the InCHIANTI study. *Spine,* 32**,** E809.

Downar, J. & Daskalakis, Z. J. 2013. New targets for rTMS in depression: a review of convergent evidence. *Brain stimulation,* 6**,** 231-240.

Du, S., Hu, L., Dong, J., Xu, G., Chen, X., Jin, S., Zhang, H. & Yin, H. 2017. Self-management program for chronic low back pain: A systematic review and meta-analysis. *Patient Education and Counseling,* 100**,** 37-49.

Du, S., Yuan, C., Xiao, X., Chu, J., Qiu, Y. & Qian, H. 2011. Self-management programs for chronic musculoskeletal pain conditions: a systematic review and meta-analysis. *Patient education and counseling,* 85**,** e299-e310.

Dubois, J.-D., Cantin, V., Piche, M. & Descarreaux, M. 2016. Physiological and Psychological Predictors of Short-Term Disability in Workers with a History of Low Back Pain: A Longitudinal Study. *PloS one,* 11**,** e0165478.

Earley, P. C., Gibson, C. B. & Chen, C. C. 1999. “How did I do?” versus “How did we do?” Cultural contrasts of performance feedback use and self-efficacy. *Journal of Cross-Cultural Psychology,* 30**,** 594-619.

Eccleston, C. & Crombez, G. 1999. Pain demands attention: A cognitive–affective model of the interruptive function of pain. *Psychological bulletin,* 125**,** 356.

Edwards, R. R., Dolman, A. J., Martel, M. O., Finan, P. H., Lazaridou, A., Cornelius, M. & Wasan, A. D. 2016a. Variability in conditioned pain modulation predicts response to NSAID treatment in patients with knee osteoarthritis. *BMC Musculoskeletal Disorders,* 17 (1) (no pagination).

Edwards, R. R., Dworkin, R. H., Sullivan, M. D., Turk, D. C. & Wasan, A. D. 2016b. The role of psychosocial processes in the development and maintenance of chronic pain. *The Journal of Pain,* 17**,** T70-T92.

Edwards, R. R., Sarlani, E., Wesselmann, U. & Fillingim, R. B. 2005. Quantitative assessment of experimental pain perception: multiple domains of clinical relevance. *Pain,* 114**,** 315-319.

Egger, M., Smith, G. D., Schneider, M. & Minder, C. 1997. Bias in meta-analysis detected by a simple, graphical test. *Bmj,* 315**,** 629-634.

Eisenberg, E., Midbari, A., Haddad, M. & Pud, D. 2010. Predicting the analgesic effect to oxycodone by ‘static’and ‘dynamic’quantitative sensory testing in healthy subjects. *Pain,* 151**,** 104-109.

Eisenberger, N. I., Lieberman, M. D. & Williams, K. D. 2003. Does rejection hurt? An fMRI study of social exclusion. *Science,* 302**,** 290-292.

Ellard, D. R., Underwood, M., Achana, F., Antrobus, J. H., Balasubramanian, S., Brown, S., Cairns, M., Griffin, J., Griffiths, F. & Haywood, K. 2017. Facet joint injections for people with persistent non-specific low back pain (Facet Injection Study): a feasibility study for a randomised controlled trial. *Health technology assessment (Winchester, England),* 21**,** 1.

Elsworth, G. R., Nolte, S. & Osborne, R. H. 2015. Factor structure and measurement invariance of the Health Education Impact Questionnaire: Does the subjectivity of the response perspective threaten the contextual validity of inferences? *SAGE open medicine,* 3**,** 2050312115585041.

Engel, G. L. 1977. The need for a new medical model: a challenge for biomedicine. *Science,* 196**,** 129-136.

Enthoven, W., Roelofs, P. D., Deyo, R. A., Van Tulder, M. W. & Koes, B. W. 2016. Non-steroidal anti-inflammatory drugs for chronic low back pain. *The Cochrane Library*.

Epstein, J., Osborne, R. H., Elsworth, G. R., Beaton, D. E. & Guillemin, F. 2015. Cross-cultural adaptation of the Health Education Impact Questionnaire: experimental study showed expert committee, not back-translation, added value. *Journal of clinical epidemiology,* 68**,** 360-369.

Erickson, K. I., Voss, M. W., Prakash, R. S., Basak, C., Szabo, A., Chaddock, L., Kim, J. S., Heo, S., Alves, H. & White, S. M. 2011. Exercise training increases size of hippocampus and improves memory. *Proceedings of the National Academy of Sciences,* 108**,** 3017-3022.

Escolar-Reina, P., Medina-Mirapeix, F., Gascón-Cánovas, J. J., Montilla-Herrador, J., Valera-Garrido, J. F. & Collins, S. M. 2009. Self-management of chronic neck and low back pain and relevance of information provided during clinical encounters: an observational study. *Archives of physical medicine and rehabilitation,* 90**,** 1734-1739.

Euser, A. M., Zoccali, C., Jager, K. J. & Dekker, F. W. 2009. Cohort studies: prospective versus retrospective. *Nephron Clinical Practice,* 113**,** c214-c217.

Fabio Antonaci, M. 1998. Pressure algometry in healthy subjects: inter-examiner variability. *Scand J Rehab Med,* 30**,** 8.

Fabrigar, L. R., Wegener, D. T., Maccallum, R. C. & Strahan, E. J. 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods,* 4**,** 272.

Farrar, J. T., Young Jr, J. P., Lamoreaux, L., Werth, J. L. & Poole, R. M. 2001. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain,* 94**,** 149-158.

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**,** 2679-2690.

Fernández-Carnero, J., Binderup, A. T., Ge, H.-Y., Fernandez-De-Las-Penas, C., Arendt-Nielsen, L. & Madeleine, P. 2010. Pressure pain sensitivity mapping in experimentally induced lateral epicondylalgia. *Medicine and science in sports and exercise,* 42**,** 922-927.

Fernández-Carnero, J., Fernández-De-Las-Peñas, C., De La Llave-Rincón, A. I., Ge, H.-Y. & Arendt-Nielsen, L. 2009. Widespread mechanical pain hypersensitivity as sign of central sensitization in unilateral epicondylalgia: a blinded, controlled study. *The Clinical journal of pain,* 25**,** 555-561.

Ferrari, S., Vanti, C., Pellizzer, M., Dozza, L., Monticone, M. & Pillastrini, P. 2019. Is there a relationship between self-efficacy, disability, pain and sociodemographic characteristics in chronic low back pain? A multicenter retrospective analysis. *Archives of Physiotherapy,* 9**,** 9.

Ferraz, M. B., Quaresma, M., Aquino, L., Atra, E., Tugwell, P. & Goldsmith, C. 1990. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *The Journal of rheumatology,* 17**,** 1022-1024.

Finestone, H. M., Yanni, M. M. & Dalzell, C. J. 2015. Patients' recall of diagnostic and treatment information improves with use of the Pain explanation and treatment Diagram in an outpatient chronic pain clinic. *Pain Research and Management,* 20**,** 145-151.

Fingleton, C., Smart, K., Moloney, N., Fullen, B. & Doody, C. 2015. Pain sensitization in people with knee osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis and cartilage,* 23**,** 1043-1056.

Fornasari, D. 2012. Pain mechanisms in patients with chronic pain. *Clinical drug investigation,* 32**,** 45-52.

Freeman, M. D., Nystrom, A. & Centeno, C. 2009. Chronic whiplash and central sensitization; an evaluation of the role of a myofascial trigger points in pain modulation. *Journal of Brachial Plexus and Peripheral Nerve Injury,* 4**,** 2.

French, S. D., Green, S., Buchbinder, R. & Barnes, H. 2010. Interventions for improving the appropriate use of imaging in people with musculoskeletal conditions. *The Cochrane Library*.

Freynhagen, R., Baron, R., Gockel, U. & Tölle, T. R. 2006. Pain DETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. *Current medical research and opinion,* 22**,** 1911-1920.

Fuchs, P. N., Peng, Y. B., Boyette-Davis, J. A. & Uhelski, M. L. 2014. The anterior cingulate cortex and pain processing. *Frontiers in integrative neuroscience,* 8**,** 35.

Gatchel, R. J., Peng, Y. B., Peters, M. L., Fuchs, P. N. & Turk, D. C. 2007. The biopsychosocial approach to chronic pain: scientific advances and future directions. *Psychological bulletin,* 133**,** 581.

Gay, C., Horn, M., Bishop, M., Robinson, M. & Bialosky, J. 2015. Investigating dynamic pain sensitivity in the context of the fear-avoidance model. *European Journal of Pain,* 19**,** 48-58.

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**,** 548-556.

George, S. Z. & Beneciuk, J. M. 2015. Psychological predictors of recovery from low back pain: a prospective study. *BMC musculoskeletal disorders,* 16**,** 49.

George, S. Z., Wittmer, V. T., Fillingim, R. B. & Robinson, M. E. 2006. Fear-avoidance beliefs and temporal summation of evoked thermal pain influence self-report of disability in patients with chronic low back pain. *Journal of Occupational Rehabilitation,* 16**,** 92-105.

Georgopoulos, V., Akin-Akinyosoye, K., Zhang, W., Mcwilliams, D. F., Hendrick, P. & Walsh, D. A. 2019. Quantitative Sensory Testing (QST) and predicting outcomes for musculoskeletal pain, disability and negative affect: a systematic review and meta-analysis. *Pain*.

Giesbrecht, R. J. S. & Battié, M. C. 2005. A comparison of pressure pain detection thresholds in people with chronic low back pain and volunteers without pain. *Physical therapy,* 85**,** 1085-1092.

Giesecke, T., Gracely, R. H., Grant, M. A., Nachemson, A., Petzke, F., Williams, D. A. & Clauw, D. J. 2004. Evidence of augmented central pain processing in idiopathic chronic low back pain. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 50**,** 613-623.

Girbés, E. L., Nijs, J., Torres-Cueco, R. & Cubas, C. L. 2013. Pain treatment for patients with osteoarthritis and central sensitization. *Physical therapy,* 93**,** 842-851.

Glantz, S. A., Slinker, B. K. & Neilands, T. B. 1990. *Primer of applied regression and analysis of variance*, McGraw-Hill New York.

Glombiewski, J. A., Hartwich-Tersek, J. & Rief, W. 2010. Depression in chronic back pain patients: prediction of pain intensity and pain disability in cognitive-behavioral treatment. *Psychosomatics,* 51**,** 130-136.

Gockel, M., Lindholm, H., Niemistö, L. & Hurri, H. 2008. Perceived disability but not pain is connected with autonomic nervous function among patients with chronic low back pain. *Journal of rehabilitation medicine,* 40**,** 355-358.

Goldenberg, D. L. 2010. Pain/depression dyad: a key to a better understanding and treatment of functional somatic syndromes. *The American journal of medicine,* 123**,** 675-682.

Goldsmith, R., Wright, C., Bell, S. F. & Rushton, A. 2012. Cold hyperalgesia as a prognostic factor in whiplash associated disorders: A systematic review. *Manual Therapy,* 17**,** 402-410.

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**,** 666-671.

Goodin, B. R., Bulls, H. W., Herbert, M. S., Schmidt, J., King, C. D., Glover, T. L., Sotolongo, A., Sibille, K. T., Cruz-Almeida, Y., Staud, R., Fessler, B. J., Redden, D. T., Bradley, L. A. & Fillingim, R. B. 2014. Temporal summation of pain as a prospective predictor of clinical pain severity in adults aged 45 years and older with knee osteoarthritis: ethnic differences. *Psychosomatic medicine,* 76**,** 302-10.

Gore, M., Sadosky, A., Stacey, B. R., Tai, K.-S. & Leslie, D. 2012. The burden of chronic low back pain: clinical comorbidities, treatment patterns, and health care costs in usual care settings. *Spine,* 37**,** E668-E677.

Gracely, R., Wall, P. & Melzack, R. 1994. Textbook of pain. *pp. 315-336). New York: Churchill Livingstone: Edinburgh*.

Graven-Nielsen, T. & Arendt-Nielsen, L. 2002. Peripheral and central sensitization in musculoskeletal pain disorders: an experimental approach. *Current rheumatology reports,* 4**,** 313-321.

Graven-Nielsen, 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**,** 2193-2202.

Graven-Nielsen, T., Wodehouse, T., Langford, R., Arendt-Nielsen, L. & Kidd, B. 2012. Normalization of widespread hyperesthesia and facilitated spatial summation of deep-tissue pain in knee osteoarthritis patients after knee replacement. *Arthritis & Rheumatism,* 64**,** 2907-2916.

Griffiths, C., Foster, G., Ramsay, J., Eldridge, S. & Taylor, S. 2007. How effective are expert patient (lay led) education programmes for chronic disease. *BMJ,* 334**,** 1254–1256.

Grosen, K., Fischer, I. W. D., Olesen, A. & Drewes, A. 2013. Can quantitative sensory testing predict responses to analgesic treatment? *European journal of pain,* 17**,** 1267-1280.

Gwilym, S., Oag, H., Tracey, I. & Carr, A. 2011. Evidence that central sensitisation is present in patients with shoulder impingement syndrome and influences the outcome after surgery. *J Bone Joint Surg Br,* 93**,** 498-502.

Hadjikhani, N., Ward, N., Boshyan, J., Napadow, V., Maeda, Y., Truini, A., Caramia, F., Tinelli, E. & Mainero, C. 2013. The missing link: enhanced functional connectivity between amygdala and visceroceptive cortex in migraine. *Cephalalgia,* 33**,** 1264-1268.

Hall, T., Briffa, K., Schäfer, A., Tampin, B. & Moloney, N. 2015. Quantitative Sensory Testing: Implications for clinical practice. *In:* Jull, G., Moore, A., Falla, D., Lewis, J., Mccarthy, C. & Sterling, M. (eds.) *Grieve's Modern Musculoskeletal Physiotherapy: Vertebral Column and Peripheral Joints.* 4th ed. UK: Elsevier Health Sciences.

Han, C. & Pae, C.-U. 2015. Pain and depression: a neurobiological perspective of their relationship. *Psychiatry investigation,* 12**,** 1-8.
Hansson, P., Backonja, M. & Bouhassira, D. 2007. Usefulness and limitations of quantitative sensory testing: clinical and research application in neuropathic pain states. *Pain,* 129**,** 256-259.
Harden, R. N., Bruehl, S., Stanos, S., Brander, V., Chung, O. Y., Saltz, S., Adams, A. & Stulberg, S. D. 2003. Prospective examination of pain-related and psychological predictors of CRPS-like phenomena following total knee arthroplasty: a preliminary study. *Pain,* 106**,** 393-400.
Hart, R. P., Martelli, M. F. & Zasler, N. D. 2000. Chronic pain and neuropsychological functioning. *Neuropsychology review,* 10**,** 131-149.
Hartley, S. & Stockley, R. 2016. Collaborative goal setting with adults attending physiotherapy at a specialist neuromuscular centre: is it always appropriate? A cross-sectional survey. *Physiotherapy,* 102**,** 320-326.
Hartling, L., Pickett, W. & Brison, R. J. 2002. Derivation of a clinical decision rule for whiplash associated disorders among individuals involved in rear-end collisions. *Accident; analysis and prevention,* 34**,** 531-539.
Hartvigsen, J., Hancock, M. J., Kongsted, A., Louw, Q., Ferreira, M. L., Genevay, S., Hoy, D., Karppinen, J., Pransky, G. & Sieper, J. 2018. What low back pain is and why we need to pay attention. *The Lancet,* 391**,** 2356-2367.
Hasenbring, M., Hallner, D. & Klasen, B. 2001. Psychological mechanisms in the transition from acute to chronic pain: over-or underrated? *Schmerz (Berlin, Germany),* 15**,** 442-447.
Haskins, R., Rivett, D. A. & Osmotherly, P. G. 2012. Clinical prediction rules in the physiotherapy management of low back pain: A systematic review. *Manual Therapy,* 17**,** 9-21.
Haugli, L., Steen, E., Laerum, E., Nygård, R. & Finset, A. 2003. Psychological distress and employment status. Effects of a group learning programme for patients with chronic musculoskeletal pain. *Psychology, health & medicine,* 8**,** 135-148.
Häuser, W., Wolfe, F., Tölle, T., Üçeyler, N. & Sommer, C. 2012. The role of antidepressants in the management of fibromyalgia syndrome. *CNS drugs,* 26**,** 297-307.
Hayden, J., Chou, R., Hogg-Johnson, S. & Bombardier, C. 2009. Systematic reviews of low back pain prognosis had variable methods and results—guidance for future prognosis reviews. *Journal of clinical epidemiology,* 62**,** 781-796. e1.
Hayden, J., Cote, P., Steenstra, I. & Bombardier, C. 2008. Identifying phases of investigation helps planning, appraising, and applying the results of explanatory prognosis studies. *Journal of clinical epidemiology,* 61**,** 552-560.
Hayden, J. A., Côté, P. & Bombardier, C. 2006. Evaluation of the quality of prognosis studies in systematic reviews. *Annals of internal medicine,* 144**,** 427-437.
Hayden, J. A., Van Der Windt, D. A., Cartwright, J. L., Côté, P. & Bombardier, C. 2013. Assessing bias in studies of prognostic factors. *Annals of internal medicine,* 158**,** 280-286.
Henriksen, M., Klokker, L., Graven-Nielsen, T., Bartholdy, C., Schjodt Jorgensen, T., Bandak, E., Danneskiold-Samsoe, B., Christensen, R. & Bliddal, H. 2014. Association of exercise therapy and reduction of pain sensitivity in patients with knee osteoarthritis: a randomized controlled trial. *Arthritis Care Res (Hoboken),* 66**,** 1836-43.
Herbert, R. D. 2014. Cohort studies of aetiology and prognosis: they're different. *Journal of physiotherapy,* 60**,** 241-244.
Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G. & Badia, X. 2011. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research,* 20**,** 1727-1736.
Herren-Gerber, R., Weiss, S., Arendt-Nielsen, L., Petersen-Felix, S., Di Stefano, G., Radanov, B. P. & Curatolo, M. 2004. Modulation of Central Hypersensitivity

by Nociceptive Input in Chronic Pain After Whiplash Injury. *Pain Medicine,* 5**,** 366-376.
Hidalgo-Lozano, A., Fernández-De-Las-Peñas, C., Alonso-Blanco, C., Ge, H.-Y., Arendt-Nielsen, L. & Arroyo-Morales, M. 2010. Muscle trigger points and pressure pain hyperalgesia in the shoulder muscles in patients with unilateral shoulder impingement: a blinded, controlled study. *Experimental brain research,* 202**,** 915-925.
Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L. & Sterne, J. A. 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Bmj,* 343**,** d5928.
Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. 2003. Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal,* 327**,** 557.
Hill, J. C., Dunn, K. M., Lewis, M., Mullis, R., Main, C. J., Foster, N. E. & Hay, E. M. 2008. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Care & Research,* 59**,** 632-641.
Hitchings, A., Lonsdale, D., Burrage, D. & Baker, E. 2018. *The top 100 drugs: clinical pharmacology and practical prescribing*, Elsevier Health Sciences.
Hochman, J., Davis, A., Elkayam, J., Gagliese, L. & Hawker, G. 2013. Neuropathic pain symptoms on the modified painDETECT correlate with signs of central sensitization in knee osteoarthritis. *Osteoarthritis and Cartilage,* 21**,** 1236-1242.
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**,** 523-531.
Hoy, D., Brooks, P., Blyth, F. & Buchbinder, R. 2010a. The epidemiology of low back pain. *Best practice & research Clinical rheumatology,* 24**,** 769-781.
Hoy, D., March, L., Brooks, P., Woolf, A., Blyth, F., Vos, T. & Buchbinder, R. 2010b. Measuring the global burden of low back pain. *Best practice & research Clinical rheumatology,* 24**,** 155-165.
Hübscher, M., Moloney, N., Leaver, A., Rebbeck, T., Mcauley, J. H. & Refshauge, K. M. 2013. Relationship between quantitative sensory testing and pain or disability in people with spinal pain—A systematic review and meta-analysis. *PAIN®,* 154**,** 1497-1504.
Hung, C.-I., Liu, C.-Y. & Fu, T.-S. 2015. Depression: An important factor associated with disability among patients with chronic low back pain. *The International Journal of Psychiatry in Medicine,* 49**,** 187-198.
Iasp 1979. IASP Pain Terms. *Pain***,** 247-252.
Iasp. 2010. Global Year Against Acute Pain. *Acute Pain Fact Sheets* [Online]. Available: https://s3.amazonaws.com/rdcms-iasp/files/production/public/Content/ContentFolders/GlobalYearAgainstPain2/AcutePainFactSheets/1-Problem.pdf [Accessed 02/05/2020].
Iasp. 2017. *IASP Terminology* [Online]. USA: Intrernational Association for the Study of Pain. Available: https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698#Peripheralsensitization [Accessed 15/05/2020 2020].
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**,** 2098-2107.
Izumi, M., Petersen, K. K., Laursen, M. B., Arendt-Nielsen, L. & Graven-Nielsen, T. 2017. Facilitated temporal summation of pain correlates with clinical pain intensity after hip arthroplasty. *Pain,* 158**,** 323-332.
Jafari, M. & Ansari-Pour, N. 2019. Why, When and How to Adjust Your P Values? *Cell Journal (Yakhteh),* 20**,** 604.

Janssen, M., Pickard, A. S., Golicki, D., Gudex, C., Niewada, M., Scalone, L., Swinburn, P. & Busschbach, J. 2013. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Quality of Life Research,* 22**,** 1717-1727.

Jensen, M. P., Turner, J. A., Romano, J. M. & Karoly, P. 1991. Coping with chronic pain: a critical review of the literature. *Pain,* 47**,** 249-283.

Jensen, M. P., Turner, J. A., Romano, J. M. & Strom, S. E. 1995. The chronic pain coping inventory: development and preliminary validation. *Pain,* 60**,** 203-216.

Jordan, K., Jinks, C. & Croft, P. 2006. Health care utilization: measurement using primary care records and patient recall both showed bias. *Journal of clinical epidemiology,* 59**,** 791-797. e2.

Jorge, L. L., Gerard, C. & Revel, M. 2009. Evidence of memory dysfunction and maladaptive coping in chronic low back pain and rheumatoid arthritis patients: challenges for rehabilitation. *Eur J Phys Rehabil Med,* 45**,** 469-477.

Juch, J. N., Maas, E. T., Ostelo, R. W., Groeneweg, J. G., Kallewaard, J.-W., Koes, B. W., Verhagen, A. P., Van Dongen, J. M., Huygen, F. J. & Van Tulder, M. W. 2017. Effect of radiofrequency denervation on pain intensity among patients with chronic low back pain: the mint randomized clinical trials. *Jama,* 318**,** 68-81.

Jull, G., Kenardy, J., Hendrikz, J., Cohen, M. & Sterling, M. 2013. Management of acute whiplash: A randomized controlled trial of multidisciplinary stratified treatments. *Pain* [Online], 154. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/783/CN-00918783/frame.html.

Jull, G., Sterling, M., Kenardy, J. & Beller, E. 2007. Does the presence of sensory hypersensitivity influence outcomes of physical rehabilitation for chronic whiplash?--A preliminary RCT. *Pain* [Online], 129. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/676/CN-00577676/frame.html.

Kattoor, J., Gizewski, E. R., Kotsis, V., Benson, S., Gramsch, C., Theysohn, N., Maderwald, S., Forsting, M., Schedlowski, M. & Elsenbruch, S. 2013. Fear conditioning in an abdominal pain model: neural responses during associative learning and extinction in healthy subjects. *PloS one,* 8.

Katz, M. H. 2011. *Multivariable analysis: a practical guide for clinicians and public health researchers*, Cambridge university press.

Keedy, N. H., Keffala, V. J., Altmaier, E. M. & Chen, J. J. 2014. Health locus of control and self-efficacy predict back pain rehabilitation outcomes. *The Iowa orthopaedic journal,* 34**,** 158.

Keefe, F. J., Lefebvre, J. C. & Starr, K. R. From the gate control theory to the neuromatrix: Revolution or evolution? Pain Forum, 1996. Elsevier, 143-146.

Keene, O. N. 1995. The log transformation is special. *Statistics in medicine,* 14**,** 811-819.

Kennedy, D. L., Kemp, H. I., Ridout, D., Yarnitsky, D. & Rice, A. S. 2016. Reliability of conditioned pain modulation: a systematic review. *Pain,* 157**,** 2410.

Kenshalo Jr, D. & Isensee, O. 1983. Responses of primate SI cortical neurons to noxious stimuli. *Journal of neurophysiology,* 50**,** 1479-1496.

Kidd, B. & Urban, L. 2001. Mechanisms of inflammatory pain. *British journal of anaesthesia,* 87**,** 3-11.

Kim, J.-Y., Kim, S.-H., Seo, J., Kim, S.-H., Han, S. W., Nam, E. J., Kim, S.-K., Lee, H. J., Lee, S.-J. & Kim, Y.-T. 2013. Increased power spectral density in resting-state pain-related brain networks in fibromyalgia. *PAIN®,* 154**,** 1792-1797.

Klede, M., Handwerker, H. O. & Schmelz, M. 2003. Central origin of secondary mechanical hyperalgesia. *Journal of neurophysiology,* 90**,** 353-359.

Knofczynski, G. T. & Mundfrom, D. 2008. Sample sizes when using multiple linear regression for prediction. *Educational and Psychological Measurement,* 68**,** 431-442.
Koes, B., Van Tulder, M. & Thomas, S. 2006. Diagnosis and treatment of low back pain. *Bmj,* 332**,** 1430-1434.
Koleck, M., Mazaux, J.-M., Rascle, N. & Bruchon-Schweitzer, M. 2006. Psycho-social factors and coping strategies as predictors of chronic evolution and quality of life in patients with low back pain: a prospective study. *European Journal of Pain,* 10**,** 1-11.
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**,** 79-88.
Kong, J., Kaptchuk, T. J., Polich, G., Kirsch, I. & Gollub, R. L. 2007. Placebo analgesia: findings from brain imaging studies and emerging hypotheses. *Reviews in the Neurosciences,* 18**,** 173-190.
Krumova, E. K., Geber, C., Westermann, A. & Maier, C. 2012. Neuropathic pain: is quantitative sensory testing helpful? *Current diabetes reports,* 12**,** 393-402.
Krupp, L. B., Larocca, N. G., Muir-Nash, J. & Steinberg, A. D. 1989. The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of neurology,* 46**,** 1121-1123.
Lackner, J. M., Carosella, A. M. & Feuerstein, M. 1996. Pain expectancies, pain, and functional self-efficacy expectancies as determinants of disability in patients with chronic low back disorders. *Journal of consulting and clinical psychology,* 64**,** 212.
Lamotte, R. H., Shain, C. N., Simone, D. A. & Tsai, E. 1991. Neurogenic hyperalgesia: psychophysical studies of underlying mechanisms. *Journal of neurophysiology,* 66**,** 190-211.
Latremoliere, A. & Woolf, C. J. 2009. Central sensitization: a generator of pain hypersensitivity by central neural plasticity. *The Journal of Pain,* 10**,** 895-926.
Laupacis, A., Sekar, N. & Stiell, I. G. 1997. Clinical prediction rules - A review and suggested modifications of methodological standards. *Jama-Journal of the American Medical Association,* 277**,** 488-494.
Lavand'homme, P. 2011. The progression from acute to chronic pain. *Current Opinion in Anesthesiology,* 24**,** 545-550.
Lawrance, L. & Mcleroy, K. R. 1986. Self-efficacy and health education. *Journal of School Health,* 56**,** 317-321.
Lawrence, I. & Lin, K. 1992. Assay validation using the concordance correlation coefficient. *Biometrics***,** 599-604.
Lazarus, R. S. & Folkman, S. 1984. *Stress, appraisal, and coping*, Springer publishing company.
Learmonth, Y., Dlugonski, D., Pilutti, L., Sandroff, B., Klaren, R. & Motl, R. W. 2013. Psychometric properties of the fatigue severity scale and the modified fatigue impact scale. *Journal of the neurological sciences,* 331**,** 102-107.
Lechner, S. G. & Lewin, G. R. 2009. Peripheral sensitisation of nociceptors via G-proteindependent potentiation of mechanotransduction currents. *The Journal of physiology,* 587**,** 3493-3503.
Lee, J., Ellis, B., Price, C. & Baranowski, A. 2013. Chronic widespread pain, including fibromyalgia: a pathway for care developed by the British Pain Society. *British journal of anaesthesia,* 112**,** 16-24.
Lee, M. C., Zambreanu, L., Menon, D. K. & Tracey, I. 2008. Identifying brain activity specifically related to the maintenance and perceptual consequence of central sensitization in humans. *Journal of Neuroscience,* 28**,** 11642-11649.
Leeuw, M., Goossens, M. E., Linton, S. J., Crombez, G., Boersma, K. & Vlaeyen, J. W. 2007. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. *Journal of behavioral medicine,* 30**,** 77-94.

Leino, P. & Magni, G. 1993. Depressive and distress symptoms as predictors of low back pain, neck-shoulder pain, and other musculoskeletal morbidity: a 10-year follow-up of metal industry employees. *Pain,* 53**,** 89-94.
Lemay, K. R., Tulloch, H. E., Pipe, A. L. & Reed, J. L. 2019. Establishing the minimal clinically important difference for the hospital anxiety and depression scale in patients with cardiovascular disease. *Journal of cardiopulmonary rehabilitation and prevention,* 39**,** E6-E11.
Leresche, L., Turner, J. A., Saunders, K., Shortreed, S. M. & Von Korff, M. 2013. Psychophysical tests as predictors of back pain chronicity in primary care. *The journal of pain : official journal of the American Pain Society,* 14**,** 1663-70.
Lethem, J., Slade, P., Troup, J. & Bentley, G. 1983. Outline of a fear-avoidance model of exaggerated pain perception—I. *Behaviour research and therapy,* 21**,** 401-408.
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**,** 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**,** 936-944.
Li, Z., Wang, J., Chen, L., Zhang, M. & Wan, Y. 2013. Basolateral amygdala lesion inhibits the development of pain chronicity in neuropathic pain rats. *PloS one,* 8.
Liao, J. J. & Lewis, J. W. 2000. A note on concordance correlation coefficient. *PDA journal of pharmaceutical science and technology,* 54**,** 23-26.
Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J. & Moher, D. 2009. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *Annals of Internal Medicine,* 151**,** W65-W94.
Liddle, S. D., Baxter, G. D. & Gracey, J. H. 2009. Physiotherapists' use of advice and exercise for the management of chronic low back pain: a national survey. *Manual therapy,* 14**,** 189-196.
Lim, H. S., Chen, P. P., Wong, T. C., Gin, T., Wong, E., Chan, I. S. & Chu, J. 2007. Validation of the Chinese version of pain self-efficacy questionnaire. *Anesthesia & Analgesia,* 104**,** 918-923.
Linton, S. J. & Shaw, W. S. 2011. Impact of psychological factors in the experience of pain. *Physical therapy,* 91**,** 700-711.
Lluch, E., Nijs, J., Courtney, C. A., Rebbeck, T., Wylde, V., Baert, I., Wideman, T. H., Howells, N. & Skou, S. T. 2018. Clinical descriptors for the recognition of central sensitization pain in patients with knee osteoarthritis. *Disability and rehabilitation,* 40**,** 2836-2845.
Lluch, E., Torres, R., Nijs, J. & Van Oosterwijck, J. 2014a. Central sensitization in patients with osteoarthritis pain: A systematic literature review. *Pain Practice,* 14**,** 34-35.
Lluch, E., Torres, R., Nijs, J. & Van Oosterwijck, J. 2014b. Evidence for central sensitization in patients with osteoarthritis pain: A systematic literature review. *European Journal of Pain (United Kingdom),* 18**,** 1367-1375.
Lluch Girbés, E., Duenas, L., Barbero, M., Falla, D., Baert, I. A., Meeus, M., Sánchez-Frutos, J., Aguilella, L. & Nijs, J. 2016. Expanded distribution of pain as a sign of central sensitization in individuals with symptomatic knee osteoarthritis. *Physical therapy,* 96**,** 1196-1207.
Lorig, K. 2003. Self-management education: more than a nice extra. LWW.
Lorig, K., Ritter, P. L. & Plant, K. 2005. A disease-specific self-help program compared with a generalized chronic disease self-help program for arthritis patients.

Arthritis Care & Research: Official Journal of the American College of Rheumatology, 53**,** 950-957.
Lorig, K. R. & Holman, H. R. 2003. Self-management education: history, definition, outcomes, and mechanisms. *Annals of behavioral medicine,* 26**,** 1-7.
Lorig, K. R., Ritter, P., Stewart, A. L., Sobel, D. S., Brown Jr, B. W., Bandura, A., Gonzalez, V. M., Laurent, D. D. & Holman, H. R. 2001a. Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Medical care,* 39**,** 1217-1223.
Lorig, K. R., Ritter, P. L., Laurent, D. D. & Plant, K. 2006. Internet-based chronic disease self-management: a randomized trial. *Medical care***,** 964-971.
Lorig, K. R., Sobel, D. S., Ritter, P. L., Laurent, D. & Hobbs, M. 2001b. Effect of a self-management program on patients with chronic disease. *Effective clinical practice: ECP,* 4**,** 256-262.
Lorig, K. R., Sobel, D. S., Stewart, A. L., Brown Jr, B. W., Bandura, A., Ritter, P., Gonzalez, V. M., Laurent, D. D. & Holman, H. R. 1999. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Medical care,* 37**,** 5-14.
Luna, I. E., Kehlet, H., Petersen, M. A. & Aasvang, E. K. 2017. Clinical, nociceptive and psychological profiling to predict acute pain after total knee arthroplasty. *Acta anaesthesiologica Scandinavica,* 61**,** 676-687.
Lundblad, H., Kreicbergs, A. & Jansson, K.-Å. 2008. Prediction of persistent pain after total knee replacement for osteoarthritis. *Bone & Joint Journal,* 90**,** 166-171.
Macdonald, R. R. 2002. Missing data--Quantitative applications in the social sciences. *British Journal of Mathematical & Statistical Psychology,* 55**,** 193.
Machado, G. C., Maher, C. G., Ferreira, P. H., Pinheiro, M. B., Lin, C.-W. C., Day, R. O., Mclachlan, A. J. & Ferreira, M. L. 2015. Efficacy and safety of paracetamol for spinal pain and osteoarthritis: systematic review and meta-analysis of randomised placebo controlled trials. *Bmj,* 350**,** h1225.
Mackinnon, D. 2012. *Introduction to statistical mediation analysis*, Routledge.
Mallen, C. D., Peat, G., Thomas, E., Dunn, K. M. & Croft, P. R. 2007. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract,* 57**,** 655-661.
Maniadakis, N. & Gray, A. 2000. The economic burden of back pain in the UK. *Pain,* 84**,** 95-103.
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.
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**,** 83-92.
Manusov, E. G. 2012. Evaluation and diagnosis of low back pain. *Primary Care: Clinics in Office Practice,* 39**,** 471-479.
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.
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**,** 1757-1768.
Martinez, V., Fletcher, D., Bouhassira, D., Sessler, D. I. & Chauvin, M. 2007. The evolution of primary hyperalgesia in orthopedic surgery: Quantitative sensory testing and clinical evaluation before and after total knee arthroplasty. *Anesthesia and Analgesia,* 105**,** 815-821.

Maunsell, E., Lauzier, S., Brunet, J., Pelletier, S., Osborne, R. H. & Campbell, H. S. 2014. Health-related empowerment in cancer: Validity of scales from the Health Education Impact Questionnaire. *Cancer,* 120**,** 3228-3236.
May, S. 2010. Self-management of chronic low back pain and osteoarthritis. *Nature Reviews Rheumatology,* 6**,** 199-209.
Mayer, T. G., Neblett, R., Cohen, H., Howard, K. J., Choi, Y. H., Williams, M. J., Perez, Y. & Gatchel, R. J. 2012. The development and psychometric validation of the central sensitization inventory. *Pain Practice,* 12**,** 276-285.
Mcbeth, J. & Jones, K. 2007. Epidemiology of chronic musculoskeletal pain. *Best practice & research Clinical rheumatology,* 21**,** 403-425.
Mcginn, T., Wyer, P., Wisnivesky, J., Devereaux, P., Stiell, I., Richardson, S. & Guyatt, G. 2008. Advanced topics in diagnosis: clinical prediction rules. *Users' Guides to the Medical Literature: Manual for Evidence-Based Clinical Practice, 2nd edn. Columbus, OH: McGraw-Hill***,** 491-505.
Mckenzie, R. 1981. The lumbar spine: mechanical diagnosis and therapy. *New Zealand: Spinal Publications***,** 164.
Mcphee, M. E., Vaegter, H. B. & Graven-Nielsen, T. 2020. Alterations in pronociceptive and antinociceptive mechanisms in patients with low back pain: a systematic review with meta-analysis. *Pain,* 161**,** 464-475.
Meeus, M. & Nijs, J. 2007. Central sensitization: a biopsychosocial explanation for chronic widespread pain in patients with fibromyalgia and chronic fatigue syndrome. *Clinical rheumatology,* 26**,** 465-473.
Meier, P. M., Berde, C. B., Dicanzio, J., Zurakowski, D. & Sethna, N. F. 2001. Quantitative assessment of cutaneous thermal and vibration sensation and thermal pain detection thresholds in healthy children and adolescents. *Muscle & nerve,* 24**,** 1339-1345.
Melzack, R. 2001. Pain and the neuromatrix in the brain. *Journal of dental education,* 65**,** 1378-1382.
Melzack, R. & Casey, K. L. 1968. Sensory, motivational, and central control determinants of pain: a new conceptual model. *The skin senses,* 1**,** 423-43.
Melzack, R. & Wall, P. D. 1965. Pain mechanisms: a new theory. *Science,* 150**,** 971-979.
Mendonca, M., Simis, M., Grecco, L., Battistella, L., Baptista, A. & Fregni, F. 2016. Transcranial direct current stimulation combined with aerobic exercise to optimize analgesic responses in fibromyalgia: A randomized placebo-controlled clinical trial. *Frontiers in human neuroscience* [Online], 10. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/406/CN-01141406/frame.html.
Mercado, A. C. 2004. Coping behaviour and outcome in two pain populations: A general adult population with neck or low back pain and individuals suffering from pain due to motor-vehicle injuries. *Dissertation Abstracts International: Section B: The Sciences and Engineering,* 64**,** 3534.
Merskey, H. E. 1986. Classification of chronic pain: Descriptions of chronic pain syndromés and definitions of pain terms. *Pain*.
Meyer, K., Sprott, H. & Mannion, A. F. 2008. Cross-cultural adaptation, reliability, and validity of the German version of the Pain Catastrophizing Scale. *Journal of Psychosomatic Research,* 64**,** 469-478.
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**,** 55.
Michopoulos, I., Douzenis, A., Kalkavoura, C., Christodoulou, C., Michalopoulou, P., Kalemi, G., Fineti, K., Patapis, P., Protopapas, K. & Lykouras, L. 2008. Hospital Anxiety and Depression Scale (HADS): validation in a Greek general hospital sample. *Annals of general psychiatry,* 7**,** 4.

Miles, C. L., Pincus, T., Carnes, D., Homer, K. E., Taylor, S. J., Bremner, S. A., Rahman, A. & Underwood, M. 2011. Can we identify how programmes aimed at promoting self-management in musculoskeletal pain work and who benefits? A systematic review of sub-group analysis within RCTs. *European journal of pain,* 15**,** 775. e1-775. e11.

Millan, M. J. 2002. Descending control of pain. *Progress in neurobiology,* 66**,** 355-474.

Miller, R. P., Kori, S. H. & Todd, D. D. 1991. The Tampa Scale: a Measure of Kinisophobia. *The Clinical journal of pain,* 7**,** 51.

Miller, S. M. 2012. Low back pain: pharmacologic management. *Primary Care: Clinics in Office Practice,* 39**,** 499-510.

Miller, W. R. & Seligman, M. E. 1975. Depression and learned helplessness in man. *Journal of abnormal psychology,* 84**,** 228.

Miró, J., De La Vega, R., Gertz, K. J., Jensen, M. P. & Engel, J. M. 2019. The role of perceived family social support and parental solicitous responses in adjustment to bothersome pain in young people with physical disabilities. *Disability and rehabilitation,* 41**,** 641-648.

Mlekusch, S., Schliessbach, J., Camara, R. J. A., Arendt-Nielsen, L., Juni, P. & Curatolo, M. 2013. Do central hypersensitivity and altered pain modulation predict the course of chronic low back and neck pain? *The Clinical journal of pain,* 29**,** 673-80.

Moayedi, M. & Davis, K. D. 2013. Theories of pain: from specificity to gate control. *Journal of neurophysiology,* 109**,** 5-12.

Monticone, M., Ferrante, S., Giorgi, I., Galandra, C., Rocca, B. & Foti, C. 2013. Development of the Italian version of the 42-item Chronic Pain Coping Inventory, CPCI-I: Cross-cultural adaptation, factor analysis, reliability and validity. *Quality of Life Research,* 22**,** 1459-1465.

Moons, K. G. M., Royston, P., Vergouwe, Y., Grobbee, D. E. & Altman, D. G. 2009. Prognosis and prognostic research: what, why, and how? *British Medical Journal,* 338.

Moreton, B. J., Tew, V., Das Nair, R., Wheeler, M., Walsh, D. A. & Lincoln, N. B. 2015. Pain phenotype in patients with knee osteoarthritis: classification and measurement properties of painDETECT and self-report Leeds Assessment of Neuropathic Symptoms and Signs Scale in a cross-sectional study. *Arthritis care & research,* 67**,** 519-528.

Moriarty, O., Mcguire, B. E. & Finn, D. P. 2011. The effect of pain on cognitive function: a review of clinical and preclinical research. *Progress in neurobiology,* 93**,** 385-404.

Morita, R., Arakida, M., Osborne, R. H., Nolte, S., Elsworth, G. R. & Mikami, H. 2013. Adaptation and validation of the Japanese version of the Health Education Impact Questionnaire (heiQ-J) for the evaluation of self-management education interventions. *Japan Journal of Nursing Science,* 10**,** 255-266.

Morlion, B. 2011. Pharmacotherapy of low back pain: targeting nociceptive and neuropathic pain components. *Current medical research and opinion,* 27**,** 11-33.

Morris, A. L. 2004. Patients' perspectives on self-management following a back rehabilitation programme. *Musculoskeletal care,* 2**,** 165-179.

Morsø, L., Kent, P., Albert, H. B., Hill, J. C., Kongsted, A. & Manniche, C. 2013. The predictive and external validity of the STarT Back Tool in Danish primary care. *European Spine Journal,* 22**,** 1859-1867.

Moss, P., Knight, E. & Wright, A. 2016. Subjects with knee osteoarthritis exhibit widespread hyperalgesia to pressure and cold. *PloS one,* 11.

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**,** 894-907.
Murphy, S. L., Lyden, A. K., Phillips, K., Clauw, D. J. & Williams, D. A. 2011. Subgroups of older adults with osteoarthritis based upon differing comorbid symptom presentations and potential underlying pain mechanisms. *Arthritis research & therapy,* 13**,** 1.
Mutschler, I., Ball, T., Wankerl, J. & Strigo, I. A. 2012. Pain and emotion in the insular cortex: evidence for functional reorganization in major depression. *Neuroscience letters,* 520**,** 204-209.
Neblett, R., Cohen, H., Choi, Y., Hartzell, M. M., Williams, M., Mayer, T. G. & Gatchel, R. J. 2013. The Central Sensitization Inventory (CSI): establishing clinically significant values for identifying central sensitivity syndromes in an outpatient chronic pain sample. *The Journal of Pain,* 14**,** 438-445.
Newman, S., Steed, L. & Mulligan, K. 2004. Self-management interventions for chronic illness. *The Lancet,* 364**,** 1523-1537.
Neziri, A. Y., Curatolo, M., Limacher, A., Nüesch, E., Radanov, B., Andersen, O. K., Arendt-Nielsen, L. & Jüni, P. 2012. Ranking of parameters of pain hypersensitivity according to their discriminative ability in chronic low back pain. *PAIN®,* 153**,** 2083-2091.
Neziri, A. Y., Scaramozzinol, P., Andersenl, O. K., Dickensonl, A. H., Arendt-Nielsenl, L. & Curatolol, M. 2011. Reference values of mechanical and thermal pain tests in a pain-free population. *European journal of pain,* 15**,** 376-383.
Nice. 2014. Osteoarthritis: care and management in adults. *Clinical guideline* [Online]. Available: www.nice.org.uk/guidance/cg177 [Accessed 17/06/2018].
Nice 2016. Low Back Pain and Sciatica in Over 16s: Assessment and Management.
Nicholas, M. 1994. Pain self efficacy questionnaire (PSEQ). *Summary of psychometric properties from original Westmead hospital sample (1986e87). Sydney: University of Sydney Pain Management and Research Centre*.
Nicholas, M., Asghari, A., Blyth, F., Wood, B., Murray, R., Mccabe, R., Brnabic, A., Beeston, L., Corbett, M. & Sherrington, C. 2016. Long-term outcomes from training in self-management of chronic pain in an elderly population: A randomised controlled trial. *Pain*.
Nicholas, M. K. 2007a. Mental disorders in people with chronic pain: an international perspective. LWW.
Nicholas, M. K. 2007b. The pain self-efficacy questionnaire: Taking pain into account. *European journal of pain,* 11**,** 153-163.
Nicholas, M. K., Asghari, A., Blyth, F. M., Wood, B. M., Murray, R., Mccabe, R., Brnabic, A., Beeston, L., Corbett, M. & Sherrington, C. 2013. Self-management intervention for chronic pain in older adults: a randomised controlled trial. *PAIN®,* 154**,** 824-835.
Nih. 2020. *Low Back Pain Fact Sheet* [Online]. USA: National Institutes of Neurological Disorders and Stroke,. Available: https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Low-Back-Pain-Fact-Sheet [Accessed 25/04/2020 2020].
Nijs, J., De Kooning, M., Beckwee, D. & Vaes, P. 2015. The neurophysiology of pain and pain modulation: Modern pain neuroscience for musculoskeletal physiotherapists. *In:* Jull, G., Moore, A., Falla, D., Lewis, J., Mccarthy, C. & Sterling, M. (eds.) *Grieve's Modern Musculoskeletal Physiotherapy: Vertebral Column and Peripheral Joints.* 4th ed. UK: Elsevier Health Sciences.
Nijs, J., Meeus, M., Van Oosterwijck, J., Roussel, N., De Kooning, M., Ickmans, K. & Matic, M. 2011. Treatment of central sensitization in patients with 'unexplained'chronic pain: what options do we have? *Expert opinion on pharmacotherapy,* 12**,** 1087-1098.

Nijs, J., Van Houdenhove, B. & Oostendorp, R. A. 2010. Recognition of central sensitization in patients with musculoskeletal pain: application of pain neurophysiology in manual therapy practice. *Manual therapy,* 15**,** 135-141.
Nijs, J., Van Oosterwijck, J. & De Hertogh, W. 2009. Rehabilitation of chronic whiplash: treatment of cervical dysfunctions or chronic pain syndrome? *Clinical rheumatology,* 28**,** 243-251.
Noiseux, N. O., Callaghan, J. J., Clark, C. R., Zimmerman, M. B., Sluka, K. A. & Rakel, B. A. 2014. Preoperative predictors of pain following total knee arthroplasty. *The Journal of arthroplasty,* 29**,** 1383-1387.
Nolan, C. M., Longworth, L., Lord, J., Canavan, J. L., Jones, S. E., Kon, S. S. & Man, W. D. 2016. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. *Thorax,* 71**,** 493-500.
Nolte, S., Elsworth, G. R., Newman, S. & Osborne, R. H. 2013. Measurement issues in the evaluation of chronic disease self-management programs. *Quality of Life Research,* 22**,** 1655-1664.
Nolte, S. & Osborne, R. H. 2013. A systematic review of outcomes of chronic disease self-management interventions. *Quality of Life Research,* 22**,** 1805-1816.
Norman, G. R., Sloan, J. A. & Wyrwich, K. W. 2003. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical care***,** 582-592.
Norman, P. & Conner, M. 1996. The role of social cognition models in predicting health behaviours: Future directions. *Predicting health behaviour: Research and practice with social cognition models***,** 197-225.
Nussbaum, E. L. & Downes, L. 1998. Reliability of clinical pressure-pain algometric measurements obtained on consecutive days. *Physical therapy,* 78**,** 160-169.
O'brien, A. T., Deitos, A., Pego, Y. T., Fregni, F. & Carrillo-De-La-Peña, M. T. 2018. Defective endogenous pain modulation in fibromyalgia: a meta-analysis of temporal summation and conditioned pain modulation paradigms. *The Journal of Pain,* 19**,** 819-836.
O'leary, H., Smart, K. M., Moloney, N. A., Blake, C. & Doody, C. M. 2018. Pain sensitization associated with nonresponse after physiotherapy in people with knee osteoarthritis. *Pain,* 159**,** 1877-1886.
O'leary, H., Smart, K. M., Moloney, N. A. & Doody, C. M. 2017. Nervous system sensitization as a predictor of outcome in the treatment of peripheral musculoskeletal conditions: a systematic review. *Pain Practice,* 17**,** 249-266.
O'sullivan, P. 2012. It's time for change with the management of non-specific chronic low back pain. *British Journal of Sports Medicine***,** 224-227.
O'sullivan, P., Dankaerts, W., O'sullivan, K. & Kjartan, F. 2015. Multidimesional approach for the targeted management of low back pain. *In:* Jull, G., Moore, A., Falla, D., Lewis, J., Mccarthy, C. & Sterling, M. (eds.) *Grieve's Modern Musculoskeletal Physiotherapy: Vertebral Column and Peripheral Joints.* 4th ed. UK: Elsevier Health Sciences.
O'neill, T. W. & Felson, D. T. 2018. Mechanisms of osteoarthritis (OA) pain. *Current osteoporosis reports,* 16**,** 611-616.
Olesen, S. S., Graversen, C., Bouwense, S. A., Van Goor, H., Wilder-Smith, O. H. & Drewes, A. M. 2013. Quantitative sensory testing predicts pregabalin efficacy in painful chronic pancreatitis. *PloS one,* 8**,** e57963.
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.
Oliveira, V. C., Ferreira, P. H., Maher, C. G., Pinto, R. Z., Refshauge, K. M. & Ferreira, M. L. 2012. Effectiveness of self-management of low back pain: Systematic review with meta-analysis. *Arthritis care & research,* 64**,** 1739-1748.
Ono, R., Higashi, T., Suzukamo, Y., Konno, S., Takahashi, O., Tokuda, Y., Rhaman, M., Shimbo, T., Endo, H. & Hinohara, S. 2008. Higher internality of health

locus of control is associated with the use of complementary and alternative medicine providers among patients seeking care for acute low-back pain. *The Clinical journal of pain,* 24**,** 725-730.

Osborne, R. H., Elsworth, G. R. & Whitfield, K. 2007. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient education and counseling,* 66**,** 192-201.

Osman, A., Barrios, F. X., Kopper, B. A., Hauptmann, W., Jones, J. & O'neill, E. 1997. Factor structure, reliability, and validity of the Pain Catastrophizing Scale. *Journal of behavioral medicine,* 20**,** 589-605.

Ossipov, M. H., Dussor, G. O. & Porreca, F. 2010. Central modulation of pain. *The Journal of clinical investigation,* 120**,** 3779-3787.

Overmier, J. B. 2002. Sensitization, conditioning, and learning: can they help us understand somatization and disability? *Scandinavian journal of psychology,* 43**,** 105-112.

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**,** 1452-1464.

Paasche-Orlow, M. K., Schillinger, D., Greene, S. M. & Wagner, E. H. 2006. How health care systems can begin to address the challenge of limited literacy. *Journal of General Internal Medicine,* 21**,** 884-887.

Paasche-Orlow, M. K. & Wolf, M. S. 2007. The causal pathways linking health literacy to health outcomes. *American journal of health behavior,* 31**,** S19-S26.

Panagioti, M., Richardson, G., Small, N., Murray, E., Rogers, A., Kennedy, A., Newman, S. & Bower, P. 2014. Self-management support interventions to reduce health care utilisation without compromising outcomes: a systematic review and meta-analysis. *BMC health services research,* 14**,** 1.

Panksepp, J. 2003. Feeling the pain of social loss. *Science,* 302**,** 237-239.

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**,** 412-417.

Parker, S. L., Godil, S. S., Zuckerman, S. L., Mendenhall, S. K., Devin, C. J. & Mcgirt, M. J. 2015. Extent of preoperative depression is associated with return to work after lumbar fusion for spondylolisthesis. *World neurosurgery,* 83**,** 608-613.

Parsons, S., Ingram, M., Clarke-Cornwell, A. & Symmons, D. 2011. A Heavy Burden: the occurrence and impact of musculoskeletal conditions in the United Kingdom today.

Patel, A., Buszewicz, M., Beecham, J., Griffin, M., Rait, G., Nazareth, I., Atkinson, A., Barlow, J. & Haines, A. 2009. Economic evaluation of arthritis self management in primary care. *BMJ,* 339**,** b3532.

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.

Pavlaković, G. & Petzke, F. 2010. The role of quantitative sensory testing in the evaluation of musculoskeletal pain conditions. *Current rheumatology reports,* 12**,** 455-461.

Peacock, S. & Patel, S. 2008. Cultural influences on pain. *Reviews in pain,* 1**,** 6-9.

Pearl, J. & Bareinboim, E. 2014. External validity: From do-calculus to transportability across populations. *Statistical Science,* 29**,** 579-595.

Pedler, A., Kamper, S. J. & Sterling, M. 2016. Addition of posttraumatic stress and sensory hypersensitivity more accurately estimates disability and pain than fear avoidance measures alone after whiplash injury. *Pain,* 157**,** 1645-1654.

Pelletier, R., Higgins, J. & Bourbonnais, D. 2015. Is neuroplasticity in the central nervous system the missing link to our understanding of chronic musculoskeletal disorders? *BMC musculoskeletal disorders,* 16**,** 25.

Peters, M. L., Schmidt, A. J., Van Den Hout, M. A., Koopmans, R. & Sluijter, M. E. 1992. Chronic back pain, acute postoperative pain and the activation of diffuse noxious inhibitory controls (DNIC). *Pain,* 50**,** 177-187.

Petersen-Felix, S., Arendt-Nielsen, L., Bak, P., Fischer, M., Bjerring, P. & Zbinden, A. M. 1996. The effects of isoflurane on repeated nociceptive stimuli (central temporal summation). *Pain,* 64**,** 277-281.

Petersen, K. K., Arendt-Nielsen, L., Simonsen, O., Wilder-Smith, O. & Laursen, M. B. 2015. Presurgical assessment of temporal summation of pain predicts the development of chronic postoperative pain 12 months after total knee replacement. *Pain,* 156**,** 55-61.

Petersen, K. K., Graven-Nielsen, T., Simonsen, O., Laursen, M. B. & Arendt-Nielsen, L. 2016. Preoperative pain mechanisms assessed by cuff algometry are associated with chronic postoperative pain relief after total knee replacement. *Pain,* 157**,** 1400-1406.

Pettersson, S., Lundberg, I. E., Liang, M. H., Pouchot, J. & Welin Henriksson, E. 2015. Determination of the minimal clinically important difference for seven measures of fatigue in Swedish patients with systemic lupus erythematosus. *Scandinavian journal of rheumatology,* 44**,** 206-210.

Peyron, R., Laurent, B. & Garcia-Larrea, L. 2000. Functional imaging of brain responses to pain. A review and meta-analysis (2000). *Neurophysiologie Clinique/Clinical Neurophysiology,* 30**,** 263-288.

Phillips, K. & Clauw, D. J. 2011. Central pain mechanisms in chronic pain states–maybe it is all in their head. *Best Practice & Research Clinical Rheumatology,* 25**,** 141-154.

Picavet, H. & Schouten, J. 2003. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC 3-study. *Pain,* 102**,** 167-178.

Pigg, M., Baad-Hansen, L., Svensson, P., Drangsholt, M. & List, T. 2010. Reliability of intraoral quantitative sensory testing (QST). *Pain,* 148**,** 220-226.

Pinto, R. Z., Maher, C. G., Ferreira, M. L., Hancock, M., Oliveira, V. C., Mclachlan, A. J., Koes, B. & Ferreira, P. H. 2012. Epidural corticosteroid injections in the management of sciatica: a systematic review and meta-analysis. *Annals of internal medicine,* 157**,** 865-877.

Poetscher, A. W., Gentil, A. F., Lenza, M. & Ferretti, M. 2014. Radiofrequency denervation for facet joint low back pain: a systematic review. *Spine,* 39**,** E842-E849.

Portney, L. G. & Watkins, M. P. 2009. Foundations of clinical research: applications to practice.

Price, D. D., Staud, R., Robinson, M. E., Mauderli, A. P., Cannon, R. & Vierck, C. J. 2002. Enhanced temporal summation of second pain and its central modulation in fibromyalgia patients. *Pain,* 99**,** 49-59.

Promis. 2020. *Intro to PROMIS*® [Online]. USA: Northwestern University,. Available: http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis [Accessed 26/04/2020 2020].

Pssru. 2020. *Client Service Receipt Inventory (CSRI)* [Online]. University of Kent. Available: https://www.pssru.ac.uk/csri/what-is-the-csri/ [Accessed 18/04/2020 2020].

Puta, C., Schulz, B., Schoeler, S., Magerl, W., Gabriel, B., Gabriel, H. H., Miltner, W. H. & Weiss, T. 2012. Enhanced sensitivity to punctate painful stimuli in female patients with chronic low back pain. *BMC neurology,* 12**,** 98.

R Core Team 2017. R: A language and environment for statistical computing. 3.4.2 ed. Vienna, Austria: R Foundation for Statistical Computing.

Rabey, M., Poon, C., Wray, J., Thamajaree, C., East, R. & Slater, H. 2015. Pro-nociceptive and anti-nociceptive effects of a conditioned pain modulation protocol in participants with chronic low back pain and healthy control subjects. *Manual therapy,* 20**,** 763-768.

Rahim-Williams, B., Riley Iii, J. L., Williams, A. K. & Fillingim, R. B. 2012. A quantitative review of ethnic group differences in experimental pain response: do biology, psychology, and culture matter? *Pain Medicine,* 13**,** 522-540.

Rakel, B. A., Blodgett, N. P., Zimmerman, M. B., Logsden-Sackett, N., Clark, C., Noiseux, N., Callaghan, J., Herr, K., Geasland, K. & Yang, X. 2012. Predictors of postoperative movement and resting pain following total knee replacement. *PAIN®,* 153**,** 2192-2203.

Ramond, A., Bouton, C., Richard, I., Roquelaure, Y., Baufreton, C., Legrand, E. & Huez, J.-F. 2010. Psychosocial risk factors for chronic low back pain in primary care—a systematic review. *Family practice***,** cmq072.

Randolph, A. G., Guyatt, G. H., Calvin, J. E., Doig, G. & Richardson, W. S. 1998. Understanding articles describing clinical prediction tools. *Critical Care Medicine,* 26**,** 1603-1612.

Reid, M. C., Papaleontiou, M., Ong, A., Breckman, R., Wethington, E. & Pillemer, K. 2008. Self-Management strategies to reduce pain and improve function among older adults in community settings: A review of the evidence. *Pain Medicine,* 9**,** 409-424.

Reilly, B. M. & Evans, A. T. 2006. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Annals of internal medicine,* 144**,** 201-209.

Rempe, T., Wolff, S., Riedel, C., Baron, R., Stroman, P. W., Jansen, O. & Gierthmühlen, J. 2014. Spinal fMRI reveals decreased descending inhibition during secondary mechanical hyperalgesia. *PloS one,* 9**,** e112325.

Revicki, D., Hays, R. D., Cella, D. & Sloan, J. 2008. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of clinical epidemiology,* 61**,** 102-109.

Rhee, H., Belyea, M. J., Ciurzynski, S. & Brasch, J. 2009. Barriers to asthma self-management in adolescents: Relationships to psychosocial factors. *Pediatric pulmonology,* 44**,** 183-191.

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**,** 244-253.

Richard, A. A. & Shea, K. 2011. Delineation of Self-Care and Associated Concepts. *Journal of Nursing Scholarship,* 43**,** 255-264.

Riddle, D. L., Wade, J. B., Jiranek, W. A. & Kong, X. 2010. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. *Clinical Orthopaedics and Related Research®,* 468**,** 798-806.

Riegel, B., Dickson, V. V., Goldberg, L. R. & Deatrick, J. A. 2007. Factors associated with the development of expertise in heart failure self-care. *Nursing research,* 56**,** 235-243.

Roberts, K. L. & Allen, H. A. 2016. Perception and cognition in the ageing brain: a brief review of the short-and long-term links between perceptual and cognitive decline. *Frontiers in aging neuroscience,* 8**,** 39.

Rodgers, B. L. 2005. *Developing nursing knowledge: Philosophical traditions and influences*, Lippincott Williams & Wilkins.

Roland, M. & Morris, R. 1983. A study of the natural history of back pain: part I: development of a reliable and sensitive measure of disability in low-back pain. *spine,* 8**,** 141-144.

Rolke, R., Baron, R., Maier, C. A., Tölle, T., Treede, R.-D., Beyer, A., Binder, A., Birbaumer, N., Birklein, F. & Bötefür, I. 2006a. Quantitative sensory testing in

the German Research Network on Neuropathic Pain (DFNS): standardized protocol and reference values. *Pain,* 123**,** 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**,** 77-77.
Romano, J. M., Jensen, M. P. & Turner, J. A. 2003. The chronic pain coping inventory-42: reliability and validity. *Pain,* 104**,** 65-73.
Rosomoff, H. L. & Rosomoff, R. S. 1999. Low back pain: evaluation and management in the primary care setting. *Medical Clinics of North America,* 83**,** 643-662.
Rotter, J. B. 1966. Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied,* 80**,** 1.
Ryan, P. & Sawin, K. J. 2009. The individual and family self-management theory: Background and perspectives on context, process, and outcomes. *Nursing outlook,* 57**,** 217-225. e6.
Sahrmann, S. 2010. *Movement System Impairment Syndromes of the Extremities, Cervical and Thoracic Spines-E-Book*, Elsevier Health Sciences.
Salisbury, C., Chalder, M., Scott, T. M., Pope, C. & Moore, L. 2002. What is the role of walk-in centres in the NHS? *Bmj,* 324**,** 399-402.
Salomons, T. V., Moayedi, M., Erpelding, N. & Davis, K. D. 2014. A brief cognitive-behavioural intervention for pain reduces secondary hyperalgesia. *PAIN®,* 155**,** 1446-1452.
Salvetti, M., Pimenta, C., Braga, P. E. & Correa, C. F. 2012. Disability related to chronic low back pain: prevalence abd associated factors. *Revista da Escola de Enfermagem da USP*.
Samad, T. A., Moore, K. A., Sapirstein, A., Billet, S., Allchorne, A., Poole, S., Bonventre, J. V. & Woolf, C. J. 2001. Interleukin-1β-mediated induction of Cox-2 in the CNS contributes to inflammatory pain hypersensitivity. *Nature,* 410**,** 471-475.
Samwel, H. J., Evers, A. W., Crul, B. J. & Kraaimaat, F. W. 2006. The role of helplessness, fear of pain, and passive pain-coping in chronic pain patients. *The Clinical journal of pain,* 22**,** 245-251.
Sandrini, G., Alfonsi, E., Bono, G., Facchinetti, F., Montalbetti, L. & Nappi, G. 1986. Circadian variations of human flexion reflex. *Pain,* 25**,** 403-410.
Sangesland, A., Støren, C. & Vaegter, H. B. 2017. Are preoperative experimental pain assessments correlated with clinical pain outcomes after surgery? A systematic review. *Scandinavian journal of pain,* 15**,** 44-52.
Sarzi-Puttini, P., Atzeni, F. & Mease, P. J. 2011. Chronic widespread pain: from peripheral to central evolution. *Best Practice & Research Clinical Rheumatology,* 25**,** 133-139.
Sawin, K. J., Bellin, M. H., Roux, G., Buran, C. F. & Brei, T. J. 2009. The experience of self-management in adolescent women with spina bifida. *Rehabilitation Nursing,* 34**,** 26-38.
Schafer, J. L. 1997. *Analysis of incomplete multivariate data*, CRC press.
Schaible, H.-G. 2006. Peripheral and central mechanisms of pain generation. *Analgesia.* Springer.
Schliessbach, J., Arendt-Nielsen, L., Heini, P. & Curatolo, M. 2010. The role of central hypersensitivity in the determination of intradiscal mechanical hyperalgesia in discogenic pain. *Pain medicine,* 11**,** 701-708.
Schliessbach, J., Siegenthaler, A., Streitberger, K., Eichenberger, U., Nüesch, E., Jüni, P., Arendt-Nielsen, L. & Curatolo, M. 2013. The prevalence of widespread central hypersensitivity in chronic pain patients. *European journal of pain,* 17**,** 1502-1510.
Schmid, J., Theysohn, N., Gaß, F., Benson, S., Gramsch, C., Forsting, M., Gizewski, E. R. & Elsenbruch, S. 2013. Neural mechanisms mediating positive and negative treatment expectations in visceral pain: a functional magnetic

resonance imaging study on placebo and nocebo effects in healthy volunteers. *PAIN®,* 154**,** 2372-2380.
Schug, S. 2011. 2011–The global year against acute pain. SAGE Publications Sage UK: London, England.
Schuler, M., Musekamp, G., Bengel, J., Schwarze, M., Spanier, K., Gutenbrunner, C., Ehlebracht-König, I., Nolte, S., Osborne, R. & Faller, H. 2014. Measurement of stable changes of self-management skills after rehabilitation: a latent state–trait analysis of the Health Education Impact Questionnaire (heiQ™). *Quality of life research,* 23**,** 2531-2543.
Schuler, M., Musekamp, G., Faller, H., Ehlebracht-König, I., Gutenbrunner, C., Kirchhof, R., Bengel, J., Nolte, S., Osborne, R. H. & Schwarze, M. 2013. Assessment of proximal outcomes of self-management programs: translation and psychometric evaluation of a German version of the Health Education Impact Questionnaire (heiQ™). *Quality of life research,* 22**,** 1391-1403.
Schünemann, H., Hill, S., Guyatt, G., Akl, E. A. & Ahmed, F. 2011. The GRADE approach and Bradford Hill's criteria for causation. *Journal of Epidemiology & Community Health,* 65**,** 392-395.
Seifert, F. & Maihöfner, C. 2009. Central mechanisms of experimental and chronic neuropathic pain: findings from functional imaging studies. *Cellular and molecular life sciences,* 66**,** 375-390.
Sengul, Y., Kara, B. & Arda, M. N. 2010. The relationship between health locus of control and quality of life in patients with chronic low back pain. *Turkish neurosurgery,* 20**,** 180-185.
Shadish, W. R., Cook, T. D. & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference/William R. Shedish, Thomas D. Cook, Donald T. Campbell*, Boston: Houghton Mifflin.
Shaheed, C. A., Maher, C. G., Williams, K. A., Day, R. & Mclachlan, A. J. 2016. Efficacy, tolerability, and dose-dependent effects of opioid analgesics for low back pain: a systematic review and meta-analysis. *JAMA internal medicine,* 176**,** 958-968.
Shekelle, P. G., Maglione, M. A., Chodosh, J., Mojica, W., Morton, S. C., Booth, M., Roth, E., Hilton, L., Rhodes, S. & Wu, S.-Y. 2003. Chronic disease self management for diabetes, osteoarthritis, post-myocardial infarction care, and hypertension.
Sheng, J., Liu, S., Wang, Y., Cui, R. & Zhang, X. 2017. The link between depression and chronic pain: Neural mechanisms in the brain. *Neural plasticity,* 2017.
Simons, L. E. & Blount, R. L. 2007. Identifying barriers to medication adherence in adolescent transplant recipients. *Journal of Pediatric Psychology,* 32**,** 831-844.
Simons, L. E., Moulton, E. A., Linnman, C., Carpino, E., Becerra, L. & Borsook, D. 2014. The human amygdala and pain: evidence from neuroimaging. *Human brain mapping,* 35**,** 527-538.
Slade, G. D., Sanders, A. E., Ohrbach, R., Fillingim, R. B., Dubner, R., Gracely, R. H., Bair, E., Maixner, W. & Greenspan, J. D. 2014. Pressure pain thresholds fluctuate with, but do not usefully predict, the clinical course of painful temporomandibular disorder. *PAIN®,* 155**,** 2134-2143.
Smart, K. M., Blake, C., Staines, A. & Doody, C. 2011. The Discriminative validity of "nociceptive,""peripheral neuropathic," and "central sensitization" as mechanisms-based classifications of musculoskeletal pain. *The Clinical journal of pain,* 27**,** 655-663.
Smart, K. M., Blake, C., Staines, A. & Doody, C. 2012a. Self-reported pain severity, quality of life, disability, anxiety and depression in patients classified with 'nociceptive','peripheral neuropathic'and 'central sensitisation'pain. The discriminant validity of mechanisms-based classifications of low back (±leg) pain. *Manual therapy,* 17**,** 119-125.

Smart, K. M., Blake, C., Staines, A., Thacker, M. & Doody, C. 2012b. Mechanisms-based classifications of musculoskeletal pain: part 1 of 3: symptoms and signs of central sensitisation in patients with low back (±leg) pain. *Manual therapy,* 17**,** 336-344.

Snelgrove, S. & Liossi, C. 2013. Living with chronic low back pain: a metasynthesis of qualitative research. *Chronic illness,* 9**,** 283-301.

Snijders, G. F., Van Den Ende, C. H., Fransen, J., Van Riel, P. L., Stukstette, M. J., Defoort, K. C., Arts-Sanders, M. A., Van Den Hoogen, F. H. & Den Broeder, A. A. 2011. Fatigue in knee and hip osteoarthritis: the role of pain and physical function. *Rheumatology,* 50**,** 1894-1900.

Solberg Nes, L., Roach, A. R. & Segerstrom, S. C. 2009. Executive functions, self-regulation, and chronic pain: a review. *Annals of Behavioral Medicine,* 37**,** 173-183.

Solomon, P. 2001. Congruence between health professionals' and patients' pain ratings: a review of the literature. *Scandinavian Journal of Caring Sciences,* 15**,** 174-180.

Sprague, S., Leece, P., Bhandari, M., Tornetta Iii, P., Schemitsch, E. & Swiontkowski, M. F. 2003. Limiting loss to follow-up in a multicenter randomized trial in orthopedic surgery. *Controlled clinical trials,* 24**,** 719-725.

Staud, R. 2002. Evidence of involvement of central neural mechanisms in generating fibromyalgia pain. *Current rheumatology reports,* 4**,** 299-305.

Staud, R., Robinson, M. E. & Price, D. D. 2005. Isometric exercise has opposite effects on central pain mechanisms in fibromyalgia patients compared to normal controls. *Pain,* 118**,** 176-184.

Sterling, M., Hendrikz, J. & Kenardy, J. 2011. Similar factors predict disability and posttraumatic stress disorder trajectories after whiplash injury. *Pain,* 152**,** 1272-1278.

Sterling, M., Hendrikz, J., Kenardy, J., Kristjansson, E., Dumas, J. P., Niere, K., Cote, J., Deserres, S., Rivest, K. & Jull, G. 2012. Assessment and validation of prognostic models for poor functional recovery 12 months after whiplash injury: A multicentre inception cohort study. *Pain,* 153**,** 1727-1734.

Sterling, M., Jull, G., Vicenzino, B. & Kenardy, J. 2003a. Sensory hypersensitivity occurs soon after whiplash injury and is associated with poor recovery. *Pain,* 104**,** 509-517.

Sterling, M., Jull, G., Vicenzino, B. & Kenardy, J. 2003b. Sensory hypersensitivity occurs soon after whiplash injury and is associated with poor recovery. *Pain,* 104**,** 509-17.

Sterling, M., Jull, G., Vicenzino, B., Kenardy, J. & Darnell, R. 2005. Physical and psychological factors predict outcome following whiplash injury. *Pain,* 114**,** 141-148.

Suarez-Almazor, M. E., Conner-Spady, B., Kendall, C. J., Russell, A. S. & Skeith, K. 2001. Lack of congruence in the ratings of patients' health status by patients and their physicians. *Medical Decision Making,* 21**,** 113-121.

Sullivan, M., Tanzer, M., Stanish, W., Fallaha, M., Keefe, F. J., Simmonds, M. & Dunbar, M. 2009. Psychological determinants of problematic outcomes following total knee arthroplasty. *Pain,* 143**,** 123-129.

Sullivan, M. J., Adams, H., Martel, M. O., Scott, W. & Wideman, T. 2011. Catastrophizing and perceived injustice: risk factors for the transition to chronicity after whiplash injury. *Spine,* 36**,** S244-249.

Sullivan, M. J., Bishop, S. R. & Pivik, J. 1995. The pain catastrophizing scale: development and validation. *Psychological assessment,* 7**,** 524.

Suokas, A., Walsh, D., Mcwilliams, D., Condon, L., Moreton, B., Wylde, V., Arendt-Nielsen, L. & Zhang, W. 2012. Quantitative sensory testing in painful osteoarthritis: a systematic review and meta-analysis. *Osteoarthritis and Cartilage,* 20**,** 1075-1085.

Swendeman, D., Ingram, B. L. & Rotheram-Borus, M. J. 2009. Common elements in self-management of HIV and other chronic illnesses: an integrative framework. *AIDS care,* 21**,** 1321-1334.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science,* 240**,** 1285-1293.

Swinkels-Meewisse, E., Swinkels, R., Verbeek, A., Vlaeyen, J. & Oostendorp, R. 2003. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Manual therapy,* 8**,** 29-36.

Tampin, B., Bohne, T., Callan, M., Kvia, M., Melsom Myhre, A., Neoh, E., Bharat, C. & Slater, H. 2017. Reliability of the English version of the painDETECT questionnaire. *Current medical research and opinion,* 33**,** 741-748.

Taylor, J. J., Borckardt, J. J. & George, M. S. 2012. Endogenous opioids mediate left dorsolateral prefrontal cortex rTMS-induced analgesia. *Pain,* 153**,** 1219-1225.

Tesarz, J., Eich, W., Treede, R.-D. & Gerhardt, A. 2016. Altered pressure pain thresholds and increased wind-up in adult patients with chronic back pain with a history of childhood maltreatment: a quantitative sensory testing study. *Pain,* 157**,** 1799-1809.

Thomazeau, J., Rouquette, A., Martinez, V., Rabuel, C., Prince, N., Laplanche, J. L., Nizard, R., Bergmann, J. F., Perrot, S. & Lloret-Linares, C. 2016. Acute pain Factors predictive of post-operative pain and opioid requirement in multimodal analgesia following knee replacement. *European journal of pain (London, England),* 20**,** 822-32.

Tracey, I. 2005. Nociceptive processing in the human brain. *Current opinion in neurobiology,* 15**,** 478-487.

Traxler, J., Hanssen, M. M., Lautenbacher, S., Ottawa, F. & Peters, M. L. 2018. General versus pain-specific cognitions: Pain catastrophizing but not optimism influences conditioned pain modulation. *European Journal of Pain*.

Treede, R.-D., Kenshalo, D. R., Gracely, R. H. & Jones, A. K. 1999. The cortical representation of pain. *Pain,* 79**,** 105-111.

Treede, R.-D. & Magerl, W. 2000. Multiple mechanisms of secondary hyperalgesia. *Progress in brain research.* Elsevier.

Trochim, W. M. 2006. Theory of reliability.

Truchon, M. 2001. Determinants of chronic disability related to low back pain: towards an integrative biopsychosocial model. *Disability and Rehabilitation,* 23**,** 758-767.

Tunks, E., Mccain, G., Hart, L. A., Teasell, R. W., Goldsmith, C. H., Rollman, G., Mcdermid, A. & Deshane, P. 1995. The reliability of examination for tenderness in patients with myofascial pain, chronic fibromyalgia and controls. *J Rheumatol,* 22**,** 944-952.

Turk, D. C. 2002. A diathesis-stress model of chronic pain and disability following traumatic injury. *Pain Research and Management,* 7**,** 9-19.

Tzellos, T., Toulis, K., Goulis, D., Papazisis, G., Zampeli, V., Vakfari, A. & Kouvelas, D. 2010. Gabapentin and pregabalin in the treatment of fibromyalgia: a systematic review and a meta-analysis. *Journal of clinical pharmacy and therapeutics,* 35**,** 639-656.

Uddin, Z. & Macdermid, J. C. 2016. Quantitative Sensory Testing in Chronic Musculoskeletal Pain. *Pain Medicine***,** pnv105.

Uddin, Z., Macdermid, J. C., Galea, V., Gross, A. R. & Pierrynowski, M. R. 2014. The current perception threshold test differentiates categories of mechanical neck disorder. *Journal of Orthopaedic & Sports Physical Therapy,* 44**,** 532-C1.

Uddin, Z., Macdermid, J. C., Moro, J., Galea, V. & Gross, A. R. 2016. Psychophysical and Patient Factors as Determinants of Pain, Function and Health Status in Shoulder Disorders. *The Open Orthopaedics Journal,* 10.

Ursin, H. & Eriksen, H. R. 2001. Sensitization, subjective health complaints, and sustained arousal. *Annals of the New York Academy of Sciences,* 933**,** 119-129.

Urwin, M., Symmons, D., Allison, T., Brammah, T., Busby, H., Roxby, M., Simmons, A. & Williams, G. 1998. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Annals of the rheumatic diseases,* 57**,** 649-655.

Usui, C., Hatta, K., Aratani, S., Yagishita, N., Nishioka, K., Kanazawa, T., Itoh, K., Yamano, Y., Nakamura, H. & Nakajima, T. 2013. The Japanese version of the modified ACR preliminary diagnostic criteria for fibromyalgia and the fibromyalgia symptom scale: reliability and validity. *Modern rheumatology,* 23**,** 846-850.

Vaegter, H. B., Handberg, G., Emmeluth, C. & Graven-Nielsen, T. 2016. Preoperative Hypoalgesia after Cold Pressor Test and Aerobic Exercise is Associated with Pain Relief Six Months after Total Knee Replacement. *Clinical Journal of Pain. (no pagination), 2016. Date of Publication: 12 Aug 2016.* [Online]. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/807/CN-01195807/frame.html.

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**,** 775.

Van Der Hulst, M., Vollenbroek-Hutten, M. M., Groothuis-Oudshoorn, K. G. & Hermens, H. J. 2008. Multidisciplinary rehabilitation treatment of patients with chronic low back pain: a prognostic model for its outcome. *The Clinical journal of pain,* 24**,** 421-430.

Van Hout, B., Janssen, M., Feng, Y.-S., Kohlmann, T., Busschbach, J., Golicki, D., Lloyd, A., Scalone, L., Kind, P. & Pickard, A. S. 2012. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value in health,* 15**,** 708-715.

Vandenbroucke, J. P. 2008. Observational research, randomised trials, and two views of medical science. *PLoS medicine,* 5.

Veldhuijzen, D., Kenemans, J., Van Wijck, A., Olivier, B., Kalkman, C. & Volkerts, E. 2006. Processing capacity in chronic pain patients: a visual event-related potentials study. *Pain,* 121**,** 60-68.

Verdejo-García, A., López-Torrecillas, F., Calandre, E. P., Delgado-Rodríguez, A. & Bechara, A. 2009. Executive function and decision-making in women with fibromyalgia. *Archives of Clinical Neuropsychology,* 24**,** 113-122.

Villemure, C. & Bushnell, M. C. 2009. Mood influences supraspinal pain processing separately from attention. *Journal of Neuroscience,* 29**,** 705-715.

Vlaeyen, J. W. & Linton, S. J. 2012. Fear-avoidance model of chronic musculoskeletal pain: 12 years on. *Pain,* 153**,** 1144-1147.

Vohs, K. D. & Baumeister, R. F. 2016. *Handbook of self-regulation: Research, theory, and applications*, Guilford Publications.

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., Carter, A., Casey, D. C., Charlson, F. J. & Chen, A. Z. 2016. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet,* 388**,** 1545-1602.

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., Charlson, F., Davis, A., Degenhardt, L. & Dicker, D. 2015. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet,* 386**,** 743-800.

Vuilleumier, P. H., Arguissain, F. G., Manresa, J. a. B., Neziri, A. Y., Nirkko, A. C., Andersen, O. K., Arendt-Nielsen, L. & Curatolo, M. 2017. Psychophysical and electrophysiological evidence for enhanced pain facilitation and unaltered pain inhibition in acute low back pain patients. *The journal of pain,* 18**,** 1313-1323.
Walker, A. C., Tan, L. & George, S. 1995. Impact of culture on pain management: an Australian nursing perspective. *Holistic nursing practice,* 9**,** 48-57.
Walsh, D. A. & Radcliffe, J. C. 2002. Pain beliefs and perceived physical disability of patients with chronic low back pain. *Pain,* 97**,** 23-31.
Walter, S., Eliasziw, M. & Donner, A. 1998. Sample size and optimal designs for reliability studies. *Statistics in medicine,* 17**,** 101-110.
Walton, D. M., Macdermid, J. C., Nielson, W., Teasell, R. W., Reese, H. & Levesque, L. 2011. Pressure pain threshold testing demonstrates predictive ability in people with acute whiplash. *The Journal of orthopaedic and sports physical therapy,* 41**,** 658-65.
Warsi, A., Wang, P. S., Lavalley, M. P., Avorn, J. & Solomon, D. H. 2004. Self-management education programs in chronic disease: a systematic review and methodological critique of the literature. *Archives of internal medicine,* 164**,** 1641-1649.
Werner, M. U., Duun, P. & Kehlet, H. 2004. Prediction of postoperative pain by preoperative nociceptive responses to heat stimulation. *Anesthesiology,* 100**,** 115-9; discussion 5A.
Werner, M. U., Mjöbo, H. N., Nielsen, P. R. & Rudin, Å. 2010. Prediction of Postoperative Pain: A Systematic Review of Predictive Experimental Pain Studies. *Anesthesiology: The Journal of the American Society of Anesthesiologists,* 112**,** 1494-1502.
White, A. P., Arnold, P. M., Norvell, D. C., Ecker, E. & Fehlings, M. G. 2011. Pharmacologic management of chronic low back pain: synthesis of the evidence. *Spine,* 36**,** S131-S143.
Who 2003. *The burden of musculoskeletal conditions at the start of the new millennium: report of a WHO Scientific Group*, World Health Organization.
Wideman, T. H., Asmundson, G. G., Smeets, R. J. M., Zautra, A. J., Simmonds, M. J., Sullivan, M. J., Haythornthwaite, J. A. & Edwards, R. R. 2013. Re-thinking the fear avoidance model: toward a multi-dimensional framework of pain-related disability. *Pain,* 154**,** 2262.
Wijma, A. J., Van Wilgen, C. P., Meeus, M. & Nijs, J. 2016. Clinical biopsychosocial physiotherapy assessment of patients with chronic pain: The first step in pain neuroscience education. *Physiotherapy theory and practice,* 32**,** 368-384.
Wilder-Smith, O. H., Schreyer, T., Scheffer, G. J. & Arendt-Nielsen, L. 2010. Patients with chronic pain after abdominal surgery show less preoperative endogenous pain inhibition and more postoperative hyperalgesia: a pilot study. *Journal of pain & palliative care pharmacotherapy,* 24**,** 119-128.
Williams, A. C. D. C., Davies, H. T. O. & Chadury, Y. 2000. Simple pain rating scales hide complex idiosyncratic meanings. *Pain,* 85**,** 457-463.
Williamson, A. & Hoggart, B. 2005. Pain: a review of three commonly used pain rating scales. *Journal of clinical nursing,* 14**,** 798-804.
Wolfe, F., Butler, S. H., Fitzcharles, M., Häuser, W., Katz, R. L., Mease, P. J., Rasker, J. J., Russell, A. S., Russell, I. J. & Walitt, B. 2019. Revised chronic widespread pain criteria: development from and integration with fibromyalgia criteria. *Scandinavian journal of pain*.
Wolfe, F., Clauw, D. J., Fitzcharles, M.-A., Goldenberg, D. L., Häuser, W., Katz, R. L., Mease, P. J., Russell, A. S., Russell, I. J. & Walitt, B. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Seminars in arthritis and rheumatism, 2016. Elsevier, 319-329.
Wolfe, F., Clauw, D. J., Fitzcharles, M.-A., Goldenberg, D. L., Häuser, W., Katz, R. S., Mease, P., Russell, A. S., Russell, I. J. & Winfield, J. B. 2011. Fibromyalgia

criteria and severity scales for clinical and epidemiological studies: a modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. *The Journal of rheumatology,* 38**,** 1113-1122.
Woolf, C. J. 2011. Central sensitization: implications for the diagnosis and treatment of pain. *Pain,* 152**,** S2-S15.
Woolf, C. J. & Salter, M. W. 2000. Neuronal plasticity: increasing the gain in pain. *science,* 288**,** 1765-1768.
Wylde, V., Hewlett, S., Learmonth, I. D. & Dieppe, P. 2011a. Persistent pain after joint replacement: prevalence, sensory qualities, and postoperative determinants. *PAIN®,* 152**,** 566-572.
Wylde, V., Palmer, S., Learmonth, I. & Dieppe, P. 2011b. Test–retest reliability of Quantitative Sensory Testing in knee osteoarthritis and healthy participants. *Osteoarthritis and Cartilage,* 19**,** 655-658.
Wylde, V., Palmer, S., Learmonth, I. & Dieppe, P. 2013. The association between pre-operative pain sensitisation and chronic pain after knee replacement: an exploratory study. *Osteoarthritis and Cartilage,* 21**,** 1253-1256.
Wylde, V., Sayers, A., Lenguerrand, E., Gooberman-Hill, R., Pyke, M., Beswick, A., Dieppe, P. & Blom, A. 2015. Preoperative widespread pain sensitization and chronic pain after hip and knee replacement: a cohort analysis. *Pain* [Online], 156. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/024/CN-01111024/frame.html.
Yanagisawa, K., Masui, K., Furutani, K., Nomura, M., Ura, M. & Yoshida, H. 2011. Does higher general trust serve as a psychosocial buffer against social pain? An NIRS study of social exclusion. *Social Neuroscience,* 6**,** 190-197.
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**,** 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**,** 805-806.
Yarnitsky, D., Crispel, Y., Eisenberg, E., Granovsky, Y., Ben-Nun, A., Sprecher, E., Best, L.-A. & Granot, M. 2008. Prediction of chronic post-operative pain: pre-operative DNIC testing identifies patients at risk. *Pain,* 138**,** 22-28.
Yarnitsky, D. & Granot, M. 2006. Quantitative sensory testing. *Handbook of clinical neurology,* 81**,** 397-409.
Yazdi-Ravandi, S., Taslimi, Z., Jamshidian, N., Saberi, H., Shams, J. & Haghparast, A. 2013. Prediction of quality of life by self-efficacy, pain intensity and pain duration in patient with pain disorders. *Basic and clinical neuroscience,* 4**,** 117.
Yousef, A. & Al-Deeb, A. 2013. A double-blinded randomised controlled study of the value of sequential intravenous and oral magnesium therapy in patients with chronic low back pain with a neuropathic component. *Anaesthesia,* 68**,** 260-266.
Yu, C.-Y. 2002. *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes*, University of California, Los Angeles Los Angeles.
Ziegler, E., Magerl, W., Meyer, R. & Treede, R.-D. 1999. Secondary hyperalgesia to punctate mechanical stimuli: central sensitization to A-fibre nociceptor input. *Brain,* 122**,** 2245-2257.
Zigmond, A. S. & Snaith, R. P. 1983. The hospital anxiety and depression scale. *Acta psychiatrica scandinavica,* 67**,** 361-370.
Zusman, M. 2002. Forebrain-mediated sensitization of central pain pathways:'non-specific'pain and a new image for MT. *Manual therapy,* 7**,** 80-88.

Zusman, M. 2004. Mechanisms of musculoskeletal physiotherapy. *Physical Therapy Reviews,* 9**,** 39-49.

APPENDICES

Appendix 1. OVID & MEDLINE search strategy sample

1.	exp back pain/
2.	neck pain/
3.	Shoulder pain/
4.	Tennis Elbow/
5.	exp Tendinopathy/
6.	Whiplash Injuries/
7.	Sciatica/
8.	Intervertebral Disk Displacement/
9.	(pain adj3 (neck or back or shoulder? or elbow? or forearm? or wrist? or hand? or arm? or hip? or knee? or ankle? or leg? or foot or feet)).tw.
10.	(epicondylitis or tendonitis or tendinitis or bursitis or synovitis or sprain? or strain?).tw.
11.	(Whiplash or sciatica).tw.
12.	exp Joint Diseases/
13.	exp Spinal Diseases/
14.	exp Spondylarthritis/
15.	exp Arthritis, Rheumatoid/
16.	exp Osteoarthritis/
17.	osteoarthritis.mp or exp Osteoarthritis, Hip/ or exp Osteoarthritis/ or exp Osteoarthritis, Spine/ or exp Osteoarthritis, Knee/
18.	osteoarthrosis.mp.
19.	gonarthritis.mp.
20.	gonarthrosis.mp.
21.	gonitis.mp.
22.	knee pain.mp.
23.	exp Osteophyte/ or osteophyte*.mp
24.	(joint space adj6 narrow*).tw.
25.	(arthriti$ or osteoarthriti$ or osteoporo$ or bone loss$).tw.
26.	17 or 18 or 19 or 20 or 21 or 22 or 23 or 24
27.	exp Osteoporosis/
28.	ankle/ or hip/ or knee/
29.	elbow/ or wrist/ or shoulder/
30.	elbow joint/ or exp hand joints/ or hip joint/ or knee joint/ or sacroiliac joint/ or shoulder joint/
31.	exp cervical vertebrae/ or intervertebral disk/ or lumbar vertebrae/ or thoracic vertebrae/
32.	exp Back/
33.	exp Spine/
34.	(spine or spinal).tw.
35.	or/1-35
36.	quantitative sensory.mp.
37.	exp Sensory Thresholds/
38.	exp Pain Thresholds/
39.	threshold*.mp.
40.	36 or 37 or 38 or 39
41.	35 and 40

Appendix 2. Study details of associations between baseline QST and musculoskeletal pain outcomes

Author	Study Design	Sample size	Diagnosis	Site of Pathology	Baseline QST Predictor	Outcome Variable	Outcome Measures	Statistical Analysis Method	Association Modalities	*(r)* / *(ß)* / (AUC) / (χ^2) values	*p*-values	OR	95% CI Lower Limit	95% CI Higher Limit	SE
Arendt-Nielsen et al., 2016	RCT	37	OA	Knee	TS	Pain	BPI	Correlation analysis	TS to Pain (30% cut off)	**-0.639**	0.008				
					TS	Pain	BPI	*r*	TS to Pain (50% cut off)	**-0.421**	0.032				
					PPT	Pain	BPI	*r*	PPT to Pain	No Data Given	No Data Given				
					PPT	Pain	BPI	*r*	PPT to Pain	No Data Given	No Data Given				
Bar Ziv et al., 2016	Prospective Cohort	48	OA	Knee	PPI	Pain	KSS	Correlation analysis	PPI to Pain	**-0.335**	0.019				
Coombes et al., 2015	Prospective Cohort (D.A.)	41	Epicondylitis	Elbow	CPT	Pain	PRTEE	Multivariate Linear Regression	CPT to Pain (2months) [Unadjusted]	**0.77**	0.008		0.210	1.330	
					CPT	Pain	PRTEE	*ß*	CPT to Pain (2months) [Adjusted]	**0.364**	0.008		0.095	0.633	0.1373
					CPT	Pain	PRTEE	*ß*	CPT to Pain (1 year) [Unadjusted]	**0.61**	0.034		0.050	1.170	
					CPT	Pain	PRTEE	*ß*	CPT to Pain (1 year) [Adjusted]	**0.335**	0.034		0.025	0.645	0.1580
Coronado et al., 2015a	RCT	63	Shoulder Pain	Shoulder	PPT	Pain	BPI	Correlation analysis	Average PPT(acr) to Pain(4w)	**0.12**	>0.05				
					PPT	Pain	BPI	*r*	Average PPT(acr) to Pain(8w)	**-0.01**	>0.05				
					PPT	Pain	BPI	*r*	Average PPT(acr) to Pain(12w)	**0.06**	>0.05				

					PPT	Pain	BPI	*r*	Average PPT(TA) to Pain(4w)	**-0.07**	>0.05			
					PPT	Pain	BPI	*r*	Average PPT(TA) to Pain(8w)	**-0.06**	>0.05			
					PPT	Pain	BPI	*r*	Average PPT(TA) to Pain(12w)	**0.01**	>0.05			
					HPT	Pain	BPI	*r*	AverageHPT to Pain(4w)	**-0.08**	>0.05			
					HPT	Pain	BPI	*r*	Average HPT to Pain(8w)	**-0.06**	>0.05			
					HPT	Pain	BPI	*r*	Average HPT to Pain(12w)	**-0.06**	>0.05			
Coronado et al., 2015b	Prospective Cohort	68	LBP	Low Back	PPT	Pain	BPI	Multivariate Logisitc Regression	PPT (6weeks) to Pain (6months)	No Data Given	<0.05	**2.03**	1.02	4.04
Davis et al., 2013	Prospective Cohort	31	Shoulder Pain	Shoulder	EPT	Pain	VAS	Multivariate Linear Regression	PPT (Low) to Pain (Rest)	**4.9**	0.007		1.5	8.4
								AUC	PPT (Low) to Pain (Movement)	**4.1**	0.049		0.03	8.2
Edwards et al., 2016	Prospective Cohort (C.C.)	35	OA	Knee	CPM	Pain	DPI	Correlation analysis	CPM to Pain (DPI) [Change]	**-0.38**	<0.05			
					CPM	Pain	KOOS	*r*	CPM to Pain (KOOS) [Change]	**0.45**	<0.01			
					CPM	Pain	DPI	Stepwise Linear Regression	CPM to Pain (DPI) [End point]	**-0.30**	0.03			
					CPM	Pain	KOOS	*β*	CPM to Pain (KOOS) [End point]	**0.37**	0.03			
Goodin et al., 2014	Prospective Cohort	225	OA	Knee	TS	Pain	NRS	Correlation analysis	Mech TS (Knee) to Pain (Average)	**0.21**	<0.01			
					TS	Pain	NRS	*r*	Mech TS (Knee) to Pain (Worst)	**0.17**	<0.05			

TS	Pain	NRS	*r*	Mech TS (Hand) to Pain (Average)	**0.22**	<0.01
TS	Pain	NRS	*r*	Mech TS (Hand) to Pain (Worst)	**0.14**	<0.05
TS	Pain	NRS	*r*	Heat TS (Knee) [44°C] to Pain (Average)	**0.08**	>0.05
TS	Pain	NRS	*r*	Heat TS (Knee) [46°C] to Pain (Average)	**0.10**	>0.05
TS	Pain	NRS	*r*	Heat TS (Knee) [48°C] to Pain (Average)	**0.03**	>0.05
TS	Pain	NRS	*r*	Heat TS (Knee) [44°C] to Pain (Worst)	**0.11**	>0.05
TS	Pain	NRS	*r*	Heat TS (Knee) [46°C] to Pain (Worst)	**0.07**	>0.05
TS	Pain	NRS	*r*	Heat TS (Knee) [48°C] to Pain (Worst)	**0.08**	>0.05
TS	Pain	NRS	*r*	Heat TS (Forearm) [44°C] to Pain (Average)	**0.11**	>0.05
TS	Pain	NRS	*r*	Heat TS (Forearm) [46°C] to Pain (Average)	**0.16**	<0.05
TS	Pain	NRS	*r*	Heat TS (Forearm) [48°C] to Pain (Average)	**0.08**	>0.05
TS	Pain	NRS	*r*	Heat TS (Forearm) [44°C] to Pain (Worst)	**0.10**	>0.05
TS	Pain	NRS	*r*	Heat TS (Forearm)	**0.08**	>0.05

					TS	Pain	NRS	*r*	[46°C] to Pain (Worst)	**0.05**	>0.05	
					TS	Pain	NRS	*r*	Heat TS (Forearm) [48°C] to Pain (Worst)	**0.05**	>0.05	
					TS	Pain	NRS	Multivariate Linear Regression	Mech TS (Knee) [ethn adj] to Pain (Average)	**-0.28**	0.018	0.1184
					TS	Pain	NRS	*β*	Mech TS (Knee) [ethn adj] to Pain (Worst)	**-0.29**	0.046	0.1453
		103			TS	Pain	NRS	*β*	Mech TS (Knee) [nHW] to Pain (Average)	**0.02**	0.016	0.0083
		103			TS	Pain	NRS	*β*	Mech TS (Knee) [nHW] to Pain (Worst)	**0.02**	0.044	0.0099
		122			TS	Pain	NRS	*β*	Mech TS (Knee) [AfAm] to Pain (Average)	**0.01**	>0.30	
		122			TS	Pain	NRS	*β*	Mech TS (Knee) [AfAm] to Pain (Worst)	**0.01**	>0.30	
					TS	Pain	NRS	*β*	Mech TS (Hand) [ethn adj] to Pain (Average)	**-0.19**	0.084	0.1100
					TS	Pain	NRS	*β*	Mech TS (Hand) [ethn adj] to Pain (Worst)	**-0.23**	0.092	0.1365
Gwilym et al., 2011	Prospective Cohort (C.C.)	17	SI	Shoulder	MPT	Pain	OSS	Correlation analysis	MPT to Pain	**-0.03**	0.92	
Henriksen et al., 2014	RCT	48	OA	Knee	PPT	Pain	KOOS	Correlation analysis	PPT to Pain (Change)	**0.21**	0.15	
Izumi et al., 2017	Prospective Cohort (C.C.)	40	OA	Hip	TS	Pain	VAS	Correlation analysis	TS to Pain (Rest)	**0.44**	<0.05	
					TS	Pain	VAS	*r*	TS to Pain (Walk)	**0.44**	<0.05	
					TS	Pain	VAS	*r*	TS to Pain (Change)	**-0.52**	<0.05	

					PPT	Pain	VAS	*r*	PPT to Pain	No Data Given	No Data Given			
					cPPT	Pain	VAS	*r*	cPPT to Pain	No Data Given	No Data Given			
					cPTT	Pain	VAS	*r*	cPTT to Pain	No Data Given	No Data Given			
					SS	Pain	VAS	*r*	SS to Pain	No Data Given	No Data Given			
					PinPS	Pain	VAS	*r*	PinPS to Pain	No Data Given	No Data Given			
					CDT	Pain	VAS	*r*	CDT to Pain	No Data Given	No Data Given			
					CPT	Pain	VAS	*r*	CPT to Pain	No Data Given	No Data Given			
					WDT	Pain	VAS	*r*	WDT to Pain	No Data Given	No Data Given			
					CPM	Pain	VAS	*r*	CPM to Pain	No Data Given	No Data Given			
LeResche et al., 2013	Prospective Cohort	147	LBP	Low Back	PPT	Pain	NRS	Multivariate Logistic Regression	PPT (Back) to Pain	No Data Given	<0.05	**0.66**	0.44	0.96
					PPT	Pain	NRS	*OR*	PPT (Back) to Pain (age+sex adj)	No Data Given	>0.05	**0.72**	0.46	1.11
					PPT	Pain	NRS	*OR*	PPT (Thenar) to Pain	No Data Given	<0.05	**0.62**	0.40	0.92
					PPT	Pain	NRS	*OR*	PPT (Thenar) to Pain (age+sex adj)	No Data Given	>0.05	**0.70**	0.44	1.06
					CPM	Pain	NRS	*OR*	CPM to Pain	No Data Given	>0.05	**1.11**	0.77	1.62

					CPM	Pain	NRS	*OR*	CPM to Pain (age+sex adj)	No Data Given	>0.05	**1.07**	0.73	1.56	
					TS	Pain	NRS	*OR*	TS to Pain	No Data Given	>0.05	**0.92**	0.63	1.31	
					TS	Pain	NRS	*OR*	TS to Pain (age+sex adj)	No Data Given	>0.05	**0.88**	0.58	1.27	
					CPP	Pain	NRS	*OR*	CPP to Pain	No Data Given	>0.05	**1.04**	0.72	1.51	
					CPP	Pain	NRS	*OR*	CPP to Pain (age+sex adj)	No Data Given	>0.05	**0.91**	0.61	1.36	
Luna et al., 2017	Prospective Cohort	60	OA	Knee	PPT	Pain	NRS	Multivariate Logistic Regression	PPT to Pain	No Data Given	0.007	**0.99**	0.99	1.00	
					PPT	Pain	NRS	*OR*	PPT (20kPa increase) to Pain	No Data Given	0.007	**0.85**	0.76	0.96	
					PPT	Pain	NRS	*OR*	PPT (50kPa increase) to Pain	No Data Given	0.007	**0.67**	0.50	0.90	
Lundblad et al., 2008	Prospective Cohort	69	OA	Knee	EPT	Pain	VAS	Correlation analysis	EPT to Pain (Rest)	**6.34**	0.012				
					EST	Pain	VAS	χ^2	EST to Pain (Rest)	**4.00**	0.045				
					EPT	Pain	VAS	Logistic Regression Analysis	EPT to Pain (Rest)	**2.22**	0.01	**9.19**	1.69	50.07	0.87
Martinez et al., 2007	Prospective Cohort	20	OA	Knee	PPT	Pain	VAS	Correlation analysis	PPT to Pain	No Data Given	No Data Given				
					TA	Pain	VAS	*r*	TA to Pain	No Data Given	No Data Given				
					HPT	Pain	VAS	*r*	HPT to Pain	No Data Given	No Data Given				
					CPT	Pain	VAS	*r*	CPT to Pain	No Data Given	No Data Given				
Mendonca et al., 2016	RCT	32	Fibromyalgia	Widespread Pain	PPT	Pain	VNS	Correlation analysis	PPT to Pain [Change]	**-0.04**	0.7				

					PPT	Pain	VNS	Univariate Linear Regression	PPT to Pain [Change]	**0.07**	0.7			
Mlekusch et al., 2013	Prospective Cohort	169	LBP and Neck Pain	Low Back and Neck	PTT	Pain	BPI	Multivariate Linear Regression	PTT to Pain (Average) [crude]	**-0.02**	0.855	-0.23	0.19	0.107
					PTT	Pain	BPI	*β*	PTT to Pain (Average) [adjusted]	**-0.07**	0.524	-0.30	0.15	0.117
					CPM	Pain	BPI	*β*	CPM to Pain (Average) [crude]	**-0.11**	0.600	-0.51	0.29	0.204
					CPM	Pain	BPI	*β*	CPM to Pain (Average) [adjusted]	**-0.03**	0.903	-0.45	0.40	0.214
					CTT	Pain	BPI	*β*	CTT to Pain (Average) [crude]	**0.03**	0.727	-0.15	0.22	0.092
					CTT	Pain	BPI	*β*	CTT to Pain (Average) [adjusted]	**0.02**	0.851	-0.18	0.22	0.102
					PTT	Pain	BPI	*β*	PTT to Pain (Max) [crude]	**-0.04**	0.733	-0.29	0.20	0.128
					PTT	Pain	BPI	*β*	PTT to Pain (Max) [adjusted]	**-0.05**	0.709	-0.31	0.21	0.133
					CPM	Pain	BPI	*β*	CPM to Pain (Max) [crude]	**-0.23**	0.331	-0.71	0.24	0.245
					CPM	Pain	BPI	*β*	CPM to Pain (Max) [adjusted]	**-0.12**	0.632	-0.62	0.38	0.255
					CTT	Pain	BPI	*β*	CTT to Pain (Max) [crude]	**0.04**	0.725	-0.18	0.25	0.112
					CTT	Pain	BPI	*β*	CTT to Pain (Max) [adjusted]	**0.06**	0.622	-0.18	0.29	0.122
		113	LBP	Low Back	PTT	Pain	BPI	*β*	PTT to Pain (Average) [crude]	**0.03**	No Data Given	-0.21	0.28	0.122

			PTT	Pain	BPI	*β*	PTT to Pain (Average) [adjusted]	**-0.01**	No Data Given	-0.28	0.27	0.138
			CPM	Pain	BPI	*β*	CPM to Pain (Average) [crude]	**-0.40**	No Data Given	-0.80	0.00	0.204
			CPM	Pain	BPI	*β*	CPM to Pain (Average) [adjusted]	**-0.36**	No Data Given	-0.78	0.07	0.214
			CTT	Pain	BPI	*β*	CTT to Pain (Average) [crude]	**-0.02**	No Data Given	-0.23	0.20	0.107
			CTT	Pain	BPI	*β*	CTT to Pain (Average) [adjusted]	**0.00**	No Data Given	-0.26	0.25	0.133
			PTT	Pain	BPI	*β*	PTT to Pain (Max) [crude]	**0.00**	No Data Given	-0.30	0.30	0.153
			PTT	Pain	BPI	*β*	PTT to Pain (Max) [adjusted]	**0.00**	No Data Given	-0.33	0.33	0.168
			CPM	Pain	BPI	*β*	CPM to Pain (Max) [crude]	**-0.42**	No Data Given	-0.92	0.07	0.255
			CPM	Pain	BPI	*β*	CPM to Pain (Max) [adjusted]	**-0.34**	No Data Given	-0.87	0.19	0.270
			CTT	Pain	BPI	*β*	CTT to Pain (Max) [crude]	**-0.03**	No Data Given	-0.29	0.23	0.133
			CTT	Pain	BPI	*β*	CTT to Pain (Max) [adjusted]	**0.04**	No Data Given	-0.27	0.34	0.158
56	Neck Pain	Neck	PTT	Pain	BPI	*β*	PTT to Pain (Average) [crude]	**-0.06**	No Data Given	-0.43	0.32	0.189
			PTT	Pain	BPI	*β*	PTT to Pain (Average) [adjusted]	**-0.10**	No Data Given	-0.51	0.30	0.209
			CPM	Pain	BPI	*β*	CPM to Pain (Average) [crude]	**0.49**	No Data Given	-0.42	1.41	0.464

					CPM	Pain	BPI	*β*	CPM to Pain (Average) [adjusted]	**0.62**	No Data Given	-0.38	1.61	0.510
					CTT	Pain	BPI	*β*	CTT to Pain (Average) [crude]	**0.14**	No Data Given	-0.19	0.47	0.168
					CTT	Pain	BPI	*β*	CTT to Pain (Average) [adjusted]	**0.08**	No Data Given	-0.28	0.45	0.184
					PTT	Pain	BPI	*β*	PTT to Pain (Max) [crude]	**0.01**	No Data Given	-0.40	0.42	0.209
					PTT	Pain	BPI	*β*	PTT to Pain (Max) [adjusted]	**-0.02**	No Data Given	-0.48	0.44	0.235
					CPM	Pain	BPI	*β*	CPM to Pain (Max) [crude]	**0.28**	No Data Given	-0.74	1.30	0.520
					CPM	Pain	BPI	*β*	CPM to Pain (Max) [adjusted]	**0.44**	No Data Given	-0.72	1.60	0.592
					CTT	Pain	BPI	*β*	CTT to Pain (Max) [crude]	**0.20**	No Data Given	-0.16	0.56	0.184
					CTT	Pain	BPI	*β*	CTT to Pain (Max) [adjusted]	**0.17**	No Data Given	-0.23	0.58	0.204
Noiseaux et al., 2014	Prospective Cohort	215	OA	Knee	MPS (VFPI)	Pain	NRS (0-20)	Multivariate Logistic Regression	MPS to Pain	No Data Given	No Data Given			
					HPT	Pain	NRS (0-20)	*β*	HPT to Pain	No Data Given	No Data Given			
					PPT	Pain	NRS (0-20)	*β*	PPT to Pain	No Data Given	No Data Given			
Pedler et al., 2016	Prospective Cohort	91	WAD	Neck	CPT	Pain	VAS	Correlation analysis	CPT to Pain	**0.196**	<0.05			
					PPT	Pain	VAS	*r*	PPT to Pain	**-0.294**	<0.01			

					CPT	Pain	VAS	Multivariate Linear Regression	CPT to Pain	**0.02**	0.389	-0.02	0.05	
					PPT	Pain	VAS	*β*	PPT to Pain	**-0.003**	0.084	-0.007	0.004	
Petersen et al., 2015	Prospective Cohort	78	OA	Knee	TS	Pain	VAS	Correlation analysis and	TS to Pain	**0.240**	0.037			
					PPT	Pain	VAS	*r*	PPT(affected) to Pain	**-0.051**	0.657			
					PPT	Pain	VAS	*r*	PPT(unaffected) to Pain	**-0.077**	0.502			
					CPM	Pain	VAS	*r*	CPM to Pain	**-0.176**	0.123			
					TS	Pain	VAS	Multivariate Linear Regression	TS to Pain (crude)	**0.311**	0.080			0.147
					TS	Pain	VAS	*β*	TS to Pain (adjusted)	**0.289**	0.052			0.146
Petersen et al., 2016	Prospective Cohort	103	OA	Knee	PDT	Pain	VAS	Correlation analysis	PDT to Pain (Relief/Change)	**-0.216**	0.021			
					TS	Pain	VAS	*r*	TS to Pain	No Data Given	No Data Given			
					CPM	Pain	VAS	*r*	CPM to Pain	No Data Given	No Data Given			
					PPT	Pain	VAS	*r*	PPT to Pain	No Data Given	No Data Given			
					PDT	Pain	VAS	Multivariate Linear Regression	PDT to Pain (crude)	**-0.196**	0.034			
					PDT	Pain	VAS	*β*	PDT to Pain (adjusted)	**-0.222**	0.042			
Rakel et al., 2012	Prospective Cohort	215	OA	Knee	MPS (VFPI)	Pain	NRS (0-20)	Correlation analysis	MPS to Pain (Rest)	No Data Given	0.0019			
					HPT	Pain	NRS (0-20)	*r*	HPT to Pain (Postop Rest)	No Data Given	>0.10			
					PPT	Pain	NRS (0-20)	*r*	PPT to Pain (Postop Rest)	No Data Given	>0.10			

MPS (VFPI)	Pain	NRS (0-20)	Multivariate Logistic Regression	MPS to MMPain(vs Mild Pain)	No Data Given	0.010	**1.32**	1.07	1.63
MPS (VFPI)	Pain	NRS (0-20)	*OR*	MPS to SMPain(vs Mild Pain)	No Data Given	0.001	**1.42**	1.14	1.77
HPT	Pain	NRS (0-20)	*OR*	HPT to MMPain(vs Mild Pain)	No Data Given	0.076	**0.9**	0.80	1.01
HPT	Pain	NRS (0-20)	*OR*	HPT to SMPain(vs Mild Pain)	No Data Given	0.007	**0.83**	0.73	0.95
PPT	Pain	NRS (0-20)	*OR*	PPT to MMPain(vs Mild Pain)	No Data Given	0.100	**0.85**	0.70	1.03
PPT	Pain	NRS (0-20)	*OR*	PPT to SMPain(vs Mild Pain)	No Data Given	0.003	**0.66**	0.50	0.87
MPS (VFPI)	Pain	NRS (0-20)	*OR*	MPS to MiRPain(vs None)	No Data Given	0.390	**1.07**	0.92	1.24
MPS (VFPI)	Pain	NRS (0-20)	*OR*	MPS to M/SRPain(vs None)	No Data Given	0.023	**1.16**	1.02	1.32
HPT	Pain	NRS (0-20)	*OR*	HPT to MiRPain(vs None)	No Data Given	0.656	**1.03**	0.91	1.16
HPT	Pain	NRS (0-20)	*OR*	HPT to M/SRPain(vs None)	No Data Given	0.145	**0.92**	0.83	1.03
PPT	Pain	NRS (0-20)	*OR*	PPT to MiRPain(vs None)	No Data Given	0.853	**1.02**	0.82	1.27
PPT	Pain	NRS (0-20)	*OR*	PPT to M/SRPain(vs None)	No Data Given	0.117	**0.84**	0.67	1.05
MPS (VFPI)	Pain	NRS (0-20)	Logistic Regression (Incl. Pain)	MPS to MMPain(vs Mild Pain)	No Data Given	0.022	**1.30**	1.04	1.63
MPS (VFPI)	Pain	NRS (0-20)	*OR*	MPS to SMPain(vs Mild Pain)	No Data Given	0.050	**1.27**	1.00	1.61

					HPT	Pain	NRS (0-20)	*OR*	HPT to MMPain(vs Mild Pain)	No Data Given	0.066	**0.88**	0.77	1.01	
					HPT	Pain	NRS (0-20)	*OR*	HPT to SMPain(vs Mild Pain)	No Data Given	0.031	**0.85**	0.73	0.98	
					MPS (VFPI)	Pain	NRS (0-20)	Logistic Regression (Incl. Depress)	MPS to MMPain(vs Mild Pain)	No Data Given	0.020	**1.31**	1.04	1.64	
					MPS (VFPI)	Pain	NRS (0-20)	*OR*	MPS to SMPain(vs Mild Pain)	No Data Given	0.007	**1.38**	1.09	1.76	
					HPT	Pain	NRS (0-20)	*OR*	HPT to MMPain(vs Mild Pain)	No Data Given	0.055	**0.88**	0.78	1.00	
					HPT	Pain	NRS (0-20)	*OR*	HPT to SMPain(vs Mild Pain)	No Data Given	0.014	**0.83**	0.72	0.96	
Thomazeau et al., 2016	Prospective Cohort	103	OA	Knee	EST	Pain	BPI	Multivariate Linear Regression	EST to Pain (Rest)	**0.0327**	0.113				0.0205
Vaegter et al., 2016	Prospective Cohort	14	OA	Knee	CPM	Pain	NRS	Correlation analysis	CPM (MPPT) to Peak Pain (Change)	**0.08**	0.78				
					CPM	Pain	NRS	*r*	CPM (cPPT) to Peak pain (Change)	**-0.06**	0.83				
					CPM	Pain	NRS	*r*	CPM (cPTT) to Peak pain (Change)	**0.18**	0.53				
					CPM	Pain	NRS	*r*	CPM (MPPT) to Mov Pain (Change)	**0.07**	0.81				
					CPM	Pain	NRS	*r*	CPM (cPPT) to Mov Pain (Change)	**0.04**	0.90				
					CPM	Pain	NRS	*r*	CPM (cPTT) to Mov Pain (Change)	**0.42**	0.14				

					CPM	Pain	NRS	*r*	CPM (MPPT) to Rest Pain (Change)	**-0.37**	0.19			
					CPM	Pain	NRS	*r*	CPM (cPPT) to Rest Pain (Change)	**-0.02**	0.96			
					CPM	Pain	NRS	*r*	CPM (cPTT) to Rest Pain (Change)	**0.57**	0.035			
Valencia et al., 2014	Prospective Cohort	73	Shoulder Pain	Shoulder	SHPR	Pain	BPI	Correlation analysis	SHPR to Pain	**-0.080**	>0.05			
					HPT	Pain	BPI	*r*	HPT to Pain	**0.099**	>0.05			
					CPM	Pain	BPI	*r*	CPM to Pain	**-0.137**	>0.05			
					CPM (%change)	Pain	BPI	*r*	CPM (%change) to Pain	**-0.220**	>0.05			
					SHPR (5th Stim.)	Pain	BPI	Regression analysis	HPT (Change of 5th) to Pain	**-0.350**	0.003			0.01
Werner et al., 2004	Prospective Cohort	20	ACL Tear	Knee	MPT	Pain	VAS	Correlation analysis	Dec MPT to Pain(0-2days)	**0.40**	No Data Given	-0.05	0.72	
					HPT	Pain	VAS	*r*	Dec HPT to Pain(0-2days)	**-0.13**	No Data Given	-0.33	0.54	
					MPP	Pain	VAS	*r*	Inc MPP to Pain(0-2days)	**0.05**	No Data Given	-048	0.41	
					HPT	Pain	VAS	*r*	Inc HPP to Pain(0-2days)	**0.33**	No Data Given	-0.13	0.67	
					MPT	Pain	VAS	*r*	Dec MPT to Pain(3-10days)	**0.25**	No Data Given	-0.21	0.62	
					HPT	Pain	VAS	*r*	Dec HPT to Pain(3-10days)	**-0.21**	No Data Given	-0.60	0.25	
					MPP	Pain	VAS	*r*	Inc MPP to Pain(3-10days)	**0.03**	No Data Given	-0.42	0.46	

					HPP	Pain	VAS	*r*	Inc HPP to Pain(3-10days)	**0.38**	No Data Given		-0.18	0.71	
Wilder-Smith et al., 2010	Prospective Cohort	20	Postoperative pain	Abdomen	DNIC	Pain	NRS	Correlation analysis	DNIC (electric) to Pain (rest) [6 months]	**-0.68**	0.02				
					DNIC	Pain	NRS	*r*	DNIC (electric) to Pain (movement) [6 months]	**-0.63**	0.04				
Wylde et al., 2013	Prospective Cohort	51	OA	Knee	PPT	Pain	WOMAC	Correlation analysis	PPT (Knee) to Pain	**0.257**	0.078				
					PPT	Pain	WOMAC	*r*	PPT (Forearm) to Pain	**0.370**	0.008				
					HPT	Pain	WOMAC	*r*	HPT (Knee) to Pain	**0.130**	0.368				
					HPT	Pain	WOMAC	*r*	HPT (Forearm) to Pain	**0.237**	0.094				
Wylde et al., 2015	Prospective Cohort (D.A.)	254	OA	Hip	PPT	Pain	WOMAC	Multivariate Linear Regression	PPT to Pain (min adj)	**-0.110**	0.010		-0.193	-0.027	0.04
					PPT	Pain	WOMAC	*β*	PPT to Pain (mod adj)	**-0.104**	0.015		-0.187	-0.020	0.04
					PPT	Pain	WOMAC	*β*	PPT to Pain (prop pain adj)	**-0.091**	0.036		-0.176	-0.006	0.04
		239	OA	Knee	PPT	Pain	WOMAC	*β*	PPT to Pain (min adj)	**-0.063**	0.259		-0.174	0.047	0.06
					PPT	Pain	WOMAC	*β*	PPT to Pain (mod adj)	**-0.093**	0.097		-0.204	0.017	0.06
					PPT	Pain	WOMAC	*β*	PPT to Pain (prop pain adj)	**-0.053**	0.313		-0.157	0.051	0.05
Yarnitsky et al., 2008	Prospective Cohort	62	Postoperative pain	Thorax	DNIC	Pain	NRS	Correlation analysis	DNIC to Pain (acute)	**0.1469**	0.2546				
					DNIC	Pain	NRS	*r*	DNIC to Pain (chronic)	**0.3684**	0.0032				
					DNIC	Pain	NRS	Logistic Regression	DNIC to Pain (chronic) (incl. acute)	No Data Given	0.0024	**0.52**	0.33	0.77	
					DNIC	Pain	NRS	*OR*	DNIC to Pain (chronic)	No Data Given	0.0016	**0.55**	0.36	0.77	

Note: *A negative correlation or β-coefficient value indicates that a low QST value is associated with a higher level of pain.*

AAS: Activity Assessment Scale, **AUC:** Area Under Curve, **β:** Beta Coefficient of Regression, **BPI:** Brief Pain Inventory, **BPPT:** Brachial Plexus Provocation Tess, **CDT:** Cold Detection Threshold, **CES-D:** Center for Epidemiological Studies-Depression Scale, **CI:** Confidence Interval, **CPA:** Cuff Pressure Algometry, **CPM:** Conditioned Pain Modulation, **CPP:** Cold Pressor Pain, **cPPT:** Cuff Pain Pressure Threshold, **CPT:** Cold Pain Threshold, **CPTol:** Cold Pain Tolerance Threshold, **cPTT:** Cuff Pain Tolerance Threshold, **CTT:** Cold Tolerance Time, **DNIC:** Diffuse Noxious Inhibitory Control, **DPI:** Daily Pain Intensity, **EPT:** Electrical Pain Threshold, **EPThr:** Electrical Pain Detection Threshold, **EPTol:** Electrical Pain Tolerance Threshold, **ePTT:** Electrical Pain Tolerance Threshold, **EST:** Electrical Sensation Threshold, **HNCS:** Heterotopic Noxious Counter-Stimulation, **HPP:** Heat Pain Perception, **HPT:** Heat Pain Threshold, **HPT (Thr.):** Heat Pain Detection Threshold, **HPT (Tol.):** Heat Pain Tolerance Threshold, **KOOS:** Knee Injury and Osteoarthritis Outcomes Score, **kPa:** Kilopascal, **KSS:** Knee Society Score, **LBP:** Low Back Pain, **MEP:** Motor-Evoked Potentials, **MPS:** Mechanical Pain Sensitivity, **MPP:** Mechanical Pain Perception, **MPT:** Mechanical Pain Threshold, **NDI:** Neck Disability Index, **NRS:** Numerical Rating Scale, **OA:** Osteoarthritis, **ODI:** Oswestry Disability Index, **OR:** Odds Ratio, **OSS:** Oxford Shoulder Score, **PDS:** Post-traumatic Stress Diagnostic Scale, **PinPS:** Pinprick Pain Sensitivity, **PPI:** Pain Pressure Intensity, **PPT:** Pain Pressure Threshold, **PTT:** Pain Tolerance Threshold, **pPTT:** Pressure Pain Tolerance Threshold, **PRTEE:** Patient Rated Tennis Elbow Evaluation, **r:** Pearson's r Correlation-Coefficient, **ρ:** Spearman's ρ Rank-Order Correlation, **RCT:** Randomised Controlled Trial, **RMDQ:** Roland-Morris Disability Questionnaire, **SHPR:** Suprathreshold Heat Pain Response, **SE:** Standard Error, **SS:** Spreading Sensitisation, **TA:** Tibialis Anterior, **Tac.Allo:** Tactile Allodynia, **THPR:** Tonic Heat Pain Response, **TS:** Temporal Summation, **VAS:** Visual Analogue Scale, **VHPI:** Von Frey Pain Intensities, **VNS:** Visual Numeric Scale, **WAD:** Whiplash Associated Disorders, **WDT:** Warmth Detection Threshold, **WOMAC:** Western Ontario and McMaster Universities Osteoarthritis Index, **χ^2:** Determinant of significant difference

Appendix 3. Study details of associations between baseline QST and pain-related disability outcomes

Author	Study Design	Sample size	Diagnosis	Site of Pathology	Baseline QST Predictor	Outcome Variable	Outcome Measures	Statistical Analysis Method	Association Modalities	*(r) / (β)* values	*p*-values	OR	95% CI Lower Limit	95% CI Higher Limit	SE
Aasvang et al., 2010	Prospective Cohort	442	Postoperative pain	Groin	WDT	Disability	AAS	Multivariate Logistic Regression	WDT (Groin) to Disability	No Data Given	0.21	**0.8**	0.56	1.14	
					WDT	Disability	AAS	*OR*	WDT (Arm) to Disability	No Data Given	0.43	**1.12**	0.85	1.48	
					HDT	Disability	AAS	*OR*	HDT (Groin) to Disability	No Data Given	0.65	**0.94**	0.73	1.22	
					HDT	Disability	AAS	*OR*	HDT (Arm) to Disability	No Data Given	0.93	**0.99**	0.81	1.21	
					THPR	Disability	AAS	*OR*	THPR (Groin) [47°C] to Disability	No Data Given	<0.01	**1.05**	1.02	1.08	
					THPR	Disability	AAS	*OR*	THPR (Arm) [47°C] to Disability	No Data Given	0.07	**0.97**	0.95	1.00	
					THPR	Disability	AAS	*OR*	THPR [47°C] to Disability	No Data Given	0.0018	**1.28**	1.10	1.50	
					WDT	Disability	AAS	*OR*	WDT (Groin) [Change] to Disability	No Data Given	0.029	**1.07**	1.01	1.14	
Coombes et al., 2015	Prospective Cohort (D.A.)	41	Epicondylitis	Elbow	CPT	Disability	PRTEE	Multivariate Linear Regression	CPT to Pain (2months) [Unadjusted]	**0.77**	0.008		0.210	1.330	
					CPT	Disability	PRTEE	*β*	CPT to Pain (2months) [Adjusted]	**0.364**	0.008		0.095	0.633	0.1373
					CPT	Disability	PRTEE	*β*	CPT to Pain (1 year)	**0.61**	0.034		0.050	1.170	

					CPT	Disability	PRTEE	*β*	CPT to Pain (1 year)	**0.335**	0.034		0.025	0.645	0.1580
Coronado et al., 2015b	Prospective Cohort	68	LBP	Low Back	PPT	Disability	ODI	Multivariate Logistic Regression	PPT (6weeks) to Disability (6months)	No Data Given	>0.05	1.32	0.83	2.08	
Dubois et al., 2016	Prospective Cohort	77	LBP	Low Back	HNCS	Disability	RMDQ	Correlation analysis	HNCS to Disability	**-0.24**	0.038				
					HNCS	Disability	RMDQ	Multivariate Logistic Regression	HNCS to Disability	**-0.307**	0.005				0.1094
Jull et al., 2013	RCT	97	WAD	Neck	CPT	Disability	NDI	Multivariate Linear Regression	CPT (Usual Care) to Disability	**0.85**	0.004				0.26
					CPT	Disability	NDI	*β*	CPT (Intervention) to Disability	**-0.27**	No Data Given				0.25
					PPT	Disability	NDI	*β*	PPT to Disability	**-0.0007**	No Data Given				0.008
Pedler et al., 2016	Prospective Cohort	91	WAD	Neck	CPT	Disability	NDI	Correlation analysis	CPT to Disability	**0.181**	0.02				
					PPT	Disability	NDI	*r*	PPT to Disability	**-0.384**	<0.01				
					CPT	Disability	NDI	Multivariate Linear Regression	CPT to Disability	**-0.15**	0.170		-0.37	0.07	
					PPT	Disability	NDI	*β*	PPT to Disability	**-0.03**	0.011		-0.05	-0.01	
Sterling et al., 2005	Prospective Cohort	76	WAD	Neck	CPT	Disability	NDI	Multivariate Linear Regression	CPT to Disability (6months)	**0.505**	0.01				0.199
					CPT	Disability	NDI	Logistic Regression	CPT to Disability (m/s 6months)	**0.26**	0.01	1.29	1.05	1.58	0.10
Sterling et al., 2011	Prospective Cohort	155	WAD	Neck	CPT	Disability	NDI	Multivariate Linear Regression	MG: CPT (13≥°C) to Disability (12months)	**1.289**	0.0111	3.628	1.346	9.779	
					CPT	Disability	NDI	*β*	SG: CPT (13≥°C) to	**3.270**	0.0001	26.320	4.981	139.09	

									Disability (12months)					
Sterling et al., 2012	Prospective Cohort	225	WAD	Neck	CPT	Disability	NDI	Multivariate Linear Regression	CPT to Disability (12months)	**0.328**	0.014	0.42	0.58	0.133
					CPT	Disability	NDI	*β*	CPT to Disability (12months) + site	**0.302**	0.017	0.48	0.64	0.126
Valencia et al., 2014	Prospective Cohort	77	WAD	Neck	SHPR	Disability	DASH	Correlation analysis	SHPR to Disability	**-0.043**	>0.05			
					HPT	Disability	DASH	*r*	HPT to Disability	**0.047**	>0.05			
					CPM	Disability	DASH	*r*	CPM to Disability	**-0.211**	>0.05			
					CPM (%change)	Disability	DASH	*r*	CPM (%change) to Disability	**-0.249**	<0.05			
					HPT (5th Stim.)	Disability	DASH	Regression analysis	SHPR (Change of 5th) to Disability	**-0.30**	0.01			0.06
Walton et al., 2011	Prospective Cohort	45	WAD	Neck	PPT	Disability	NDI	Correlation analysis	PPT (Trap) to Disability	**-0.29**	0.06			
					PPT	Disability	NDI	*r*	PPT (Tib) to Disability	**-0.32**	0.03			
					PPT	Disability	NDI	Multivariate Linear Regression	PPT (Tib) to Disability	**-0.30**	No Data Given	-0.60	-0.06	0.138

Note: *A negative correlation or β-coefficient value indicates that a low QST value is associated with a higher level of disability.*

AAS: Activity Assessment Scale, **β:** Beta Coefficient of Regression, **CI:** Confidence Interval, **CPM:** Conditioned Pain Modulation, **CPT:** Cold Pain Threshold, **DASH:** Disability of the Arm, Shoulder and Hand Questionnaire, **HNCS:** Heterotopic Noxious Counter-Stimulation, **HPT:** Heat Pain Threshold, **kPa:** Kilopascal, **LBP:** Low Back Pain, **MPT:** Mechanical Pain Threshold, **NDI:** Neck Disability Index, **NRS:** Numerical Rating Scale, **ODI:** Oswestry Disability Index, **OR:** Odds Ratio, **PPT:** Pain Pressure Threshold, **PRTEE:** Patient Rated Tennis Elbow Evaluation, **r:** Pearson's r Correlation-Coefficient, **RCT:** Randomised Controlled Trial, **RMDQ:** Roland-Morris Disability Questionnaire, **SHPR:** Suprathreshold Heat Pain Response, **SE:** Standard Error, **THPR:** Tonic Heat Pain Response, **WAD:** Whiplash Associated Disorders, **WDT:** Warmth Detection Threshold

Appendix 4. Study details for associations between baseline QST and negative affect outcomes

Author	Study Design	Sample size	Diagnosis	Site of Pathology	Baseline QST Predictor	Outcome Variable	Outcome Measures	Statistical Analysis Method	Association Modalities	*(r)* / *(β)* values	*p*-values	OR	95% CI Lower Limit	95% CI Higher Limit	SE
Sterling et al., 2011	Prospective Cohort	155	WAD	Neck	CPT	PTSD	PDS	Multivariate Linear Regression	RG: CPT (13≥°C) to PTSD (12months)	**1.265**	0.0205	3.543	1.219	10.295	0.544
					PPT	PTSD	PDS	*β*	RG: PPT(Neck) to PTSD (12months)	**-0.003**	0.2361	0.997	0.992	1.002	0.003
					CPT	PTSD	PDS	*β*	M/SG: CPT (13≥°C) to PTSD (12months)	**2.272**	0.0027	9.699	2.217	42.435	0.753
					PPT	PTSD	PDS	*β*	M/SG: PPT(Neck) to PTSD (12months)	**-0.010**	0.0453	0.990	0.980	1.000	0.005
Goodin et al., 2014	Prospective Cohort	225	OA	Knee	TS	Depression	CES-D	Correlation analysis	Mech TS (Knee) to Depression	**0.11**	>0.05				
					TS	Depression	CES-D	*r*	Mech TS (Hand) to Depression	**0.17**	<0.05				
					TS	Depression	CES-D	*r*	Heat TS (Knee) [44°C] to Depression	**0.09**	>0.05				
					TS	Depression	CES-D	*r*	Heat TS (Knee) [46°C] to Depression	**0.10**	>0.05				
					TS	Depression	CES-D	*r*	Heat TS (Knee)	**0.07**	>0.05				

					TS	Depression	CES-D	*r*	[48°C] to Depression	**0.02**	>0.05
					TS	Depression	CES-D	*r*	Heat TS (Forearm) [44°C] to Depression	**0.08**	>0.05
					TS	Depression	CES-D	*r*	Heat TS (Forearm) [46°C] to Depression	**0.09**	>0.05
					TS	Depression	CES-D	*r*	Heat TS (Forearm) [48°C] to Depression		
Valencia et al., 2014	Prospective Cohort	91	WAD	Neck	SHPR	Depression	PHQ-9	Correlation analysis	SHPR to Depression	**-0.147**	>0.05
					HPT	Depression	PHQ-9	*r*	HPT to Depression	**-0.009**	>0.05
					CPM	Depression	PHQ-9	*r*	CPM to Depression	**-0.105**	>0.05
					CPM	Depression	PHQ-9	*r*	CPM (%change) to Depression	**0.0001**	>0.05

Note: *A negative correlation or β-coefficient value indicates that a low QST value is associated with a higher level of depression.*
β: Beta Coefficient of Regression, **CES-D:** Center of Epidemiological Studies-Depression Scale, **CI:** Confidence Interval, **CPM:** Conditioned Pain Modulation, **CPT:** Cold Pain Threshold, **HPT:** Heat Pain Threshold, **kPa:** Kilopascal, **OA:** Osteoarthritis, **OR:** Odds Ratio, **PDS:** Post-traumatic Stress Diagnostic Scale, **PHQ-9:** Patient Health Questionnaire, **PPT:** Pain Pressure Threshold, **PTSD:** Post-traumatic Stress Disorder, **r:** Pearson's r Correlation-Coefficient, **SHPR:** Suprathreshold Heat Pain Response, **SE:** Standard Error, **TS:** Temporal Summation, **WAD:** Whiplash Associated Disorders

Appendix 5. Shapiro-Wilk, skewness and kurtosis data of all baseline variables found to deviate significantly from normal distribution before as well as after logarithmic transformation.

	Sample Size n=25	Normality before transformation				Normality after transformation			
Healthy Group Rater 1		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	0.81	<0.01	1.6	5.2	**0.96**	**0.51**	**0.3**	**3.2**
	TS^{WUD}	0.82	<0.01	1.7	6.5	**0.95**	**0.28**	**-0.1**	**1.9**
	TS^{WUR}	0.61	<0.01	3.2	14.1	**0.94**	**0.19**	**0.9**	**4.4**
	CPM^{Unc}	**0.99**	**0.99**	**-0.1**	**3.1**	-	-	-	-
	$CPM^{PPT-mean}$	**0.93**	**0.08**	**-0.4**	**5.2**	-	-	-	-
	Sample Size n=25	Normality before transformation				Normality after transformation			
Healthy Group Rater 2		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	0.77	<0.01	1.8	5.7	**0.97**	**0.60**	**0.33**	**3.0**
	TS^{WUD}	0.85	<0.01	1.2	3.8	**0.95**	**0.30**	**-0.4**	**2.4**
	TS^{WUR}	0.87	<0.01	1.2	3.8	**0.98**	**0.86**	**0.1**	**2.2**
	CPM^{Unc}	**0.94**	**0.17**	**0.7**	**2.9**	-	-	-	-
	$CPM^{PPT-mean}$	**0.97**	**0.72**	**1.4**	**5.0**	-	-	-	-
	Sample Size n=25	Normality before transformation				Normality after transformation			
Patient Group		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	**0.98**	**0.77**	**0.3**	**2.9**	-	-	-	-
	TS^{WUD}	0.91	0.04	0.8	3.2	**0.91**	**0.05**	**-0.6**	**2.2**
	TS^{WUR}	0.67	<0.01	2.4	9.0	**0.93**	**0.09**	**0.6**	**2.3**
	CPM^{Unc}	**0.98**	**0.78**	**-0.3**	**2.9**	-	-	-	-
	$CPM^{PPT-mean}$	**0.97**	**0.54**	**-0.2**	**4.1**	-	-	-	-
	Sample Size n=50	Normality before transformation				Normality after transformation			
Combined Group		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	0.90	<0.01	1.3	5.0	**0.98**	**0.38**	**-0.4**	**3.8**
	TS^{WUD}	0.87	<0.01	1.4	5.3	0.95	0.04	-0.4	2.2
	TS^{WUR}	0.57	<0.01	3.4	16.3	0.92	<0.01	1.0	3.5
	CPM^{Unc}	**0.99**	**0.91**	**-0.1**	**3.1**	-	-	-	-
	$CPM^{PPT-mean}$	**0.95**	**0.05**	**-0.2**	**4.9**	-	-	-	-

$CPM^{PPT-mean}$: 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, **PPT:** Pain Pressure Detection Threshold, **TS^{WUD}:** Temporal Summation calculated as a difference, **TS^{WUR}:** Temporal Summation calculated as a ratio
Data in **bold** indicate normal distribution of data either before or after transformation. A hyphen indicates that no attempt of transformation was undertaken

Appendix 6. Shapiro-Wilk, skewness and kurtosis data of all follow-up variables found to deviate significantly from normal distribution before as well as after logarithmic transformation.

	Sample Size n=25	Normality before transformation				Normality after transformation			
Healthy Group Rater 1		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	0.67	<0.01	2.7	10.3	**0.92**	**0.05**	**0.9**	**4.6**
	TS^{WUD}	0.82	<0.01	1.5	2.0	**0.95**	**0.20**	**-0.2**	**2.0**
	TS^{WUR}	0.61	<0.01	2.9	11.7	**0.92**	**0.05**	**-0.6**	**5.4**
	CPM^{Unc}	0.80	<0.01	1.8	6.6	0.85	0.01	1.3	4.4
	$CPM^{PPT\text{-}mean}$	0.80	<0.01	2.0	7.9	**0.97**	**0.69**	**1.4**	**5.0**
	Sample Size n=25	Normality before transformation				Normality after transformation			
Patient Group		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	**0.95**	**0.27**	**0.7**	**3.1**	-	-	-	-
	TS^{WUD}	0.88	0.01	1.0	3.0	0.87	0.01	1.0	3.0
	TS^{WUR}	0.61	<0.01	2.1	6.0	0.83	<0.01	1.1	3.1
	CPM^{Unc}	0.91	0.03	1.0	3.5	**0.93**	**0.19**	**0.1**	**1.0**
	$CPM^{PPT\text{-}mean}$	0.88	0.01	1.2	4.4	**0.95**	**0.27**	**0.2**	**2.0**
	Sample Size n=50	Normality before transformation				Normality after transformation			
Combined Group		**SW**	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
	PPT	0.77	<0.01	2.5	11.4	**0.96**	**0.09**	**-0.1**	**4.8**
	TS^{WUD}	0.86	<0.01	1.2	4.0	0.94	0.02	-0.3	2.0
	TS^{WUR}	0.62	<0.01	2.4	8.1	0.91	<0.01	-0.2	5.4
	CPM^{Unc}	0.84	<0.01	1.9	8.0	0.94	0.03	0.9	3.8
	$CPM^{PPT\text{-}mean}$	0.85	<0.01	1.8	7.4	0.94	0.04	1.0	3.5

$CPM^{PPT\text{-}mean}$: 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, **PPT:** Pain Pressure Detection Threshold, **TS^{WUD}:** Temporal Summation calculated as a difference, **TS^{WUR}:** Temporal Summation calculated as a ratio
Data in **bold** indicate normal distribution before or after transformation. A hyphen (-) indicates that no attempt of transformation was undertaken

Appendix 7. Shapiro-Wilk, skewness and kurtosis data of all factors found to deviate significantly from normal distribution before as well as after logarithmic transformation.

Sample Size n=97		Normality before transformation				Normality after transformation			
		SW	SW p-value	Skewness	Kurtosis	SW	SW p-value	Skewness	Kurtosis
Dem.	AGE	**0.98**	**0.18**	**-0.2**	**2.8**	-	-	-	-
	BMI	0.94	<0.01	0.9	4.0	**1.00**	**1.00**	**-0.1**	**2.8**
CS Indices	PPT	0.95	<0.01	0.8	3.3	**0.99**	**0.90**	**-0.1**	**2.6**
	TS	0.79	<0.01	1.6	5.0	0.97	0.02	0.1	2.0
	CPM	**0.99**	**0.80**	**-0.1**	**3.0**	-	-	-	-
	CMT	**0.99**	**0.66**	**0.1**	**2.4**	-	-	-	-
Self-management Self-care	HEIQ-HDB	**0.97**	**0.06**	**0.1**	**3.0**	-	-	-	-
	HEIQ-PEL	**0.98**	**0.13**	**-0.1**	**2.9**	-	-	-	-
	HEIQ-SMI	0.93	<0.01	0.3	3.7	0.94	<0.01	0.1	3.8
	HEIQ-CAA	**0.97**	**0.09**	**-0.1**	**3.0**	-	-	-	-
	HEIQ-STA	0.94	<0.01	0.2	2.9	0.95	<0.01	0.1	2.8
	HEIQ-SIS	0.96	0.01	-0.5	3.3	**0.98**	**0.13**	**-0.1**	**2.9**
	HEIQ-HSN	0.95	<0.01	-0.5	3.9	0.96	<0.01	-0.1	3.4
	HEIQ-ED	**0.98**	**0.06**	**-0.1**	**2.6**	-	-	-	-
	PSEQ	0.95	<0.01	0.4	2.1	0.97	0.04	-0.1	2.1
	HCUQ	0.82	<0.01	1.8	6.3	0.96	<0.01	0.2	3.3
Pain	NRS	0.96	0.01	-0.4	3.4	0.97	0.03	-0.1	2.9
	PDETECT	**0.99**	**0.47**	**-0.1**	**2.4**	-	-	-	-
	Now	0.97	0.01	-0.3	3.2	0.97	0.02	-0.1	2.9
	Strongest	0.82	<0.01	-1.8	8.1	0.92	<0.01	-0.2	2.7
	Average	0.93	<0.01	-0.5	4.7	0.94	<0.01	0.1	3.8
Negative Affect	HADS Anx.	0.97	0.02	0.3	2.2	**0.98**	**0.29**	**-0.1**	**2.4**
	HADS Dep.	0.96	0.01	0.1	2.1	0.96	0.01	-0.1	2.0
	PCS	0.96	0.01	0.4	2.3	**0.98**	**0.09**	**-0.1**	**2.1**
	TSK	**0.99**	**0.78**	**0.1**	**2.5**	-	-	-	-
Limiting Factors	RMDQ	**0.97**	**0.06**	**-0.1**	**2.2**	-	-	-	-
	FSS	0.96	<0.01	-0.3	2.0	0.96	0.01	-0.1	1.9
	FSVAS	0.97	0.01	0.3	2.5	**0.98**	**0.07**	**-0.1**	**2.7**
	FMSS	0.97	0.03	0.3	2.2	**0.98**	**0.15**	**-0.1**	**2.1**
QoL	EQ-5D-5L	0.96	0.01	-0.4	2.4	0.96	0.01	-0.4	2.4
	EQ-VAS	**0.98**	**0.18**	**0.1**	**2.8**	-	-	-	-
Risk	STarT Back	0.95	<0.01	-0.2	2.0	0.95	<0.01	-0.2	2.0

BMI: Body Mass Index, **CMT:** Central Mechanisms Trait, **CPM:** Conditioned Pain Modulation, **EQ-5D-5L:** Quality of Life Instrument, **EQVAS:** Quality of Life Visual Analogue Scale, **FMSS:** Fibromyalgia Severity Scale, **FSS:** Fatigue Severity Scale, **FSVAS:** Fatigue Severity Visual Analogue Scale, **HADS:** Hospital Anxiety & Depression Scale, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **NRS:** Pain Numerical Rating Scale, **PCS:** Pain Catastrophising Scale, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **RMDQ:** Roland-Morris Disability Questionnaire, **STarT-Back:** Stratification tool, **SW:** Shapiro-Wilk Normality Test, **TS:** Temporal Summation, **TSK:** Tampa Scale of Kinesiophobia

Data in **bold** font indicate normally distributed data (p>0.05) or data that became normally distributed after logarithmic transformation. A hyphen (-) indicates that data were normally distributed before transformation and therefore no transformation occurred. Not all factors, initially found to be significantly different to normal, became normally distributed after transformation (p<0,05), indicating the potential use of non-parametric tests for their incorporation in analyses.

Appendix 8. Explanatory power of models exploring the relationship between measurements of CS-related indices and each SM/SC construct at baseline adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue as well as disability.

	Primary Predictors adjusted for additional factors	Dependent Variable									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2
CS Indices as Primary Predictors	PPT †, ‡	0.14*	0.50***	0.09*	0.41***	0.19***	0.33***	0.11*	0.58***	0.63***	0.13**
	PPT $_{\text{disability instead of pain}}$	0.19***	0.52***	0.09*	0.44***	0.21***	0.33***	0.11*	0.57***	0.68***	0.12***
	TS †, ‡	0.14**	0.52***	0.14**	0.43***	0.20***	0.33***	0.12**	0.58***	0.63***	0.12**
	TS $_{\text{disability instead of pain}}$	0.19***	0.54***	0.14**	0.45***	0.23***	0.33***	0.12**	0.56***	0.68***	0.11*
	CPM †, ‡	0.14**	0.52***	0.08*	0.47***	0.19***	0.33***	0.15**	0.59***	0.63***	0.14**
	CPM $_{\text{disability instead of pain}}$	0.19***	0.54***	0.08*	0.48***	0.22***	0.33***	0.15**	0.57***	0.68***	0.14**
	ACR †, ‡	0.16**	0.51***	0.12*	0.41***	0.19***	0.33***	0.11*	0.58***	0.62***	0.12**
	ACR $_{\text{disability instead of pain}}$	0.20***	0.53***	0.12**	0.44***	0.21***	0.33***	0.11*	0.56***	0.68***	0.12*
	>9/24 †, ‡	0.16**	0.50***	0.10*	0.42***	0.21***	0.33***	0.11*	0.58***	0.63***	0.13**
	>9/24 $_{\text{disability instead of pain}}$	0.21***	0.53***	0.10*	0.45***	0.23***	0.33***	0.12*	0.57***	0.69***	0.12***
	CMT †, ‡, ¥	0.07*	0.33**	0.001	0.21***	0.07*	0.09*	0.07*	0.46***	0.64***	0.09*
	CMT $_{\text{disability instead of pain}}$ ¥	0.12***	0.35***	0.001	0.25***	0.07*	0.09*	0.07*	0.46***	0.64***	0.09*

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation.

R^2 values represent the overall explanatory power of each regression model featuring CS indices as primary predictors and SM/SC follow-up scores as dependent variables. In primary analyses, each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). In separate exploratory analyses, baseline disability scores instead of pain and baseline scores of each SM/SC were included as independent variables separately as well as in the same models. All variables entered in the models were based on 87 observations.

† Primary predictor.

‡ Primary hypothesis models adjusted only for age, sex, pain, depression, catastrophising and fatigue.

¥ Independent variables were age, sex, pain and quality of life.

* ≤0.05, ** <0.01, ***<0.001.

Appendix 9. Shapiro-Wilk, skewness and kurtosis data of all factors found to deviate significantly from normal distribution before as well as after logarithmic transformation.

	Sample Size n=87	Normality before transformation				Normality after transformation			
		SW	**SW p-value**	**Skewness**	**Kurtosis**	**SW**	**SW p-value**	**Skewness**	**Kurtosis**
Self-management Self-care	**HEIQ-HDB**	0.96	0.01	-0.1	2.9	0.96	0.01	-0.1	2.8
	HEIQ-PEL	0.95	0.01	0.3	2.4	0.96	0.02	-0.1	2.4
	HEIQ-SMI	0.95	<0.01	-0.1	3.8	0.95	<0.01	0.1	3.6
	HEIQ-CAA	0.97	0.03	-0.1	2.5	0.97	0.03	-0.1	2.5
	HEIQ-STA	0.90	<0.01	-0.4	5.6	0.91	<0.01	0.1	4.5
	HEIQ-SIS	0.90	<0.01	-0.9	4.1	0.93	<0.01	-0.1	3.5
	HEIQ-HSN	0.95	0.01	0.1	2.4	0.96	0.01	-0.1	2.5
	HEIQ-ED	**0.98**	**0.30**	**-0.1**	**2.6**	-	-	-	-
	PSEQ	**0.98**	**0.22**	**-0.1**	**2.2**	-	-	-	-
	HCUQ	**0.98**	**0.16**	**0.1**	**2.5**	-	-	-	-
Pain	**NRS**	**0.97**	**0.06**	**-0.2**	**2.7**	-	-	-	-
	PDETECT	**0.98**	**0.16**	**-0.1**	**2.3**	-	-	-	-
	Now	0.97	0.03	-0.2	2.4	0.97	0.03	-0.1	2.4
	Strongest	0.84	<0.01	-1.5	5.5	0.94	<0.01	-0.1	2.5
	Average	0.95	<0.01	-0.7	3.7	0.97	0.03	-0.1	2.9
Negative Affect	**HADS Anx.**	0.96	0.01	0.3	2.1	**0.97**	**0.09**	**-0.1**	**2.2**
	HADS Dep.	0.95	<0.01	0.1	1.8	0.96	0.01	-0.2	1.9
	PCS	0.90	<0.01	0.9	2.9	**0.98**	**0.23**	**-0.1**	**2.4**
	TSK	**0.98**	**0.26**	**0.2**	**2.5**	-	-	-	-
Limiting Factors	**RMDQ**	0.95	<0.01	0.3	2.2	0.97	0.02	-0.1	2.2
	FSS	**0.99**	**0.55**	**0.1**	**2.4**	-	-	-	-
	FSVAS	0.96	<0.01	0.3	2.3	0.96	0.01	-0.1	2.4
	FMSS	0.95	<0.01	0.7	3.0	**0.98**	**0.39**	**-0.1**	**2.6**
QoL	**EQ-5D-5L**	0.96	0.01	-0.6	3.2	0.96	0.01	-0.6	3.2
	EQ-VAS	**0.97**	**0.06**	**-0.2**	**2.4**	-	-	-	-
Risk	**STarT Back**	0.95	<0.01	0.3	2.1	0.96	0.01	-0.1	2.3

EQ-5D-5L: Quality of Life Instrument, **EQVAS:** Quality of Life Visual Analogue Scale, **FMSS:** Fibromyalgia Severity Scale, **FSS:** Fatigue Severity Scale, **FSVAS:** Fatigue Severity Visual Analogue Scale, **HADS:** Hospital Anxiety & Depression Scale, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **NRS:** Pain Numerical Rating Scale, **PCS:** Pain Catastrophising Scale, **PSEQ:** Pain Self-efficacy Questionnaire, **RMDQ:** Roland-Morris Disability Questionnaire, **STarT-Back:** Stratification tool, **SW:** Shapiro-Wilk Normality Test, **TSK:** Tampa Scale of Kinesiophobia

Values in **bold** font indicate normally distributed data (p>0.05) or data that became normally distributed after logarithmic transformation. Hyphens indicate that data were found to be normally distributed before transformation and therefore no transformation occurred. Not all factors, initially found to be significantly different to normal, became normally distributed after transformation (p<0,05), indicating the potential use of non-parametric tests for their incorporation in analyses.

Appendix 10. Bivariate correlations between baseline and follow-up SM/SC constructs as well as between each psychological variable aimed to be used as independent variables in multivariable regression models.

Dependent Variables at Baseline	Dependent Variables at follow-up									
	HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
HEIQ-HDB	**0.48*** **	-	-	-	-	-	-	-	-	-
HEIQ-PEL	-	**0.58*** **	-	-	-	-	-	-	-	-
HEIQ-SMI	-	-	**0.34*** **	-	-	-	-	-	-	-
HEIQ-CAA	-	-	-	**0.58*** **	-	-	-	-	-	-
HEIQ-STA	-	-	-	-	**0.39*** **	-	-	-	-	-
HEIQ-SIS	-	-	-	-	-	**0.54*** **	-	-	-	-
HEIQ-HSN	-	-	-	-	-		**0.54*** **	-	-	-
HEIQ-ED	-	-	-	-	-	-	-	**0.57***	-	-
PSEQ	-	-	-	-	-	-	-	-	**0.72*** **	-
HCUQ	-	-	-	-	-	-	-	-	-	0.16

HCUQ: Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PSEQ:** Pain Self-efficacy Questionnaire

Values in **bold** indicate statistical significance.
* ≤0.05, ** <0.01, ***<0.001.
Yellow colour indicates the association of β<0.50
Green colour indicates the association of β>0.50

Appendix 11. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, disability, depression, catastrophisation and fatigue as well as the baseline SM/SC score of each dependent SM/SC variable.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Quantitative Sensory Testing	*PPT* †	0.03	0.001	0.08	0.02	0.12	**0.68***	0.02	-0.07	1.28	0.22
	Baseline SM/SC	**0.41*****	**0.002****	**1.28*****	**0.27***	**1.48*****	**0.71*****	**0.06*****	**0.40****	**10.35*****	0.51
	Age	-0.001	-0.001	-0.006	-0.001	**-0.03****	0.01	0.001	0.007	0.001	0.06
	Sex	0.06	0.001	-0.03	0.08	0.24	0.69	0.02	-0.005	1.27	1.39
	Disability	-0.02	-0.001	0.01	**-0.02***	-0.02	0.002	-0.001	**0.03***	-0.33	**0.25****
	Depression	-0.004	-0.001	-0.02	-0.05	**-0.16***	-0.10	-0.002	**0.08***	-0.68	**0.55****
	Catastrophisation	0.002	-0.001	-0.04	-0.01	0.10	-0.15	-0.007	-0.01	0.22	-0.11
	Fatigue	0.001	0.001	**0.002***	0.001	**0.004***	-0.001	0.001	-0.001	**0.03***	0.002
	TS †	1.32	0.001	0.28	-0.55	0.71	-2.60	0.32	**1.95***	11.31	**-11.48***
	Baseline SM/SC	**0.41*****	**0.002****	**1.33*****	**0.26***	**1.46*****	**0.72*****	**0.06*****	**0.38****	**13.03*****	0.48
	Age	-0.004	-0.001	-0.007	-0.001	**-0.03****	0.004	0.001	0.006	-0.03	**0.06***
	Sex	0.02	-0.001	-0.09	0.07	0.13	0.23	0.001	0.02	0.12	1.36
	Disability	-0.02	**-0.001***	0.01	**-0.02***	-0.02	-0.006	-0.002	**0.03***	-0.38	**0.26*****
	Depression	0.005	-0.001	-0.01	-0.05	**-0.16***	-0.09	0.001	**0.09****	-0.63	**0.48***
	Catastrophisation	-0.007	-0.001	-0.04	-0.01	0.09	-0.19	-0.01	-0.01	0.07	-0.07
	Fatigue	0.001	0.001	**0.002***	0.001	**0.004***	-0.001	0.001	-0.002	0.03	0.003
	CPM †	0.001	0.001	0.001	-0.001	0.002	-0.001	0.001	-0.001	-0.006	**0.01***
	Baseline SM/SC	**0.41*****	**0.002***	**1.30*****	**0.28****	**1.43****	**0.73*****	**0.06*****	**0.37****	**13.33*****	0.39
	Age	-0.003	-0.001	-0.007	-0.001	**-0.03****	0.002	0.001	0.007	-0.02	0.05
	Sex	0.05	-0.001	-0.08	0.06	0.16	0.20	0.003	0.03	0.19	1.32
	Disability	-0.01	-0.001	0.01	**-0.02***	-0.01	-0.01	-0.001	**0.03***	-0.38	**0.29*****
	Depression	-0.003	-0.001	-0.02	-0.05	**-0.16***	-0.07	-0.002	**0.08***	-0.67	**0.55****
	Catastrophisation	-0.005	-0.001	-0.05	-0.02	0.09	-0.21	-0.009	0.002	0.16	-0.17
	Fatigue	0.001	0.001	**0.002***	0.001	**0.004***	-0.001	0.001	-0.001	0.03	0.002

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (baseline scores of each SM/SC outcome, age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Appendix 12. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, disability, depression, catastrophisation and fatigue as well as the baseline SM/SC score of each dependent SM/SC variable.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Body Manikin	*ACR* †	-0.17	-0.001	0.10	0.04	0.30	0.41	0.01	-0.01	4.01	-0.43
	Baseline SM/SC	**0.43*** **	**0.002** **	**1.26*** **	**0.27***	**1.41** **	**0.73*** **	**0.06*** **	**0.39** **	**12.77*** **	0.50
	Age	-0.004	-0.001	-0.007	-0.001	**-0.03** **	0.002	0.001	0.008	-0.02	0.05
	Sex	0.09	0.001	-0.12	0.05	0.06	0.09	-0.001	0.05	-0.80	1.33
	Disability	-0.01	-0.001	0.009	**-0.02***	-0.02	-0.01	-0.001	**0.03***	-0.42	**0.25** **
	Depression	0.001	-0.001	-0.02	-0.05	**-0.17***	-0.08	-0.002	**0.08***	-0.81	**0.56** **
	Catastrophisation	-0.001	-0.001	-0.04	-0.02	0.10	-0.20	-0.008	-0.006	0.13	-0.13
	Fatigue	0.001	0.001	**0.002***	0.001	**0.004***	-0.001	0.001	-0.001	0.03	0.002
	>9/24 †	-0.13	-0.001	-0.12	0.02	0.04	-0.07	-0.06	-0.16	0.47	0.47
	Baseline SM/SC	**0.43*** **	**0.002** **	**1.35*** **	**0.27***	**1.43** **	**0.73*** **	**0.06*** **	**0.38** **	**13.16*** **	0.48
	Age	-0.003	-0.001	-0.006	-0.001	**-0.03** **	0.002	0.001	0.008	-0.02	0.05
	Sex	0.07	0.001	-0.06	0.06	0.13	0.22	0.02	0.08	0.18	1.10
	Disability	-0.01	-0.001	0.01	**-0.02***	-0.02	-0.008	-0.001	**0.04** **	-0.38	**0.24** **
	Depression	0.005	-0.001	-0.008	-0.05	**-0.17***	-0.06	0.002	**0.09** **	-0.71	**0.53** **
	Catastrophisation	-0.003	-0.001	-0.05	-0.02	0.09	-0.21	-0.01	-0.007	0.13	-0.12
	Fatigue	0.001	0.001	**0.002***	0.001	**0.004***	-0.001	0.001	-0.001	0.03	0.001
Trait Score	*CMT* †	-0.01	-0.001	0.12	-0.18	**0.71***	-1.02	0.03	0.06	-0.44	0.80
	Baseline SM/SC	**0.39*** **	**0.003*** **	**1.34*** **	**0.33*** **	**1.87*** **	**0.73*** **	**0.06*** **	**0.37** **	**12.72*** **	0.28
	Age	-0.004	-0.001	-0.003	0.001	-0.02	-0.001	0.001	0.002	0.03	0.02
	Sex	0.04	0.001	-0.07	0.10	0.08	0.43	-0.004	0.02	0.82	0.97
	Disability	-0.008	-0.001	0.02	-0.02	-0.007	-0.01	0.001	0.03	-0.27	**0.21** **
	Quality of Life	0.25	**0.005***	0.62	0.20	**1.68***	0.12	**0.20***	-0.48	2.84	-2.85

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (baseline scores of each SM/SC outcome, age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Appendix 13. Multivariable models exploring the relationship between baseline measurements of distinct QST modalities and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue as well as programme participation.

	Variables used as adjustments	Dependent Variables									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Quantitative Sensory Testing	*PPT* †	0.04	0.001	0.09	0.03	0.10	**0.81***	0.006	-0.08	1.28	0.08
	Programme	0.22	0.001	0.33	0.04	**0.85****	-0.10	-0.03	-0.07	**-6.52***	**3.46*****
	Age	-0.005	-0.001	-0.01	-0.004	**-0.04*****	0.008	0.001	0.004	-0.05	0.03
	Sex	-0.14	-0.001	-0.22	0.03	-0.17	-0.13	-0.05	0.06	1.27	0.38
	Pain	**-0.04***	-0.001	0.008	-0.003	-0.03	-0.13	-0.001	0.02	-0.36	0.16
	Depression	-0.05	**-0.001*****	-0.07	**-0.11*****	**-0.32*****	**-0.50****	-0.01	**0.10****	**-2.14****	0.35
	Catastrophisation	-0.006	-0.001	-0.03	-0.03	0.09	0.11	-0.002	**0.08***	-0.26	-0.07
	Fatigue	-0.001	-0.001	**0.002***	-0.001	0.003	0.004	0.001	-0.001	0.01	0.001
	TS †	1.33	0.001	-0.74	-1.10	-0.16	-4.61	0.04	**2.20***	13.36	**-10.44***
	Programme	0.22	0.001	0.31	0.04	**0.84****	-0.22	-0.03	-0.05	**-6.66***	**3.42*****
	Age	-0.006	-0.001	-0.01	-0.003	**-0.04*****	-0.001	0.001	0.003	-0.08	0.04
	Sex	-0.19	-0.001	-0.27	0.03	-0.24	-0.64	-0.06	0.09	0.13	0.49
	Pain	**-0.04***	-0.001	0.01	-0.001	-0.02	-0.11	-0.001	0.02	-0.35	0.17
	Depression	-0.04	**-0.001****	-0.07	**-0.12*****	**-0.32*****	**-0.51****	-0.01	**0.11****	**-2.04****	0.29
	Catastrophisation	-0.02	-0.001	-0.03	-0.02	0.09	0.06	-0.003	**0.07***	-0.45	-0.01
	Fatigue	-0.001	-0.001	**0.002***	-0.001	0.003	0.005	0.001	-0.001	0.01	0.002
	CPM †	0.001	**0.001***	0.001	0.001	0.002	0.001	0.001	**-0.002***	0.01	0.008
	Programme	0.23	0.001	0.32	0.05	**0.87****	-0.20	-0.03	-0.08	**-6.54***	**3.56*****
	Age	-0.004	-0.001	-0.01	-0.004	**-0.04*****	-0.003	0.001	0.004	-0.06	0.03
	Sex	-0.15	-0.001	-0.28	0.02	-0.21	-0.70	-0.05	0.09	0.51	0.44
	Pain	**-0.04***	-0.001	0.009	-0.003	-0.03	-0.12	-0.001	0.02	-0.35	0.15
	Depression	-0.04	**-0.001****	-0.07	**-0.11*****	**-0.31*****	**-0.48****	-0.01	**0.09****	**-2.10****	**0.37***
	Catastrophisation	-0.01	-0.001	-0.04	-0.03	0.08	0.03	-0.003	**0.09****	-0.40	-0.09
	Fatigue	-0.01	-0.001	**0.002***	-0.001	0.003	0.004	0.001	-0.001	0.01	0.001

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between QST modalities and SM/SC constructs was adjusted for the same baseline variables (programme participation, age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Appendix 14. Multivariable models exploring the relationship between baseline measurements of other CS-related indices and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue as well as programme participation.

	Variables used as adjustments	Dependent Variable									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		β	β	β	β	β	β	β	β	β	β
Body Manikin	*ACR* †	-0.07	-0.001	0.21	0.05	0.42	0.73	-0.001	0.02	4.17	-0.21
	Programme	0.21	0.001	0.34	0.05	**0.88****	-0.14	-0.03	-0.05	**-6.27***	**3.43*****
	Age	-0.005	-0.001	-0.01	-0.004	**-0.04*****	-0.003	0.001	0.005	-0.06	0.03
	Sex	-0.14	-0.001	**-0.35***	-0.008	-0.37	-0.94	-0.06	0.11	-0.96	0.39
	Pain	-0.04	-0.001	0.005	-0.004	-0.03	-0.13	-0.001	0.02	-0.44	0.17
	Depression	-0.04	**-0.001****	-0.07	**-0.11*****	**-0.33*****	**-0.50****	-0.01	**0.10****	**-2.21****	0.36
	Catastrophisation	-0.009	-0.001	-0.04	-0.03	0.08	0.03	-0.003	**0.08***	-0.39	-0.07
	Fatigue	-0.001	-0.001	**0.002***	-0.001	0.003	0.003	0.001	-0.001	0.009	0.001
	>9/24 †	-0.05	-0.001	-0.06	0.01	0.12	0.33	-0.02	-0.12	0.89	0.55
	Programme	0.22	0.001	0.32	0.04	**0.83****	-0.24	-0.03	-0.04	**-6.78***	**3.39*****
	Age	-0.005	-0.001	-0.01	-0.004	**-0.04*****	-0.005	0.001	0.005	-0.07	0.03
	Sex	-0.15	-0.001	-0.27	0.006	-0.26	-0.78	-0.05	0.15	0.16	0.21
	Pain	**-0.04***	-0.001	0.01	-0.003	-0.03	-0.12	-0.001	0.02	-0.35	0.16
	Depression	-0.04	**-0.001***	-0.06	**-0.11*****	**-0.32*****	**-0.50****	-0.01	**0.10****	**-2.17****	0.32
	Catastrophisation	-0.01	-0.001	-0.04	-0.03	0.09	0.04	-0.003	**0.08***	-0.36	-0.07
	Fatigue	-0.01	-0.001	**0.002***	-0.001	0.003	0.004	0.001	-0.001	0.01	-0.001
Trait Score	*CMT* †	-0.17	-0.002	0.06	**-0.33***	0.07	-0.79	0.05	**0.36***	-4.61	-0.004
	Programme	0.25	0.001	0.31	0.03	**0.72***	-0.45	-0.03	-0.03	**-6.29***	**3.61*****
	Age	-0.004	0.001	-0.005	-0.001	-0.02	0.02	0.002	0.001	0.03	0.009
	Sex	-0.13	-0.001	-0.24	0.09	-0.16	-0.40	-0.06	0.02	1.78	0.27
	Pain	-0.03	-0.001	0.03	0.01	0.002	-0.09	0.005	-0.002	**0.05**	**0.13**
	Quality of Life	0.37	**0.006****	**0.86***	**0.69***	**1.84***	1.14	**0.29****	**-0.94***	**15.99***	-2.58

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation

Beta values represent standardized (β) regression coefficients for each variable entered simultaneously in the model and express their association with each SM/SC outcome at 3-months follow-up. Each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (programme participation, age, sex, pain, depression, catastrophisation and fatigue). Bivariate correlations between each independent variable and each SM/SC outcome are presented in **Table 35** and **Table 36**. Multicollinearity testing yielded VIF values ranging from 1.2 to 2.5 for all independent variables indicating not significant multicollinearity between them. All variables entered in the models were based on 87 observations.

† Primary predictor. Values in **bold** indicate statistical significance. * ≤0.05, ** <0.01, ***<0.001. Yellow colour indicates the association between a variable and an SM/SC outcome of β<0.50, Green colour indicates the association between a variable and an SM/SC outcome of β>0.50

Appendix 15. Explanatory power of models exploring the relationship between baseline measurements of CS-related indices and each SM/SC construct at 3-months follow-up adjusted for baseline age, sex, pain, depression, catastrophisation and fatigue as well as baseline scores of each SM/SC and disability.

	Primary Predictors adjusted for additional factors	Dependent Variable									
		HEIQ-HDB	HEIQ-PEL	HEIQ-SMI	HEIQ-CAA	HEIQ-STA	HEIQ-SIS	HEIQ-HSN	HEIQ-ED	PSEQ	HCUQ
		R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2	R^2
CS Indices as Primary Predictors in Distinct Regression Models	PPT †, ‡	0.02	0.30***	0.05	0.29***	0.13*	0.23***	0.02	0.34***	0.34***	0.29***
	PPT baseline dependent variable score	0.23**	0.37***	0.18**	0.37***	0.24***	0.42***	0.32***	0.41***	0.52***	0.30***
	PPT disability instead of pain	0.03	0.34***	0.08	0.35***	0.13*	0.21**	0.03	0.38***	0.42***	0.33***
	PPT baseline dependent variable and disability	0.22***	0.38***	0.19**	0.40***	0.24***	0.41***	0.32***	0.44***	0.53***	0.34***
	TS †, ‡	0.04	0.29***	0.04	0.31***	0.13*	0.18**	0.02	0.36***	0.34***	0.32***
	TS baseline dependent variable score	0.24***	0.36***	0.18*	0.37***	0.23***	0.38***	0.32***	0.43***	0.51***	0.33***
	TS disability instead of pain	0.06	0.33***	0.04	0.36***	0.13*	0.16**	0.03	0.41***	0.43***	0.36***
	TS baseline dependent variable and disability	0.24***	0.38***	0.17**	0.40***	0.23***	0.37***	0.32***	0.46***	0.52***	0.37***
	CPM †, ‡	0.04	0.33***	0.04	0.30***	0.14**	0.17**	0.04	0.36***	0.34***	0.31***
	CPM baseline dependent variable score	0.25***	0.37***	0.18**	0.36***	0.24***	0.38***	0.31***	0.42***	0.51***	0.31***
	CPM disability instead of pain	0.04	0.35***	0.04	0.35***	0.14**	0.15**	0.04	0.39***	0.42***	0.38***
	CPM baseline dependent variable and disability	0.23***	0.38***	0.18**	0.40***	0.24***	0.36***	0.31***	0.44***	0.52***	0.38***
	ACR †, ‡	0.02	0.29***	0.06	0.29***	0.15**	0.18**	0.02	0.33***	0.36***	0.29***
	ACR baseline dependent variable score	0.25***	0.36***	0.18**	0.36***	0.24***	0.38***	0.31***	0.40***	0.52***	0.30***
	ACR disability instead of pain	0.04	0.33***	0.06	0.36***	0.15**	0.16**	0.03	0.38***	0.45***	0.33***
	ACR baseline dependent variable and disability	0.24***	0.38***	0.18***	0.40***	0.25***	0.37***	0.31***	0.43***	0.54***	0.34***
	>9/24 †, ‡	0.02	0.30***	0.04	0.29***	0.13*	0.17**	0.03	0.33***	0.33***	0.30***
	>9/24 baseline dependent variable score	0.24***	0.38***	0.18***	0.36***	0.23***	0.38***	0.33***	0.41***	0.51***	0.31***
	>9/24 disability instead of pain	0.03	0.34***	0.04	0.36***	0.13*	0.15**	0.03	0.39***	0.42***	0.33***
	>9/24 baseline dependent variable and disability	0.23***	0.39***	0.18**	0.40***	0.23***	0.36***	0.33***	0.44***	0.52***	0.34***
	CMT †, ‡, ¥	0.05	0.33***	0.01	0.28***	0.03	0.12**	0.10*	0.31***	0.35***	0.28***
	CMT baseline dependent variable score ¥	0.25***	0.42***	0.15**	0.40***	0.25***	0.38***	0.35***	0.39***	0.50***	0.28***
	CMT disability instead of pain ¥	0.06	0.35***	0.01	0.31***	0.03	0.11*	0.09*	0.35***	0.40***	0.32***
	CMT baseline dependent variable and disability ¥	0.24***	0.42***	0.15**	0.41***	0.24***	0.38***	0.35***	0.41***	0.50***	0.32***

CMT: Central Mechanisms Trait, **CPM**: Conditioned Pain Modulation, **HCUQ:** Health Care Utilisation Questionnaire, **HEIQ:** Health Education Impact Questionnaire, **HEIQ-CAA:** Constructive Attitudes & Approaches, **HEIQ-ED:** Emotional Distress, **HEIQ-HDB:** Health Directed Behaviour, **HEIQ-HSN:** Health Services Navigation, **HEIQ-PEL:** Positive & Active Engagement in Life, **HEIQ-SIS:** Social Integration and Support, **HEIQ-SMI:** Self-monitoring & Insight, **HEIQ-STA:** Skill & Technique Acquisition, **PPT:** Pain Pressure Detection Threshold, **PSEQ:** Pain Self-efficacy Questionnaire, **SE:** Standard Error, **TS:** Temporal Summation.
R^2 values represent the overall explanatory power of each regression model featuring CS indices as primary predictors and SM/SC follow-up scores as dependent variables. In primary analyses, each model between the CS indices and SM/SC constructs was adjusted for the same baseline variables (age, sex, pain, depression, catastrophisation and fatigue). In separate exploratory analyses, baseline disability scores instead of pain and baseline scores of each SM/SC were included as independent variables separately as well as in the same models. All variables entered in the models were based on 87 observations.
† Primary predictor. ‡ Primary hypothesis models adjusted only for age, sex, pain, depression, catastrophising and fatigue. ¥ Independent variables were age, sex, pain and quality of life.
* ≤0.05, ** <0.01, ***<0.001.