Imperial College London

Which Trial Design is Best for the Study of Physical, Psychological, and Self-Management Interventions for People with Persistent Pain?

Developing Best-Practice Trial Methods across the Pragmatic-Explanatory Continuum.

---

by

David Hohenschurz-Schmidt

A Thesis Submitted to the Department of Surgery and Cancer, Imperial College London,

in Fulfilment of the Requirements for the Degree of

Doctor of Philosophy (PhD)

March 2023

London, UK

Supervisors: Prof Andrew SC Rice, Imperial College London; Dr Jerry Draper-Rodi, University College of Osteopathy; Prof Alison McGregor, Imperial College London; Prof Lene Vase, Aarhus University; Dr Whitney Scott, King's College London

Statement of Originality

I hereby confirm that the work presented in this document is my own, except where I have clearly indicated otherwise through appropriate references or statements about the contribution of others. I have not used any unauthorised assistance.

Declaration of Interest

The thesis author has received honoraria from osteopathic higher education institutions for teaching activities and works in private osteopathic practice. The author was also in receipt of a personal PhD stipend and other charitable funding as part of this project, as detailed below. For the work on pragmatic trials in pain research, the author received a personal honorarium from the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT). Unrelated to this work, the author has received consulting fees from Altern Health Ltd., a company developing a Digital Therapeutics device for pain management.

Funding and the Role of Funders

The PhD project was funded by the Alan and Sheila Diamond Charitable Trust, including a personal PhD stipend to the author. Additional funding for individual project aspects was obtained through the author in the form of a personal £20,000 project grant from *The Osteopathic Foundation* (oF) and a personal £2,500 travel fellowship by the *Society for Back Pain Research*. Both grants were awarded as part of open competitions. Quarterly project reports were submitted to the oF and, as part of the funding agreement, the research will be presented at funder-related conferences once completed. These funders had no role in the planning, conduct, or reporting of the project. A honorarium of US$ 17,000 was received from IMMPACT for the work on pragmatic trials in pain research. This funder did take an active role in the planning, conduct, and reporting of the related study aspects as described below.

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non-Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Word Count

This thesis consists of approximately 85,000 words, which excludes appendices and reference list, but includes tables, footnotes, and captions.

Thesis Abstract

This PhD project aimed to address major challenges of designing and conducting randomised controlled trials (RCTs) for physical, psychological, and self-management (PPS) interventions for people with pain. Such challenges include designing adequate 'sham' control interventions and ensuring participant and provider blinding to control for context-dependent effects. Studies are also commonly criticised for limited generalisability of results beyond the immediate study population and setting; high generalisability being a typical objective of so-called pragmatic trials. The primary objective of this PhD was to provide evidence-based guidance for trial design that is specific to PPS interventions and covers best-practice methods across the pragmatic-explanatory spectrum. A secondary objective was to explore the developed recommendations on the example of a feasibility trial protocol, studying a multimodal intervention for people with painful diabetic neuropathy.

The project involved two systematic reviews of trial methods. One describing self-declared 'pragmatic' and 'comparative effectiveness' pain trials, and another PPS trials employing 'sham' or 'placebo controls'. The latter involved a meta-regression analysis, which demonstrated that the level of similarity between experimental and control interventions can bias trial outcomes. Informed by these reviews, two independent consensus processes were conducted. One process led to two guidance statements for pragmatic trials in pain research. This was based on expert-led discussions as part of the dedicated 2020 Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) meeting. The second consensus process involved a modified Delphi survey with trial designers and placebo effect researchers, interviews with patient partners, and live consensus discussions. This process resulted in a comprehensive guideline for control intervention design, conduct, and reporting in PPS trials. The so-called 'CoPPS Statement' was published in the BMJ and adopted by the EQUATOR network library. Finally, a multimodal osteopathy-based intervention was created, and a feasibility trial protocol written to test the usability of the guidance frameworks. The systematic reviews,

consensus processes, and feasibility trial protocol development are described in detail in this thesis.

As a result of this PhD, trial designers now have access to two comprehensive evidence-based frameworks to guide their decision-making in the design, conduct, and analysis phases of PPS trials. End-users of trials can employ these frameworks to guide their interpretation of such studies. Overall, these novel frameworks address major methodological criticisms of earlier trials in a manner that is specific to PPS interventions and may allow for more rigorous and transparent future research.

Furthermore, this project highlighted that while nonpharmacological interventions are distinct from drugs, RCTs of PPS interventions can still be of high quality. The developed methodological guidance statements facilitate this by accounting for differences to drug research, providing rigorously developed and innovative resources for researchers designing and conducting RCTs of PPS interventions for people with pain and other conditions. The usability evaluation as part of the feasibility trial protocol development showed the CoPPS Statement to be comprehensive and clear for the design of control interventions in an efficacy-focused trial. The IMMPACT statements were less prescriptive but were found relevant for trials across pragmatic-explanatory spectrum. However, the distinctiveness of complex nonpharmacological intervention research will remain a challenge for the involved professions, for evidence synthesis, and for policy decision-making.

Project limitations include a lack of practical testing of the developed recommendation frameworks prior to their publication and insufficient involvement of people from regions other than Europe, the US, and Australasia. Future research needs to determine the applicability of the developed frameworks to other therapeutic areas and settings. Their long-term impact on trial quality and practical usefulness will also have to be evaluated to identify potential needs for refinement.

Table of Contents

Thesis Abstract **4**

Abbreviations and Acronyms **13**

Chapter 1: Thesis Introduction **15**

Chapter summary *15*

Fundamental terms and concepts *16*

Thesis background *18*

Thesis structure *24*

Situating the researcher and the PhD project *25*

Efficacy Section: Control Interventions in Efficacy and Mechanistic Trials **27**

Section summary *27*

Section background *28*

Chapter 2: Systematic Review of (Sham) Control Intervention Methods and Meta-Analysis **32**

Methods *32*

Results *53*

Summary of review findings *80*

Discussion of review findings *82*

Limitations of the review *84*

Chapter 3: Guideline Development for Control Intervention Design, Implementation, and Reporting ... **89**

Background *89*

Methods *91*

Results *109*

Efficacy Section: Discussion and Limitations *128*

Effectiveness Section: Pragmatic and Comparative Effectiveness Trials 138

Section Summary 138

Section background 139

Chapter 4: Systematic Review of Methods in Pragmatic and Comparative Effectiveness Trials 145

Introduction 145

Methods 146

Results 152

Discussion 176

Conclusion and next steps 179

Chapter 5: Development of Best-Practice Considerations for Pragmatic Trials 181

Introduction 181

Methods 185

Results 187

Effectiveness Section: Discussion and Limitations 223

Practical Section: Applying Recommendations to a Specific Example 229

Chapter 6: Trial Protocol and Control Intervention for a Novel Painful Diabetic Polyneuropathy Treatment 229

Chapter and section summary 229

Introduction 231

Methods 236

Results 244

Discussion and Limitations 290

Discussion Section 295

Chapter 7: General Discussion, Limitations, and Outlook 295

Chapter summary 295

Discussion of main project outputs *296*

Future opportunities and remaining challenges for PPS research *304*

Limitations and alternative approaches *306*

Learning journey and reflections *309*

Conclusion *311*

Acknowledgements **312**

Appendices **314**

Appendix 1: Systematic review and crowd-related content *314*

Appendix 2: Consensus methods, additional content *316*

Appendix 3: IMMPACT Statement, additional content *338*

Appendix 4: NeuOst intervention, additional content *339*

Appendix 5: Supplementary information for NeuOst feasibility trial *351*

Appendix 6: Academic development and awards *365*

Appendix 7: Outreach activities and additional contributions *370*

Appendix 8: Peer-reviewed publications *376*

References **384**

List of tables

Table 1: Glossary. 16

Table 2: Characteristics of crowd members 36

Table 3: Eligibility criteria for inclusion into the systematic review. 39

Table 4: Features for which the level of similarity between control and index interventions was rated during the data extraction process. 45

Table 5: Overview of included studies.. 55

Table 6:Overview of employed control interventions per therapy type. 57

Table 7: Matching of provider-related factors. 63

Table 8: Additional features of control intervention design and implementation. 64

Table 9: Reporting of control intervention development processes and theoretical considerations. 65

Table 10: Meta-analyses results of test over control interventions per therapy subgroup and type of outcome measure.. 74

Table 11: Regression models with similarity ratings.. 76

Table 12: Method of rating the resemblance between control and experimental interventions. 86

Table 13: Rating scale employed during Delphi rounds. 98

Table 14: Eligibility criteria for panellists 100

Table 15: Baseline demographic data of initial panel. 109

Table 16: Participant characteristics of patient consultation interviews. 112

Table 17: Checklist of recommendations for the development and implementation of a control intervention in efficacy and mechanistic trials of physical, psychological, and self-management interventions (CoPPS Checklist). 116

Table 18: Eligibility criteria. 147

Table 19: Study characteristics regarding treatment modalities, participant demographics, and trial settings. 153

Table 20: Methods of trial design. 157

Table 21: Reporting of selected items of the 2010 update of the Consolidated Standards of Reporting Trials (CONSORT) statement (Schulz et al., 2010)[1] and all items of the extension for the reporting of pragmatic trials..... 159

Table 22: Correlation analysis between total sample size at randomisation and domain-specific PRECIS-2 ratings. 173

Table 23: Correlation analyses amongst trial methods with ratings of trial pragmatism (PRECIS-2 scores), randomisation methods, and analysis method. 174

Table 24: A 'typical' pragmatic or comparative effectiveness RCT in pain research. 176

Table 25: Possible bias in clinical research and proposed considerations for methods to minimise bias in pragmatic trials of pain treatments. 191

Table 26: Ethical considerations in pragmatic trials 198

Table 27: Concomitant treatments in pragmatic analgesia trials..... 208

Table 28: Recommendation for the reporting of pragmatic trials. 222

Table 29: *NeuOst* working group. 244

Table 30: Components of the NeuOst intervention. 252

Table 31: Participant eligibility criteria for the *NeuOst* feasibility RCT. 261

Table 32: Completed CoPPS Checklist for the control intervention used in the *NeuOst* feasibility trial..... 267

Table 33: Primary feasibility trial outcomes and progression rules. 277

Table 34: Usability evaluation of the IMMPACT statements..... 283

Table 35: Usability evaluation of the CoPPS Statement 287

Table 36 Patient consultation interview script 317

Table 37: Summary of consensus workshops. 320

Table 38: The lived-experience perspective on NeuOSt..... 330

Table 39: Questions to consider before using blinding. 335

Table 40: Alternative designs to parallel-group RCTs potentially suitable for pragmatic research questions. 338

Table 41: Intervention components of NeuOst. 339

Table 42: *NeuOst* intervention protocol and fidelity checklist. 351

Table 43: PRECIS-2 table for NeuOst feasibility trial 362
Table 44: Conference contributions and presentations. 366
Table 45: Awards and personal grants.. 368
Table 46: Overview of teaching and supervision activities 369

List of figures

Figure 1: Number of PubMed entries per year relating to clinical trials of physiotherapy, osteopathy, chiropractic, manual therapy, rehabilitation, or psychotherapy 20

Figure 2: PRISMA flow diagram for control interventions review. 54

Figure 3: Similarity between tested intervention and placebo control intervention 60

Figure 4: Average similarity ratings categorized by therapy type 62

Figure 5: Forest plot showing all studies for Bang's blinding findex (BI) 69

Figure 6: Risk of Bias assessment for all comparisons with 20 or more participants per group 72

Figure 7: Flowchart of CoPPS guidance generation process 93

Figure 8: PRISMA flow diagram for pragmatic trials review 152

Figure 9: Average PRECIS-2 scores per domain for all included trials. 167

Figure 10: Flowchart of IMMPACT XXIV guidance development. 185

Figure 11: MRC Framework for developing and evaluating complex interventions. 237

Figure 12: Process flowchart of NeuOst intervention development and evaluation. 238

Figure 13: Flowchart of the NeuOst feasibility trial. 260

Figure 14: An example of the use of Microsoft Teams chat space for the organisation of data extraction between crowd-members. 314

Figure 15: Analysis map of NeuOst patient partner interviews 329

Figure 16: Feedback from a Delphi panellist and CoPPS co-author 334

Figure 17: Logic model of NeuOst complex intervention. 350

Abbreviations and Acronyms

Abbreviation / acronym	Meaning
ACT	Acceptance and Commitment Therapy
ACTTION	Analgesic, Anesthetic, and Addiction Clinical Trials, Translations, Innovations, Opportunities, and Networks
AMED	Allied and Complementary Medicine Database
ASPIRE	Applying Surgical Placebo in Randomised Evaluations checklist
CBT	Cognitive-bahavioural therapy
CENTRAL	Cochrane Central Register of Controlled Trials
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CONSORT	CONsolidated Standards of Reporting Trials
CoPPS	Recommendations for the Development, Implementation, and Reporting of Control Interventions in Efficacy and Mechanistic Trials of Physical, Psychological, and Self-Management Therapies
DGSS	Deutsche Gesellschaft für Schmerzforschung (German Pain Research Society)
DM	Diabetes Mellitus
DN4	Doleur neuropathique en 4 questions (diagnostic questionnaire for neuropathic pain)
DPN	Diabetic peripheral neuropathy / Diabetic polyneuropathy
EHR	Electronic health record
EQUATOR	Enhancing the QUAlity and Transparency Of health Research
fMRI	Functional magnetic resonance imaging
GAPPA	Global Alliance of Partners for Pain Advocacy
GOsC	General Osteopathic Council (UK regulatory body of the osteopathic profession)
GRIPP	Guidance for Reporting Involvement of Patients and the Public
HbA1c	Glycated haemoglobin
HIV	Human Immunodeficiency Virus
IASP	International Association for the Study of Pain
ICMJE	International Committee of Medical Journal Editors
IMMPACT	Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials
ITT	Intention-to-treat (analysis method)
MeSH	Medical Subject Headings
MRC	Medical Research Council
MSK	Musculoskeletal
NCOR	National Council for Osteopathic Research
NeuOst	Augmented osteopathic care for people with painful diabetic neuropathy (name of an intervention)
NHS	National Health Service
NIH	US National Institutes of Health
NIHR	UK National Institute for Health and Care Research
PCORI	Patient-Centered Outcomes Research Institute
pDPN	Painful diabetic peripheral neuropathy / painful diabetic polyneuropathy

PEDro	Physiotherapy Evidence Database (PEDro)
PIS	Participant information sheet
PPI	Patient and public involvement
PPS	Physical, psychological, and self-management interventions
PRECIS	PRagmatic–Explanatory Continuum Indicator Summary
PRISMA (and PRISMA-P)	Preferred Reporting Items for Systematic Reviews and Meta-Analyses (or Protocols)
PROSPERO	International Prospective Register of Systematic Reviews
RCT	Randomised clinical trial
RoB	Risk of bias
SIPS	Society for Interdisciplinary Placebo Studies
SWOT analysis	A project management technique used to identify Strengths, Weaknesses, Opportunities, and Threats regarding an economic situation or a project
T2DM	Type 2 Diabetes Mellitus
TAU	Treatment or care as usual
TAU	Treatment as usual
TIDieR	Template for Intervention Description and Replication
UCO	University College of Osteopathy (London, UK)

Chapter 1: Thesis Introduction

Chapter summary

The opening chapter of this thesis provides a comprehensive overview of the PhD project's background and its significance in improving clinical trial methods for physical, psychological, and self-management (PPS) intervention research. The project aligns with a growing interest in nonpharmacological and multimodal interventions for people with pain, especially persistent pain.

Randomised controlled trials (RCTs) are an important source of evidence about an intervention's effect under ideal circumstances, studied in explanatory trials, and under conditions more akin to normal clinical practice (studied in trials with a more 'pragmatic' attitude to trial design). However, despite an increasing number of trials, research on PPS interventions suffers from methodological challenges which create uncertainty about treatment effects and hinder their consideration for clinical guidelines and in health-policy decision-making. This chapter will describe these challenges in detail, presenting the current state of debate at both ends of the pragmatic-explanatory continuum of trials. In doing so, key differences to drug research are outlined and fundamental concept are introduced.

Additionally, the chapter will outline the scientific approach used to identify gaps in the field and to develop comprehensive solutions to address these challenges. Finally, the thesis structure will be presented to provide a clear roadmap for the reader, and the thesis author is introduced to situate the project within their clinical and academic background.

Fundamental terms and concepts

Before describing current debates and developments regarding clinical trials in nonpharmacological intervention research for people with pain, several fundamental concepts and terms need to be introduced.

First, the project refers to 'physical, psychological, and self-management interventions'. Abbreviated 'PPS', the term includes any physical and manual therapies, exercise, psychological interventions, and education or other self-management approaches (See glossary in Table 1 for a more comprehensive definition). This group of interventions shares many challenges of trial design and conduct. Not included in the PPS term, and thus beyond the scope of this thesis, are other nonpharmacological interventions, notably surgery, devices, nutrition, and acupuncture, for which different challenges and opportunities for trial design exist.

The glossary in Table 1 provides an overview of other relevant terms and concepts relating to clinical trial design and the phenomena studied or controlled for in trials. These terms are discussed in additional nuance in the relevant thesis sections.

Table 1: Glossary. Definitions and explanations were derived from the provided references and/or they were agreed upon by the participants of the expert consensus processes of this PhD project. The same definitions are published in Hohenschurz-Schmidt et al. (2023a, 2023d). RCT, randomised controlled trial.

Terms	Definition, explanation, and examples
Physical, psychological, and self-management (PPS) therapies	The PPS term covers non-surgical, nonpharmacological interventions, including all forms of manual and physical therapy, exercise and rehabilitation therapy, conversation-based and psychological therapies, mind–body, spiritual, religious, and other non-material healing practices, and educational interventions. The term excludes any interventions that require the skin to be pierced surgically or with needles or the ingestion or introduction of substances (drugs, supplements, nutritional interventions). Because other opportunities and challenges for trial design exist, therapies based on meridian and *Qi* concepts (including traditional acupuncture and

	acupressure) are also excluded, as well as any interventions in which therapists rely on the use of devices (e.g., ultrasound, transcutaneous electrical nerve stimulation, laser, transcranial stimulation, shockwave, spinal stimulation, splinting, and braces, but not exercises using, for example, elastic bands).
Efficacy	Efficacy is the extent to which an intervention provides benefit under ideal circumstances ("Can it work?") (Haynes, 1999)
Explanatory RCT	An RCT that tests the benefits and/or harms of a treatment under relatively ideal conditions, aimed primarily at investigating a scientific or biological problem (Schwartz and Lellouch, 1967; Loudon et al., 2015a; Ford and Norrie, 2016).
Effectiveness	Effectiveness assesses whether an intervention is beneficial when provided under usual circumstances of healthcare practice ("Does it work in practice?") (Haynes, 1999)
Pragmatic RCT	An RCT intended to directly inform clinical or health policy decision-making, usually investigating therapeutic effectiveness or comparative effectiveness under conditions similar to clinical practice (Schwartz and Lellouch, 1967; Loudon et al., 2015a; Ford and Norrie, 2016). Used interchangeably with 'effectiveness RCT'.
Internal validity	"*Internal validity describes the […] accuracy of the study results by minimizing error. Thus, internal validity is the degree to which changes in the dependent variable can be attributed to the intervention and is maximized by decreasing bias using design features such as random assignment, allocation concealment, and blinding*" (p. 164) (Wright et al., 2021).
Generalisability	The degree to which trial results may be considered valid in and applicable to participants, practitioners, interventions, outcome measures, and settings outside the respective trial (Cronbach and Shapiro, 1982; Rothwell, 2005a, 2010). Used interchangeably with 'external validity' and 'applicability' (Rothwell, 2005a, 2010)
Placebo and nocebo effect	"*The placebo and nocebo effect [are changes in health outcomes that are] specifically attributable to placebo and nocebo mechanisms, [such as] the neurobiological and psychological mechanisms of expectancies [and learning]. These mechanisms are shaped, for example, by verbal instruction, or nonverbal or situational cues that affect treatment expectancies.*"(p.3) (Evers et al., 2018) For details on placebo and nocebo mechanisms, see, e.g., Colloca and Barsky (2020), Benedetti (2020), and Benedetti et al. (2022).

Control intervention (in the context of efficacy or mechanistic trials)	Definition: Procedures delivered to trial participants in the control group, specifically designed to test the efficacy or mechanism of the test intervention. Although also known as 'sham', 'placebo', and 'attention' control interventions, this thesis specifically avoids these terms based on an expert consensus on best-practice terminology described in Chapter 3. These control interventions are distinct from usual care, other recognised treatment, and no-treatment controls. Although the latter are also commonly called 'controls', they do not control well for expectancy effects, do not allow for blinding, and are neither specifically designed for a given trial nor very common in efficacy or mechanistic trials. Such comparator arms would be better termed a 'comparator intervention.'

Thesis background

The need for better nonpharmacological intervention research

Two phenomena form the immediate backdrop to this thesis: First, decision-makers' interest into nonpharmacological treatment options for people living with pain has risen sharply in the last decade, not least due to the continuingly high public health burden from pain (Pizzo et al., 2011; Rice et al., 2016) and because of the US opioid catastrophe which makes the shortcomings of analgesic drugs strikingly evident (Krashin et al., 2016; Tucker et al., 2020; Wewege et al., 2023). Similarly, more and more commonly performed surgical procedures for persistently painful conditions are shown to be no more effective than a surgical placebo (Thorlund et al., 2015; Gu et al., 2017; Lähdeoja et al., 2020; Karjalainen et al., 2022). These developments narrow the focus for effective or new treatment options for primary pain to physical, psychological, and self-management interventions – the topic of this thesis – and possibly devices and traditional therapies.

Secondly, as more funding and research efforts are invested, the need for improved RCTs in the field of PPS becomes ever more apparent. Current clinical guidelines recommend various non-drug therapies as a first-line treatment for low back and chronic musculoskeletal pain

(Bernstein et al., 2017; Kligler et al., 2018; Qaseem et al., 2017). However, most recommendations are based on low or moderate quality evidence (NICE, 2016; Stochkendahl et al., 2018), a widespread concern in PPS interventions (Fregni et al., 2010; Chen et al., 2013; Kumar et al., 2013; Finnegan-John et al., 2013; Cerritelli et al., 2016; Alvarez et al., 2021; Skelly et al., 2020). A lack of high-quality research means that the role of many of these therapies in the prevention, treatment, and management of people with pain is unclear. While small sample sizes were often the main point of critique, other issues have gained attention even in large clinical trials, notably concerns regarding the generalisability of findings and adequacy of 'sham' control interventions (Boutron et al., 2004, 2008; Bushnell et al., 2021; Lavazza et al., 2021; Maddox et al., 2022). In both areas, attempts have been made to conduct high-quality RCTs, namely with large pragmatic trials and with carefully designed control interventions in efficacy trials. Nonetheless, there is insufficient best-practice guidance for pragmatic trials as well as for specifically controlled efficacy trials in PPS. More fundamentally, when researchers want to investigate the effects of a PPS intervention, they are faced with the basic decision of whether to conduct a trial comparing two or more recognised treatments with one another, or a treatment against a so-called 'sham' control. As we will see later, it appears that these decisions were often made based on practical or ideological considerations, and less for scientific reasons; not least because definite guidance was missing when this PhD project commenced. Once decided, methodological frameworks available for pragmatic trials had considerable shortcomings and did not consider specific challenges of nonpharmacological intervention research. Unifying recommendations for the design of high-quality efficacy trials of complex PPS interventions was lacking entirely. Solutions were required that spanned the explanatory-pragmatic continuum of trial design, were broad enough to apply to the heterogeneous field of PPS interventions, and that clearly addressed the challenges arising from applying methods developed for drug research to complex nonpharmacological interventions.

Challenges of physical, psychological, and self-management intervention trials

Many PPS interventions currently on offer have been developed organically through clinical practice, not as well-structured research endeavours (Liem, 2018; Nicholls, 2017; Schiller, 2021). Nonetheless, since the 1980s, the evidence-based practice movement and quests for professional recognition have promoted the testing through clinical trials in many PPS areas (Leach, 2006; Moseley et al., 2020), including pain research (Figure 1).

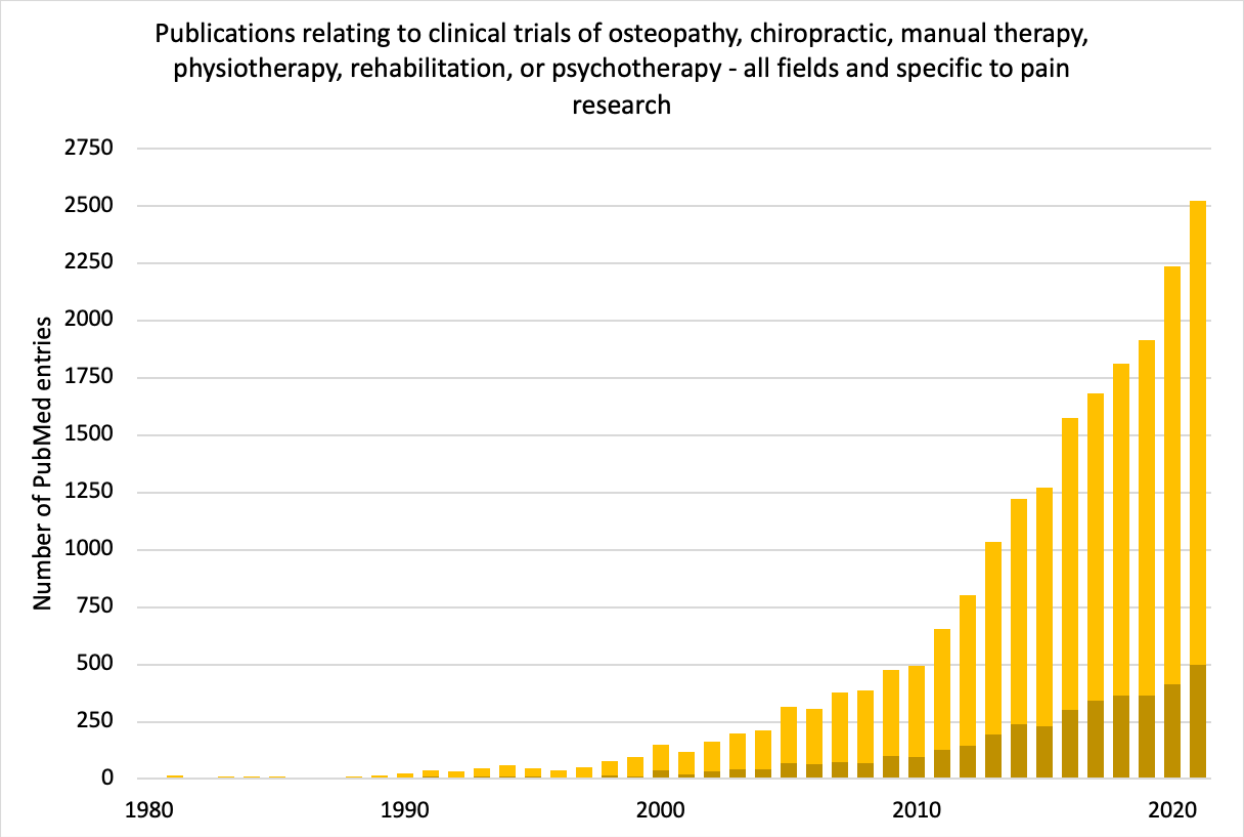


Figure 1: Number of PubMed entries per year relating to clinical trials of physiotherapy, osteopathy, chiropractic, manual therapy, rehabilitation, or psychotherapy (in general, and in pain research). The dark parts of the bars indicate trials related to 'pain', lighter parts are results independent of a specific patient population. Complete Search query: (("osteopathy"[Title/Abstract]) OR ("chiropractic"[Title/Abstract]) OR ("physiotherapy"[Title/Abstract]) OR ("psychotherapy"[Title/Abstract]) OR ("rehabilitation"[Title/Abstract]) OR ("manual therapy"[Title/Abstract])) AND (("clinical trial"[Title/Abstract]) OR ("controlled trial"[Title/Abstract]) OR ("RCT"[Title/Abstract])) AND ("pain"[Title/Abstract]); or without the final pain-related search term (dark bar portions). Limits: 1980 to 2021. Date of search: 09 August 2022.

While RCTs have become increasingly common in the past 30 years, methodological criticism of PPS trials still abounds. Chief amongst those are concerns about adequate bias control on one end (Boutron et al., 2008; Alvarez et al., 2021) – where blinding is identified as a major challenge in complex intervention trials (Boutron et al., 2004) – and applicability to clinical practice on the other end (Huebschmann et al., 2019; Gordon et al., 2020a). Arguably, many methodological challenges stem from difficulties in applying the clinical trial model from drug research to the evaluation of complex interventions. Compared to pharmacological interventions, PPS interventions are usually more complex and of an inherently participatory and interpersonal character. These characteristics may require adaptations of evaluation methodology in nonpharmacological clinical trials. The lack of direct transferability of trial methods from drug to complex intervention research complicates the consideration of PPS interventions in clinical guidelines or health care policies, where comparisons against other care options are invariably made, directly or indirectly (Cherkin, 2021).

At the explanatory end of the spectrum of RCTs, there is the challenge of designing adequate 'sham' control interventions in PPS trials: To start with, these controls cannot be expected to blind providers as a placebo pill might (Boutron et al., 2004). Further, such control interventions cannot be expected to be completely 'inert' due to their complex nature (being modelled on treatments that consist of many components that may each act through multiple mechanisms (Bialosky et al., 2017a; Skivington et al., 2021)). Lastly, the rich context of nonpharmacological control interventions (Moerman, 2013) may lead to larger placebo effects than pill-based placebos (Meissner et al., 2013; Vase et al., 2015b). Triallists have responded to these challenges with a plethora of approaches to 'sham' control intervention design, some more and some less credible for trial participants. When poorly conducted, such control interventions may lead to exaggerated effect sizes, doubts about the validity of findings, and not provide clear information about the efficacy of supposedly therapeutically active components of the tested interventions (Hohenschurz-Schmidt et al., 2021, 2022a). Conversely, when control interventions closely resemble the tested treatment, they may be so contextually rich as to lead to smaller between-group effects than what would be considered relevant in placebo-controlled drug trials (Baskin et al., 2003; Kaptchuk et al., 2006; Zhang et al., 2008; Meissner et al., 2013).

On the other end of the pragmatic-explanatory continuum, pragmatic trials of PPS interventions have become more common in the field of pain research and other areas (Lewis, 2016; Sepehrvand et al., 2019; Gordon et al., 2020b; Hohenschurz-Schmidt et al., 2022c). A typical view was that conventional sham-controlled trials were inadequate in addressing the complete spectrum of clinical inquiries. A pragmatic trial, on the other hand, can investigate if a particular therapy works as well as, or better than, established care when tested in a diverse population and in real-world settings (Cherkin, 2021), regardless of the underlying mechanisms of the treatment's effectiveness. Pragmatic and comparative effectiveness trials of complex interventions share several challenges with drug research, including variable adherence and treatment delivery in real-world settings, and the recruitment of large samples. However, while drug development follows a tightly regulated pipeline (where pragmatic, or phase IV, trials are usually reserved for post-approval testing), no such regulatory standards exist for PPS interventions. Instead, many PPS interventions have become routinely available without rigorous efficacy testing. The attractiveness of pragmatic trials for the PPS field is probably also partly due to the challenges of conducting efficacy trials (Williams et al., 2021). At the same time, pragmatic trials arguably sacrifice internal validity for the replication of real-world clinical practice; and prominent guidelines for pragmatic trial design touch only briefly on the issue of bias control (Loudon et al., 2015a). This may lead to trials that resemble clinical practice but are at a high risk of bias. Taken together, this situation in pragmatic trial research holds the danger of uninterpretable, costly trials of PPS interventions that have no solid evidence to support their efficacy or supposed mechanisms of action.

Returning to the importance of acknowledging differences between drug and complex intervention research: Some methodological differences between drug and PPS trials are inevitable, but this may not mean that high-quality RCTs are impossible in PPS research. At this point in time, the significance of necessary differences in trial design for the interpretation of findings remains largely unclear. For example, one can expect to obtain smaller effect sizes from pragmatic trials of complex interventions than from most placebo-controlled drug trials (Keefe et al., 2022), but there is no meta-analytical evidence that would tell us by how much these

effects differ. Indeed, where trial designers uncritically copied trial methods from drug research, the results were often negative and expenses large: Hawe et al. (2004), discuss the example of an overly standardised complex intervention which, in their view, undermined effectiveness in a pragmatic trial; Cherkin (2021) criticises the choice of headache frequency as a primary outcome measure in a trial for mindfulness-based stress reduction, arguing that this measure is commonly used in drug trials but does not reflect actual patient values nor the objectives of the tested intervention (which, Cherkin argues, is improved functioning and response to pain rather than symptom reduction). Finally, Cherkin raises the question of the appropriate control condition - an issue we will return to.

There is a lack of best-practice frameworks to guide trial design in PPS research, as well as the adequate interpretation of findings. The above challenges of PPS research not only mean that the conducted trials are often methodologically weak and designs poorly justified, but also that pertinent research questions are not answered or that non-RCT research designs are chosen. While alternative approaches (such as observational studies, single-case experimental designs, process evaluations, mechanistic experiments, qualitative approaches, etc.) may contribute useful complementary information, they cannot substitute the clinically relevant causal data regarding efficacy and effectiveness provided by high-quality RCTs and their meta-analysis (Jonas, 2005; Djulbegovic and Guyatt, 2017). It is therefore in the interest of patients, PPS professions, and society at large to improve the quality and usefulness of RCTs in the field of PPS interventions for people with pain.

Thesis structure

The present PhD project tackled the methodological challenges of PPS trials from both ends of the explanatory-pragmatic spectrum. In one work stream, the focus was on how internal validity of efficacy trials could be increased through better control intervention design, conduct, and reporting. This work is presented in chapters 2 & 3 - the 'Efficacy Section'. In a parallel work stream, described in chapters 4 & 5 - the 'Effectiveness Section', the question was how trials that aim to directly inform clinical or health policy decision-making need to be improved to meet their objectives. Both sections contribute to the fundamental question and thesis title of "Which trial design is best", and under which conditions.

Apart from the present general background chapter, each thesis section will detail the pertinent challenges for PPS research in the respective area. Then, systematic reviews are presented that provide a complete up-to-date picture of 'sham' control methods on one hand, and methods in pain-related pragmatic and comparative effectiveness trials on the other. Using these insights to develop evidence-based and expert-informed recommendation frameworks, the next chapter in each section describes the methods and results of guideline development processes. In each case, these culminated in comprehensive decision-making frameworks that can guide trial designers and that acknowledge the particularities of PPS interventions and pain research.

Next, chapter 6 (the 'Practical Section') shows how these frameworks can inform the development of a trial protocol, using the example of a newly developed multimodal intervention for people with painful diabetic neuropathy. This trial development showed that both frameworks can successfully guide the fundamental decision whether to conduct an efficacy-focused trial or a more pragmatic trial. Deciding on an efficacy-focused RCT with a control intervention, it was not only found that the developed methodological guideline for control interventions was highly useful and comprehensive, but it was also found that the general recommendation to consider internal and external validity for each design decision was highly relevant also in efficacy-orientated RCTs. This project illustrated that rigorous RCTs of PPS interventions are possible despite differences to drug research. Regardless, there is no one-size-

fits-all answer to the question of "which trial design is best", but the project has contributed to the methodological improvement across the spectrum of research questions and trial designs.

At the end of each chapter, and in the dedicated final chapter 7, project methods and outputs are reflected on critically. Chapter 7 also contextualises the project findings, identifies future research needs, and outlines the author's learning journey.

Appendices provide additional detail on individual research steps, and a brief account of the author's additional contributions as well as their engagement with the scientific community, clinicians, and the public. The main peer-reviewed articles published as part of this PhD project are also appended.

Situating the researcher and the PhD project

By describing how this project was initiated and who the researcher is, potential influences on the course and content of the research are made transparent.

The project was funded through a charitable trust with long-standing links to the University College of Osteopathy (UCO). With the objective to strengthen the project and the UCO's research environment, the funds were used to develop a PhD studentship at the Pain Research Group of Prof Andrew SC Rice at Imperial College London, with co-supervision from Dr Jerry Draper-Rodi, Senior Research Fellow at UCO. After an open application process, the author was awarded the studentship in 2019. The author is a UCO graduate of 2014 (then known as the British School of Osteopathy) and has practised as an osteopath since; first in private practice in Germany, where they were also involved in teaching (theoretical and clinical), and then again in London since 2017. The author's clinical practice evolved from a rather biomechanical approach initially, to including craniosacral approaches influenced by a not evidence-based paediatric osteopathy diploma. After a change in environment and initial exposure to contemporary pain research, the author's clinical practice then again developed away from these esoteric models to a psychologically oriented framework where communication, education, and mind-body exercises are used to promote patients' understanding of and engagement with their situation; and manual therapy and dry needling to provide short-term pain relief.

In 2018, the author obtained a MSc in neuroscience from King's College London. The respective MSc thesis involved the analysis of functional magnetic resonance imaging (fMRI) and physiological data, assessing the relationship between functional brain networks, pain sensation, and the autonomic nervous system (Hohenschurz-Schmidt et al., 2020a). Throughout the PhD project, the author worked in clinical practice in Brixton, South London (seeing about five patients a week) and continued teaching on osteopathic courses. In their academic work, the author was thus very aware of the complex realities of clinical practice, the need for high-quality clinical trials, the challenges of communicating research to clinicians and students, and the potential for research to conflict with the professional identity of healthcare professionals. The author hopes that such awareness is reflected in the main academic outputs of this project. It has certainly informed several clinician- and patient-oriented side projects and engagement activities (see Appendix 7: Outreach activities and additional contributions').

Apart from this biographic situating of the researcher and the PhD project, ethical and philosophical stances will be presented at relevant instances in the thesis.

Efficacy Section: Control Interventions in Efficacy and Mechanistic Trials

Section summary

Efficacy trials in physical and psychological intervention research with specifically-designed control interventions (so-called 'sham' or 'placebo controls') are either rarely conducted or subject to methodological criticism. Since blinding effectiveness and participant expectations are usually not measured, it is also often unclear whether the control interventions account for the biases they were intended to address. This project's dedicated systematic review and meta-analysis of methods also showed that the degree of similarity between control and tested intervention influences trial outcomes. This novel finding underlines the fundamental importance of control intervention design and conduct for PPS efficacy and mechanistic trials, also suggesting a need for standardisation.

A rigorous consensus process with experts from PPS trials, researchers of placebo effects, and patient partners resulted in the following key recommendation: Control interventions in efficacy and mechanistic trials of PPS interventions should resemble the tested treatment in all components other than those whose effect the trial is designed to study. Apart from this similarity principle, the newly developed CoPPS Statement (Recommendations for the Development, Implementation, and Reporting of **Co**ntrol Interventions in Efficacy and Mechanistic Trials of **P**hysical, **P**sychological, and **S**elf-Management Therapies) postulates a best-practice framework for development, conduct, reporting, and patient involvement. In doing so, the CoPPS Statement provides a basis for more high-quality efficacy trials in psychological and exercise intervention research and unifies control intervention design in all areas of PPS efficacy and mechanistic research. In next steps, the guidance needs to be tested practically on exemplary trial scenarios.

Section background

"*Efficacy is the extent to which an intervention does more good than harm under ideal circumstances*" (p. 652), thus relating to the question "Can this treatment work?" (Haynes, 1999). Efficacy or explanatory RCTs, then, test the benefits and/or harms of a treatment under relatively ideal conditions, aimed primarily at investigating a scientific or biological problem (Schwartz and Lellouch, 1967; Loudon et al., 2015a; Ford and Norrie, 2016). Even more than 'ideal' circumstances to demonstrate an intervention effect, the concept of efficacy trials implies a focus on bias control to maximise the validity of findings (Dworkin et al., 2021). In the life cycle of complex interventions, efficacy trials play a different role to implementation or 'pragmatic' studies that examine treatment effectiveness under usual circumstances of healthcare practice, asking "Does it work in practice?" (p. 652, Haynes, 1999; Skivington et al., 2021). Identifying four phases of complex intervention research (development or identification of the intervention, feasibility, evaluation, and implementation testing), the 2021 update of the Medical Research Council guidance for developing and evaluating complex interventions contents that "*[a] research programme might begin at any phase, depending on the key uncertainties about the intervention in question*" (p. 3, Skivington et al., 2021).[1] Key uncertainties which efficacy trials are designed to answer include whether the examined intervention contains elements that can produce clinical benefit (or harm) in themselves, i.e., beyond the placebo effect and other confounding factors (such as heterogeneity in patient populations, treatment delivery and adherence, provider expertise, and use of concomitant interventions). Relatedly, mechanisms of action may be investigated in explanatory trials (then often called 'mechanistic' trials).

To answer such research questions, efficacy trials, per definition, invest considerable methodological effort into bias control (Thorpe et al., 2009; Keefe et al., 2022). Apart from measures to reduce heterogeneity, major concerns are assessment bias, performance bias, and

[1] Whether efficacy counts as a 'key uncertainty' when an intervention is already routinely employed in clinical practice is contentious – more so in expert circles than in the published literature (as experienced during the expert consultations as part of this PhD) – the argument being that "it doesn't matter how it works as long as it works in practice". This is, however, a slippery slope both ethically and economically, and a best-practice approach will be proposed in the chapter on pragmatic trials.

the placebo effect. Common design features are therefore blinding and specifically designed control interventions to facilitate blinding and/or balance expectancy and learning effects between groups. These control interventions are known as 'sham' or 'placebo controls',[2] and represent one of the main challenges of efficacy trials in PPS research.[3] In mechanistic trials, these controls may be used to isolate the treatment components to be tested. A concept from drug research, the essential notion is that such control interventions be 'inert' for the ingredient of interest and otherwise resemble the tested treatment (Hróbjartsson and Miller, 2011). Unfortunately, translation of this principle from drug to complex intervention research is laden with problems. First, mechanisms of PPS therapies are often unclear, or there may be multiple mechanisms interacting in complex ways, making their isolation and relative assessment by using a control intervention difficult (Ojanen, 1994; Vase et al., 2015; Cerritelli et al., 2016). Secondly, when designed for complex interventions, control interventions are unlikely to be truly 'inert' (Moerman, 2013). Indeed, what may be a control intervention component in one trial (e.g., touch in a spinal manipulation trial or education in a cognitive-behavioural intervention trial) may well be the investigational treatment in another (e.g., touch in a trial of therapeutic touch or education in a trial of pain neuroscience education) (Cerritelli et al., 2016). Thirdly, many PPS therapies rely on the active participation of patients and their interaction with providers. Consequently, provider blinding is often impossible and participant blinding difficult (Boutron et al., 2004), thereby increasing the risk of biased trial outcomes. At times, the argument that blinding is impossible is even used to justify the lack of efficacy trials (Williams et al., 2021) and unblinded trials are considered "*as the best possible [quality of evidence] for these interventions*" (p. 328, NICE, 2016). As will be seen below, the existence of some high-quality efficacy RCTs in all areas of PPS research invalidates such reasoning.

[2] Based on own consensus work and justified below, the terms 'sham' and 'placebo control' will be avoided in favour of more descriptive language or the phrase 'control interventions in efficacy or mechanistic trials' (See chapter 3 for detail and the glossary in Table 1).

[3] Trials of surgery, devices, needle-based interventions, and therapies employing meridian rationales (e.g., traditional acupuncture and reflexology) offer good opportunities for blinding and comprehensive guidance exists (Felson et al., 2016; Braithwaite et al., 2020a; Beard et al., 2020; Zhang et al., 2022). These interventions are not discussed as part of this project.

Where employed, the types of control interventions in PPS efficacy trials vary widely (Boutron et al., 2007; Cerritelli et al., 2016). Sometimes, procedures are highly similar to the respective test interventions; other trials employ dissimilar 'fake' devices or attention control procedures that provide participants with therapist attention or unrelated educational content. Understandably, such variance in methods and lack of quality standards exposes the field to criticism (Bueno and Harrison, 2019; Furukawa et al., 2014; Pagoto et al., 2013; Hohenschurz-Schmidt et al., 2021, 2022a) and leaves questions of treatment mechanism unanswered. In particular, non-matching control interventions do not reliably distinguish specific treatment effects from context-dependent effects and may undermine blinding – thus potentially introducing bias (Boutron et al., 2007; Vase et al., 2015a; Puhl et al., 2017); Alternatively employed comparative effectiveness designs typically address different research questions than efficacy and mechanistic trials (Ford and Norrie, 2016; Zwarenstein et al., 2017; Dal-Ré et al., 2020). What constituted a high-quality control intervention in mechanistic or efficacy PPS trials was unclear when this project began. Also, the impact of design choices on trial outcomes was largely speculative and not backed up by empirical evidence (e.g., from meta-analyses).

Repeatedly, recommendations had been put forward to guide control intervention design (Baskin et al., 2003; Hart and Bagiella, 2012; Aycock et al., 2018; Braithwaite et al., 2020b; D'Alessandro et al., 2021), but these were therapy-specific, rarely based on empirical evidence or expert consensus, disregarded problems of intervention complexity or translation of methods from drug research, and ultimately failed to change research practice (Cerritelli et al., 2016; Braithwaite et al., 2019; Hohenschurz-Schmidt et al., 2022b). Notably, Baskin et al. (2003) introduced the concept of 'structural equivalence' to enhance matching between control and experimental treatments. Other authors also presented a range of features for which conditions should be similar or even 'indistinguishable', from the number of treatments, to procedural steps in the application of interventions, to the personal interactions with therapists and staff (Hart and Bagiella, 2012; Felson et al., 2016; Puhl et al., 2017; Aycock et al., 2018; Braithwaite et al., 2020a, 2020b). Recently, reporting guidelines for sham interventions were published,

encompassing many of these features (Howick et al., 2020). Not specific to nonpharmacological intervention research, it is questionable whether the Template for Intervention Description and Replication for placebo and sham controls (TIDieR-Placebo) is sufficiently comprehensive to reflect the challenges in efficacy trials of physical, psychological, and self-management interventions. Also, TIDieR-Placebo refers exclusively to the reporting, not the design or conduct of control interventions.

In sum, the field of PPS required an evidence-based, expert-informed framework for the design, conduct, and reporting of control interventions in efficacy and mechanistic trials. Such a framework had to provide unifying principles across PPS interventions, including theoretical, practical, and ethical considerations, while acknowledging the fundamental differences to drug and other nonpharmacological research areas.

To achieve this objective, it was first necessary to obtain a comprehensive overview of currently employed methods. This was provided by a systematic review of methods of published trials. At the same time, the hypothesised connected between control intervention design and trial outcomes was tested through meta-regression. Based on this analysis, a consensus process with experts in respective trials and in the science of the placebo effect was conducted, along with individuals experiencing persistent pain. As will be seen, there was consensus on numerous important aspects of control interventions in efficacy trials. It is hoped that the resulting guidance is widely applicable and will promote higher-quality efficacy trials in PPS.

Chapter 2: Systematic Review of (Sham) Control Intervention Methods and Meta-Analysis

In this second chapter, the research methods of the systematic review for this project part will be described, before presenting the guideline development process in the third chapter. Throughout the methods and discussion sections, the focus will be on justifying design choices and reflecting on alternative approaches. The detailed results of the systematic review are published (Hohenschurz-Schmidt et al., 2023b, 2023c), and main results are presented here in paraphrased form.

Methods

Systematic reviews are "*reviews of existing research using explicit, accountable rigorous research methods*" (p. 2, Gough et al., 2017). To become familiar with this methodology and able to train the project's reviewers in key methodological steps, the thesis author attended a three-day course at the Centre for Reviews and Dissemination (CRD) at the University of York. In developing the review's protocol, the CRD's guidance (Akers et al., 2009), the Cochrane Handbook (Higgins et al., 2019), and dedicated meta-analysis literature (e.g., Borenstein et al., 2021) were consulted.

Review objectives

A systematic review of methods was conducted and is reported according to the 2020 Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) statement (Page et al., 2021). The review's objectives were to identify common and less common control intervention designs in efficacy or mechanistic RCTs of PPS for a clinical population of patients with pain, commonly called 'sham', 'placebo' or 'attention controls'. Further, the review aimed to provide a detailed similarity assessment across 25 features for which matching between

control and experimental treatments has been argued to be important (see below), allowing for comparisons between therapy types. Also, we aimed to identify studies that reported on blinding effectiveness and control intervention validation studies. Finally, to provide quantitative data for the guideline development process, the impact of these control methods on trial results was formally examined.

Previous reviews had focussed on specific fields, such as acupuncture, physiotherapy, or manual therapy (Braithwaite et al., 2018, 2019; Webster et al., 2019; D'Alessandro et al., 2022), limiting the possibility to learn from approaches to shared challenges across PPS trials. No previous review had ever examined the hypothesis that similarity between control and tested interventions influenced trial outcomes such as pain reports, attrition, or blinding success.

The review was published as two separate manuscripts in the journal PAIN. The first article described control interventions in efficacy trials, their similarity to the tested PPS interventions, and other aspects relevant to their development, implementation, and interpretation (Hohenschurz-Schmidt et al., 2023b). The second paper presented meta-analyses of intervention effects compared to changes in control groups, and meta-regression analyses examining whether trial outcomes could be predicted by the (dis)similarity between control and index interventions (Hohenschurz-Schmidt et al., 2023c).

Protocol and registration

The protocol was pre-registered with the International Prospective Register of Systematic Reviews (PROSPERO, registration ID: CRD42020206590).

The reviewer crowd

Crowd science is research where some scientific tasks are performed by volunteers who are not professional scientists (Scheliga et al., 2018). The method has been applied to systematic reviews, which contain several 'microtasks' that can be completed by non-professionals (Thomas

et al., 2017). Based on specific training, crowd-based reviews enrol the help of a dedicated group of reviewers working together online (Noel-Storr, 2019). Next to time savings, crowds can achieve high sensitivity and specificity during citation screening (Noel-Storr et al., 2021).

When planning the review, it became clear that the number and complexity of efficacy trials of PPS interventions for pain would make the review very labour-intensive. The expected amount of eligibility screening and data extraction would only be feasible with a crowd-based approach. Here, the employed crowd methods are reported alongside crowd-related procedural outcomes.

Informed by experiences from other members of the research group (Soliman et al., 2019) and available literature (Scheliga et al., 2018), the thesis author set out to design a framework for such a collaboration: Crowd-reviewers worked as volunteers in return for training, research experience, and co-authorship. No further incentives were provided. Formally, the criteria for volunteers as per recruitment advert were:

- A biomedical or therapeutic background, at least at completed BSc level
- A vague understanding of systematic reviews and the willingness to learn more
- The ability to commit to the project both in terms of time and scientific rigour

Crowd-members were recruited through social media and personal networks (including the UCO research department, Prof Lene Vase's network at Aarhus University, and alumni mailing lists of an MSc in Neuroscience and a Master in Osteopathy). During the kick-off meeting, the following 'working contract' was agreed between the thesis author and the volunteers:

- Requirements for co-authorship: average contribution of at least 3 hrs per week for the entire course of screening and data extraction (allowing for holidays)
- DHS (the thesis author) reserves the right to end the collaboration at any point (primarily if quality of work does not suffice). Concerns will be flagged up in writing at least twice beforehand.
- Co-authors must review and approve manuscript prior to publication

- Other contributions will be acknowledged in any publication's 'Acknowledgments' section
- The professional development of all reviewers is a declared objective of this collaboration

Once enrolled, training encompassed an introduction to systematic review methods and the present project, a tutorial on bias in clinical trials, and in-depth practice of reference screening and data extraction on the employed web-platform (https://www.covidence.org/). Further, the author hosted a participatory workshop on placebo effects and control interventions, and optional drop-in sessions twice a week. A reviewer handbook was available to ensure homogenous screening and data extraction, which was constantly updated as needs for clarification were identified. Microsoft Teams and Google Drive were used for workshops and virtual collaborative working. Volunteers were free to discontinue the project at any time. However, completion of the training phase was necessary to enter the official project, and a minimum contribution for co-authorship. Above all, these criteria ascertained the quality of the volunteers' work, further ensured by regular process-checking through the thesis author and pairwise performance of all review steps (Gartlehner et al., 2020).

Throughout, I was very conscious of the moral and practical tensions between requiring committed and capable volunteers and my inability to reimburse them for their work.[4] My recruitment therefore focused on three groups of people: recent MSc graduates or current medical students who were aspiring to a research career, hoping they would obtain tangible benefits from co-authorship and additional research experience; clinicians who wanted to conduct research but did have other income; and professional researchers who had otherwise salaried positions but were available and motivated to participate. Importantly, I provided training and checked in with volunteers to ensure that their participation was indeed worthwhile. This included making their participation fun – through a team-building approach, lively use of a chat space in Microsoft Teams (Appendix 1: Systematic review and crowd-related content'), regular updates and personal emails, and as many synchronous meetings as desired during which I was available for discussions beyond the present project.

[4] For ease of reading, the first-person singular is used for reflective sections.

Thirty-nine volunteers commenced and 26 completed the training. Twenty-five began the official project stage (Crowd overview in Table 2) of which eighteen persisted until the end of the project to now be co-authors on the resulting publications. Compared to other crowd-based reviews, this was a very high retention rate (46% vs. 6.2% in Soliman et al., 2021).

Table 2: Characteristics of crowd members who completed the training and made a significant contribution to the project (n = 25; one person completed the training but did not commence the official project).

Crowd feature	**N (total n = 25)**
Highest academic degree	PhD (1), M.Sc (18), M.Ost, integrated Master in osteopathy (4), BSc (2)
Academic or professional background (multiple options possible)	Osteopathy (11), Neuroscience (10), Medicine (3), Psychology (2), Sports science (2), Physiotherapy (1), Biochemistry (1), Pharmacology (1)
Country of residence (during project)	UK (17), DK (2), FR (2), AUS (1), GER (1), A (1), B (1)
Previous systematic review experience	None (13), Some partial experience (8), conducted one full review (2), conducted and/or participated in several reviews (2)
Age group	<30 (17), 30-39 (5), 40-49 (2), >50 (1)

In the 5-month period from July 2020 to early December 2020, the crowd provided a total of 1,045 labour hours (20 hours of title/abstract screening, 60 hours of full-text screening, and 965 hours of data extraction, as per individuals' weekly time logs and not including the thesis author's contribution). This is equivalent to 26 full-time working weeks. If paid, the associated cost would have been between £17,828 and £30,044 (using Imperial College hourly rates for low and high-preparation student demonstrator roles; £17.06 and £28.75 per hour, respectively).

Where crowd members disengaged and informed the first author of their reasons, this was usually due to a lack of time or personal circumstances. Others dropped out during the training phase without giving a reason. On reflection, the training phase also gave potential

volunteers an opportunity to familiarise themselves with the nature of the work and the anticipated workload. For everyone concerned, disengagement at this early stage was preferable to later dropouts. Overall, however, feedback from volunteers was overwhelmingly positive (See volunteer feedback in Appendix 1: Systematic review and crowd-related content).

To summarily reflect on a crowd-based approach to systematic reviews, I believe that reimbursement should be the norm for any kind of labour and that crowd-science risks exploitation, not least through academia's incentive structures that promote self-exploitation (Brienza, 2016). Despite concerns about complicity in this system, the project's results indicate that a mutually beneficial situation was generated, where volunteers' ambitions were met, and a valuable project was made feasible. More specifically, the crowd-based approach provided an opportunity for early-career researchers to develop their skillset and for healthcare professionals to develop a deeper understanding of research and critical appraisal: Crowd science "*as a form of communication and knowledge exchange between volunteers and scientists*" (p. 516, Scheliga et al., 2018, referring to Bonney et al., 2016). Personally, I also enjoyed the interactions with a large and diverse team and acquired important leadership skills for subsequent project steps and my future career.

Eligibility criteria for reviewed studies

Returning to the systematic review itself, it included RCTs of PPS interventions for adults living with pain, irrespective of gender, underlying pathology or pain severity and duration (Table 3). At least one pain-related primary outcome measure had to be reported.

It was challenging to decide which interventions were appropriate to include. The umbrella 'non-pharmacological, non-surgical' (used in the original PhD project proposal) was soon found to be inadequate, as the terms were too imprecise, and the negative definition was

deemed undesirable[5]. Also, Braithwaite et al. (2019) had just published their systematic review of sham control methods in dry needling trials, making apparent the different opportunities and challenges compared to, for example, manual therapy, exercise-based interventions, or psychotherapy. As device-based interventions, acupuncture solutions focused on technological 'tricks', such as blunt or retracting needles, that did not seem transferable to other nonpharmacological areas. Iterative test searches, pilot eligibility screening, and extensive reading further demonstrated that meridian-based interventions (both acupuncture and acupressure) could be excluded as these trials could test mechanisms through avoidance of 'real' acupuncture points (Vase et al., 2015a). It was further found that device trials had recent and comprehensive guidance (Felson et al., 2016), leading the thesis author to exclude these interventions but informing the approach to data extraction with many potentially relevant items. In the end, the thesis author developed the PPS umbrella term, which also avoided the negative phrasing of 'non-pharmacological, non-surgical', and defined it as follows: As described in the glossary (Table 1) and used as inclusion criterion for the review (Table 3), PPS interventions include all forms of manual and physical therapy, exercise and rehabilitation therapy, conversation-based and psychological therapies, body-mind, spiritual, religious, other non-material healing practices, web-based therapies, relaxation, and educational interventions (the latter two were classified as 'self-management'). The review excluded studies in which pharmacological or drug interventions formed the mainstay of treatment and studies of surgical or otherwise invasive interventions. Further, all therapies relying on the permanent introduction of some form of matter into the body were excluded. Device and needle-based therapies were not eligible. Implanted and externally applied devices, all acupuncture modalities, and therapies based on assumed reflex points or energy meridians were excluded.

[5] This is aligned with attempts in the field to define unconventional medical approaches by 'What is' rather than 'What is not' – for example preferring the term 'integrative' over 'complementary' or 'alternative' therapies (Ng et al., 2016). Although likely a simplification, Ng et al. (2016) also ascribe different political agendas to authors using either terminology. 'What is' authors, they state, declared their objective to be the reform of the standard medical model, integrating the studied therapies if found to be effective; Conversely, 'What is not' authors were "*often advocat[ing] for greater scientific research of unconventional therapies with a stated goal of determining the safety and efficacy, among other aspects, of such therapies*" (p. 15). Ng et al. do not explore that these objectives may not be mutually exclusive.

Table 3: Eligibility criteria for inclusion into the systematic review.

Population	Interventions	Comparator	Outcomes	Design	Timeframe
Included:					
Any pain Adults	Physical Therapies Psychological interventions Self-Management	Placebo - Sham - Attention controls	Pain-related primary	RCTs	January 2008 – 24 November 2021
Excluded:					
Experimental pain	Surgery Devices Acupuncture Meridian therapy	If sole comparator: Other designated intervention groups (comparative effectiveness) Waitlist / no-treatment Treatment as usual		Pilot and feasibility RCTs (as defined by the primary study authors)	

To be eligible, trials had to employ a 'sham' control intervention. Classification depended on authors' self-report, where authors either used one of the common terms ('sham', 'attention' or 'placebo' control) or stated that the control intervention was designed to control for elements of the tested intervention.

Excluded were non-randomised studies, observational studies, cross-sectional studies, case-control, case-series, and case-report studies. Pilot or feasibility RCTs were excluded, except for validation studies that specifically evaluated aspects of the control interventions in an adult

population of patients with pain, irrespective of employed outcome measures. This decision was taken to omit trials that were not designed with the objective to test efficacy, but still capture validation studies as potentially more rigorous evaluations of blinding methods. Concerned about introducing bias from small studies (Dechartres et al., 2013), trials with less than 20 participants per arm were originally excluded. Before submitting the review for publication, however, the review was presented to the panellists of the Delphi study described in the next chapter. Especially Dr Heike Gerger, a meta-researcher from Erasmus MC (Rotterdam), provided extensive methodological feedback, which led to a re-analysis and inclusion of smaller trials for sensitivity analyses (confirming findings from primary analyses but increasing the review's comprehensiveness).

For included studies, trial protocols were consulted where available and required for additional method information.

The first reporting guideline for nonpharmacological therapy trials was published in February 2008 (Boutron et al., 2008). Therefore, this review systematically assessed studies published from 2008 onwards.

Data sources

The following databases were searched from 2008 to November 2021 (initial search conducted 23 June 2020, then updated; latest search: 24 November 2021): MEDLINE®, EMBASE, PsychInfo, the Cochrane Database of Systematic Reviews, the Cochrane Central Register of Controlled Trials (CENTRAL), NIH Clinicaltrials.gov, AMED (Allied and Complementary Medicine), CINAHL (nursing and allied health), the Physiotherapy evidence database (pedro.org.au), ostmed.dr (ostmed-dr.oclc.org), osteopathic research web (osteopathic-research.com), and the index to chiropractic literature (chiroindex.org).

Search strategy

Aligning with the PICOS framework (Higgins et al., 2019), this review's target Population, Interventions, Control conditions, Outcomes, and Study designs are encompassed by the following keywords, around which the search strategy was built:

(Pain OR painful conditions) AND Physical, Psychological, Self-management therapies (incl. specific therapy and technique names) AND (placebo control OR sham control OR attention control) AND controlled clinical trials. Limit: 2008-present.

The search strategy was developed based on existing literature and with database experts. It is provided in full for each database in the supplemental material of the final publications (Hohenschurz-Schmidt et al., 2023b, 2023c), and so is a spreadsheet of search results per database. The search strategy was developed for MEDLINE® first, using PubMed and Ovid interfaces, and then adapted to other databases. Development of the search strategy was paralleled by refinement of the inclusion criteria: Once decided which broad subject headings each of the PICOS categories should contain, own search strings were developed or searches of previous reviews were adapted (e.g., the intervention-related search strings of a review on the 'non-drug, non-surgical treatment of shoulder conditions' [Hawk et al., 2017]), and Boutron et al., [2007] and Amatya et al., [2018]). Vice versa, the reading of existing literature and insights from the search string development often led to clarifications in the review protocol and the PICOS framework. Refinement of the search string was iterative, adding or removing individual terms (free text or index terms) and assessing how this influenced the number and quality of retrieved references. If the number of additional hits was large, specificity of the terms was evaluated, mainly by visual inspection of retrieved abstracts (Higgins et al., 2019, chapter 4.4). Search terms were modified, for example by moving up or down the Medical Subject Heading (MeSH) 'tree' (https://meshb.nlm.nih.gov/treeView), exploring synonyms, or considering terms not captured by index terms. Finally, combined search strategies were trialled through test searches and screening of a subset of abstracts. Here, the author examined whether the draft

search returned relevant studies which had been identified through manual searches before, and omitted irrelevant study designs. Ambivalent cases which might require refinement of eligibility criteria were screened for; and the number of potentially included studies counted, estimating the total number of included studies. During test-screening, the thesis author also counted certain features of retrieved studies to determine the specificity of the search (for example, the number of pain-related studies, experimental pain designs, drug trials, surgical trials, RCTs, etc.). Subject experts were consulted regularly for therapy- or topic-specific evaluation of the search (usually members of the supervisory team); as was the local librarian, who mainly checked the technical correctness and logic of the string and helped with adaptation to additional databases and search interfaces.

Study selection

Eligibility screening was performed on Covidence.org. Two independent reviewers drawn from the pool of specifically trained research contributors (the 'crowd') independently screened each study. Compared to single-reviewer screening, duplicate screening reduces the risk of missing relevant studies considerably (Gartlehner et al., 2020). Disagreements were resolved by a third reviewer. The screening was first performed based on study title and abstract. Full-text eligibility was assessed in a second step.

Data extraction

The data extraction process also required a minimum of two independent reviewers. Discrepancies were resolved through discussion or, if not possible, by consulting a third independent reviewer. Here, a bespoke Google Form (a questionnaire tool) was used for data extraction, and Google Sheets for reconciliation.

Publications reporting multiple relevant control interventions were extracted independently for each pair of index intervention and control, with data from an intervention

arm used twice for comparisons with control interventions if required (sample size of the respective group was halved to account for duplicate inclusion). Where a single control group acted as comparator for multiple tested interventions, data were extracted from the test intervention that most resembled the control intervention.

The data extraction form followed from the respective research questions, thus containing the following domains: bibliographic data, general study design, trial reporting, control and blinding methods, trial results (pain-related outcomes but also attrition and blinding effectiveness), and risk of bias (using the Cochrane Risk of Bias tool 2, RoB 2) (Sterne et al., 2019). It was not deemed necessary to assess studies according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system, since no recommendations regarding intervention efficacy were made (Guyatt et al., 2011).

During data extraction, resemblance between control and experimental treatments was rated for 25 features. Pain-related outcome measures were extracted for sham control and active intervention groups at baseline and earliest follow-up after treatment. Pain-related outcomes can be unidimensional (usually pain intensity rating scales) or multidimensional (e.g., questionnaires assessing disease-related function, pain interference or quality of life) (Turk et al., 2003). In the absence of research on whether these are differentially susceptible to placebo effects, we strove to extract both a unidimensional and a multidimensional outcome measure per study. Specifically, data were sampled for the declared primary pain-related outcome, irrespective of its nature. Where available, another pain-related outcome of the opposite dimensionality was extracted. Where authors did not declare a primary outcome measure, data for a unidimensional and a multidimensional measure were extracted if available. Direction of effect was considered. Where necessary, data were extracted from figures using the Adobe Reader measurement tool. Authors were contacted via email if data were missing that were required for the calculation of effect sizes or if data appeared erroneous. For cross-over designs, only results from the first phase were extracted.

Data extraction was trialled using a sample of potentially eligible studies. Results of the pilot testing informed the final approach to data extraction, with detailed annotations for extraction items available to reviewers, and reliability monitored throughout (Higgins et al., 2019).

Data analysis

The review included a predominantly descriptive part, providing an overview of blinding methods used in the field of PPS therapies for pain, with a detailed description of control interventions, their development and reported rationale, their similarity to respective tested interventions, and compliance with relevant reporting guidelines (Boutron et al., 2017). The second part of the analysis consisted of meta-analyses of treatment effects compared to control interventions, differential attrition between groups, and measures of blinding effectiveness. Further, meta-regression analyses investigated potential links between blinding methods and these outcomes. Where appropriate, data were sub-grouped by therapy type. These methods are further detailed below.

Similarity index and ratings

A high degree of similarity between control and test intervention is commonly assumed to be a desirable feature of controlled efficacy and mechanistic trial designs (Turner et al., 1994; Baskin et al., 2003; Machado et al., 2008; Braithwaite et al., 2020a; Felson et al., 2016; Hart and Bagiella, 2012; Puhl et al., 2017; Aycock et al., 2018). Whilst some authors have used concepts of 'indistinguishability' and 'structural equivalence' to denote different levels of similarity (Baskin et al., 2003; Machado et al., 2008; Puhl et al., 2017), this project drew on such work to define 25 features across which control and treatment interventions may be compared. Assessed features are listed in Table 4 and were based on a review of the following pertinent literature: (Turner et al., 1994; Baskin et al., 2003; Safer and Hugo, 2006; Tough et al., 2009; Campbell et al., 2011; Hart and Bagiella, 2012; Hróbjartsson et al., 2014; Chaibi et al., 2015; Felson et al., 2016; Boutron et

al., 2017; Puhl et al., 2017; Aycock et al., 2018; Braithwaite et al., 2020b, 2020a). Similarity ratings were based on two reviewers' evaluation of how similar tested and control interventions were for each of the 25 items. For their ratings, reviewers were asked to rely on reported information as much as possible and avoid guessing. They were also able to rate interventions as 'partly' similar if this was the case but not to a large extent. Descriptive evaluations were converted into numerical values to allow for meta-analysis. Specifically, 'Yes' (similar) and 'No' (dissimilar) evaluations were rated as 2 and -2, respectively. 'Probably Yes' and 'Probably No' were awarded 1 and -1 points, 'Partly' was awarded 1 point, and 0 points were given for each item that could not be rated due to insufficient information. Non-applicable items were not rated.

Table 4: Features for which the level of similarity between control and index interventions was rated during the data extraction process. Items were only rated if relevant information was reported or could be inferred from trial reports with relative certainty.

Similarity items	Comments and explanations
Number of treatments / treatment sessions	
Treatment duration	*Duration of individual treatment sessions.*
Treatment frequency	
Assessments and re-assessments	*This included baseline testing and assessments as well as assessments performed at each appointment.*
Co-interventions	*Which other interventions were offered or permissible to participants in respective groups.*
Intervention standardization / level of protocol flexibility	*The degree of freedom the therapist and patient had in applying or complying with the therapy.*
Fidelity monitoring	*Monitoring whether the treatment protocol was followed, e.g., through video-taping individual sessions or asking patients / therapists about the treatment.*
Level of individualisation to the patient	*The degree to which treatment or control interventions were adapted to patient characteristics and / or preferences.*
Application modes	*The means of delivering interventions, such as by hand (manual therapies), device (e.g., sham ultrasound), talking, video, etc.*
Delivery formats	*Referring to group vs. individual sessions*
Information provided to participants as to the efficacy of the intervention they receive	*E.g., if authors tried to create a certain expectancy by (mis)informing patients about treatment efficacy. Reviewers answered 'Yes' if the patients were told that all treatments are efficacious or that only active treatments are tested.*

Thematic content of sessions	*What patients and providers talked about.*
Body area(s) addressed	*Included criteria for which body parts were addressed, e.g., symptomatic vs. asymptomatic. Not applicable for cognitive therapies.*
Physical procedures performed or undergone	*In the case of talking therapies, this may include simply sitting and talking.*
Cognitive function targeted or attention focus	*Refers to what providers tried to change in the patient's mind and / or the patients' attention focus during interventions. This includes focusing on the same body area because the therapist is touching the same areas in both groups.*
Level of patient participation (physical or cognitive)	*The level of active engagement and participation required from participants during interventions, on a spectrum from completely passive to fully self-directed.*
Types of procedural steps per treatment session	*Examples of 'procedural steps' into which interventions can be deconstructed, are: verbal case history taking, assessment and diagnostic tests, single or multiple therapy procedures, and appointment closure.*
Number of procedural steps per treatment session	*As for previous question, judging the number of such steps only, not the content.*
Equipment / devices employed	*For example, if exercise groups employed Thera bands, did the control group do so, too? Other devices are smartphones, PC, books, leaflets, ultrasound, needles etc.* *Not applicable if no devices employed in either group.*
Treatment environment (locations, clinical setting)	*Rated as similar if all groups attended the same treatment centre and room. For remotely delivered interventions (online, phone etc.), this was rated as 'similar' if all patients were in their respective home environments.*
Treatment-specific sensory cues	*Whether the interventions involved physical sensations that were comparable between groups. Reviewers were asked to consider e.g., sound levels, touch, temperature, vibration, movement, pain sensations, and smells.*
Other therapy-specific considerations, incl. replication of side effects	*This relatively open question invited reviewers to think critically about whether there was anything special about the therapy and whether the authors had taken that into consideration when designing the control intervention. Reviewers were asked to provide detail if choosing to answer this question.*
Personal interactions with therapy providers (i.e., excl. admin, research, and other medical staff)	*The extent to which participants met providers in each group.*

Personal interactions with personnel other than therapists (i.e., admin, other medical, and research staff)	*Judged to be likely similar if all participants attended the same clinics and if trial personnel were blinded. Otherwise, authors had to report attempts to standardise such interactions.*
Follow-up contact frequency and mode	*Procedures for outcome assessments.*
Provider-related similarity (analysed separately)	
Same providers	*Whether therapy and control interventions were provided by the same therapists.*
Provider expertise	*Expertise referred to educational levels and professional qualifications.*
Provider experience	*Years in practice.*
Trial-specific training	*Any training providers received prior to the trial on how to deliver the treatment and / or control interventions.*
Provider behaviour	*Only if reported to be deliberately matched between control and corresponding intervention groups, this was also relevant if the same providers applied both interventions.*

Apart from item-specific ratings, each trial's total ratings were divided by the number of rated items to produce a single value, encompassing similarity across all applicable items. This was for illustrative purposes only as it is unclear whether all items can be weighted equally. Further challenges of and reflections on this assessment of similarity are presented in the below discussion section.

Possible values of the item-specific group averages and the overall similarity average ranged from -2 (dissimilar across all studies / rated items) to 2 (similar). Data for individual items and the overall index were synthesised as means and standard deviations for each therapy group. Apart from their descriptive presentation, these data were used as covariates for meta-regression modelling as described below.

Meta-analysis: placebo responses and treatment effects

Primary meta-analyses and meta-regression analyses included all studies with more than 20 participants per arm, with smaller studies added for sensitivity analyses. For primary meta-

analyses, we synthesised outcome data as per below, and presented them in forest plots alongside a study-specific a risk of bias (RoB) assessment (Sterne et al., 2019).

For each included control group, placebo responses were calculated as standardised mean changes (SMC) from baseline to first follow-up after the end of the treatment period, for both a uni- and a multidimensional pain-related outcome measure, where available. To calculate SMCs, a change score was divided by the pooled standard deviations. Compared to standardised mean differences (SMDs), SMCs take into consideration the correlation between baseline and follow-up assessments and were preferable as those measurements are not independent (Higgins et al., 2019, section 10.5.2).

Between-group differences (treatment effects) were calculated as the difference in SMCs between active and control groups for the same measures and timepoints. Meta-analyses of between-group differences were performed per therapy subgroup and for uni- and multidimensional measures, separately. For each group of therapies, summary effects were calculated using random effects models weighted by the inverse of the outcome's variance and plotted using RevMan 5 software (The Cochrane Collaboration, 2020). The heterogeneity of overall effects was estimated using Tau^2 (T^2) and I^2 statistics and tested for significance using Z statistics (Borenstein et al., 2021).

Meta-analysis: blinding success and blinding indices

Evaluation of blinding success provides the arguably most robust measure of control intervention quality and we meta-analysed blinding effectiveness in a subgroup of studies where sufficient data were reported. During data extraction, studies that indicated the effectiveness of the employed blinding methods were identified, for example studies that had patients guess their group allocation or rate the treatment credibility.

Ideally, blinding effectiveness is examined through allocation guessing, i.e., by asking participants to guess which trial group they believe they were allocated to. This in turn allows for

the calculation of blinding indices (BI). Such indices estimate the level of blinding success beyond chance level, and thus allow for a more nuanced assessment of blinding effectiveness than the report of correct/incorrect guesses per group. As input, such calculations require data of guessed group allocations per study group, either in a 2x2 (Possible guesses: treatment group / control group), 2x3 (Treatment group / control group / Don't know), or 2x5 format (Certainly treatment group / probably treatment group / probably control group / certainly control group / Don't know). Two indices are commonly available: James' BI and Bang's BI. After some consideration and reading (James et al., 1996; Bang et al., 2010; Braithwaite et al., 2018; Colagiuri et al., 2019), we opted to employ Bang's BI. First, this index was deemed less sensitive to 'Don't know' data than James' BI, which was not expected to be commonly reported and around which there is uncertainty whether 'don't know' signifies true blinding or is just a more socially acceptable answer than the disclosure of non-blinding (Bang et al., 2010). Further, Bang's BI enables comparisons of proportionate guessing between study arms (Braithwaite et al., 2018), not least when indices from each study arm are combined in a manner proposed by Colagiuri et al. (2019), using Hedge's g. Finally, Bang's BI is easier to interpret and classification rules for successful vs. unsuccessful blinding have been proposed (Bang et al., 2010; Moroz et al., 2013).

In short, where group guesses were reported in a manner that allowed for the calculation of Bang's BI, the index was calculated for active and control groups individually (Bang et al., 2004) and a ratio of Bang's BI calculated as Hedge's g for each comparison between test and control group (Colagiuri et al., 2019). The effect size of Hedge's g was estimated irrespectively of the interventions studied, using meta-analysis methods as above.

Meta-analysis: expectancy and treatment credibility measures

To allow for further discussion and sample currently employed methods of assessing control intervention quality, methods other than allocation guessing were extracted and reported descriptively where available. Mainly, this was expected to include measures of treatment credibility, expectancy, and satisfaction.

On the suggestion of a manuscript reviewer and not specified in the original protocol, these reports were examined for the possibility of data pooling and meta-analysis of expectancy measures. Knowing that the most commonly employed questionnaire by Devilly and Borkovec (2000) evaluated both expectancy and credibility aspects of interventions, the following criteria were specified for meta-analysis and regression with similarity ratings: Reported expectancy data had to be i) clearly attributable to expectation of treatment benefit only (i.e., not presented as a compound measure with questions on treatment credibility, or evaluated as treatment satisfaction), ii) sampled after at least one exposure to the test or control intervention but not after the final of multiple treatment sessions (to avoid confounding with treatment satisfaction), and iii) reported in full and per trial group.

Multiple meta-regression analysis: the role of similarity between intervention and sham controls in predicting trial outcomes or blinding effectiveness

Originally, simple correlation analyses had been planned to examine the relationship between similarity ratings and trial outcomes. Based on interim feedback from subject experts, we deviated from the pre-registered protocol to perform a more rigorous meta-regression analysis instead. Specifically, to assess the potential relationship between the trial results and how (dis)similar sham and active interventions were, for each individual therapy subgroup, meta-regression analyses were performed using methods-of-moments random effects models (Borenstein et al., 2021; Wilson, 2005) with studies weighted by the inverse of their variances in outcomes. An advantage of meta-regression over simple correlation analyses is that it allows for the analysis of multiple potential predicting variables simultaneously (Higgins et al., 2019, section 9.6.4), assuming a sufficient number of included studies (Thompson and Higgins, 2002). Meta-regression also allows for the appropriate weighting of studies according to the precision of their effect estimates.

The SMC between sham control and active intervention groups was used as the dependent variable and models were weighted by the inverse of its variance. Individual similarity

ratings were used as independent variables or potential 'predictors' of the SMCs' unexplained heterogeneity. Put simply, it was tested if similarity ratings could account for some of the differences in pain-related outcomes between trials. In particular, this method examined differences that were not likely due to the 'true' difference in treatment efficacy but due to other factors, for example control methods (i.e., the 'unexplained heterogeneity').

Models were built for each therapy subgroup individually, identifying potential predictive variables from the pool of similarity ratings. The process to decide whether similarity items should be included in regression models was as follows: First, non-parametric correlation analyses between all similarity ratings for each subgroup were performed, to identify a selection of variables with little interdependence between each other. Selected items were then used individually as predictors in a regression model, and those items with the largest individual predictive value were then combined. In doing so, it was attempted to maximise the overall predictive power of each model while keeping the number of included variables as small as possible. This iterative adjustment of the model resulted in models of supposed best fit.

Meta-regression modelling was also performed for a subgroup of studies for which Bang's blinding index could be calculated, testing if similarity features could predict variance in the studies' blinding success as well as testing if the blinding index could predict the variance in studies' effect sizes. These analyses were irrespective of the type of therapy tested in the trials. Similarly, two meta-regression models were computed to examine if differences in attrition between studies predicted trial outcomes and, further, whether the degree of similarity between active and sham interventions predicted the degree of differential attrition. The results of meta-regression models are presented per therapy group below. Primary analyses were performed with large studies only (20 or more participants per group) and sensitivity-tested using the complete dataset and/or excluding studies that did not formally qualify as outliers (i.e., were included in the meta-analysis and primary modelling), but whose confidence intervals did not overlap with those of the aggregate effect.

The initial plan was to use a trial-level average of similarity ratings for meta-regression analyses. However, the thesis author later decided to use ratings for individual items only, given concerns about the validity of a compound score. Notably, many of the individual items' ratings were inter-correlated so that an overall score would have been biased. Also, an equal weighting of all items as part of an average was seen as an undue assumption. In retrospect, this approach also enabled a more nuanced discussion than what would have been possible with a compound score.

Results

Crowd performance

As already noted above, the crowd consisted of twenty-five members who began the official project stage, of which eighteen persisted until the end of the project. They provided a total of 1,045 labour hours (26 full-time working weeks) over an approximately 5-month period. Inter-rater reliability was moderate for the title/abstract screening stage (Cohen's $\kappa = 0.58$) and substantial for full-text screening (Cohen's $\kappa = 0.7$), with percentage agreements of 95.2% and 88.6%, respectively. During data extraction, crowd reviewers extracted similar data for an average 85% of the extractable items (κ not calculable).

Flowchart and description of included studies

Figure 2 shows the study selection process. The final review included 194 publications (plus protocols where available), extracting data for 197 unique 'sham' control interventions and 198 comparisons between control and experimental interventions. Manual therapy trials were most common, and there were multiple psychological and rehabilitation trials. Mostly, patients with musculoskeletal pain were recruited. Included trials' characteristics are illustrated in Table 5. Of all included trials, 121 were included in the primary meta-analyses, with a further 49 smaller studies used for secondary sensitivity analyses.

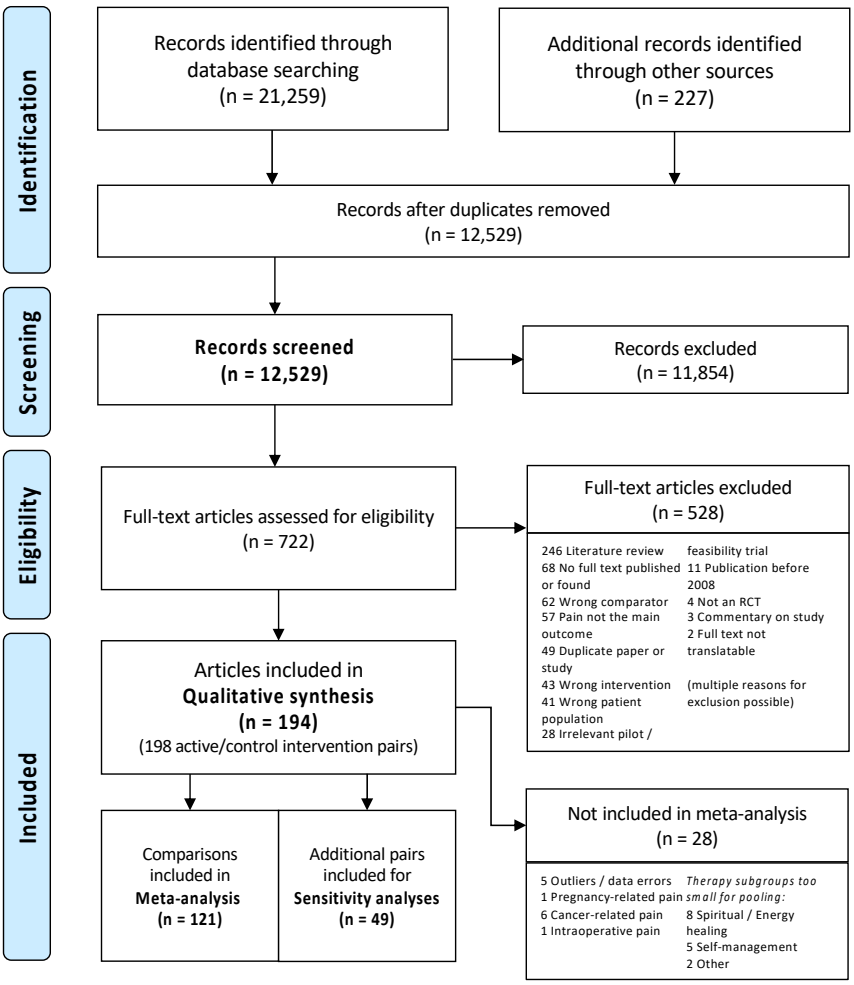


Figure 2: PRISMA flow diagram of the systematic search and selection process. A list of all studies excluded at the full-text screening stage is provided in a Supplementary Digital Content file of the resulting publications.

Table 5: Overview of included studies. The types of therapies, intervention complexity, and pain population are provided for the entire sample (N = 194 publications with 198 comparisons between treatment and a sham control). Special cases: In one trial, data from the active intervention group was used twice to compare it with two different sham controls: Bialosky et al. (2014) used a 'standard' and an 'enhanced' sham control. Three publications reported more than one trial: D'Souza et al. (2008) studied two groups with different types of headaches, and Assefi et al.'s (2008) publication included two active interventions and a matching sham control each. Finally, Sharpe et al. (2012) reported two trials in a single publication, which were treated entirely independently here. In general, only patients that informed the present analyses are counted in this table, patients were not counted twice, and analyses of reporting refer to individual trials.

Therapy types	**n of studies**	**%**
Manual therapy with spinal manipulation	48	24.2
Craniosacral therapy and gentle myofascial release	22	11.1
Other Manual Therapy	64	32.3
Rehabilitation / Physiotherapy	22	11.1
Self-management	5	2.5
Cognitive-behavioural and other psychotherapy	27	13.6
Spiritual / energetic / esoteric healing	8	4.0
Other	2	1.0
Intervention complexity [+]	**n**	**%**
Simple	112	56.6
Complex	86	43.4
Pain descriptor	**n**	**%**
Musculoskeletal pain	121	61.1
Diffuse chronic pain	18	9.1
Cancer-related pain	6	3.0
Visceral pain	5	2.5
Neuropathic pain	5	2.5
Pregnancy-related pain	1	0.5
Not specified	1	0.5
Sample size at randomization	**Median**	**Q1 / Q3**
Overall sample size (all trial arms combined)	64	40 / 101
Sample size per trial arm (only groups included in review)	27	19.5 / 46
Registered trial protocol available	**n**	**%**
Registered	114	57.9
Group design		
Parallel-group	163	92.9

Cross-over	14	7.1
Number of study conditions per trial *	**n**	**%**
2	139	71.6
3	46	23.7
4	8	4.1
5	1	0.5
6	1	0.5
Additional non-sham comparators included	**n**	**%**
Active comparator (comparative effectiveness) ±	29	14.9
No treatment / waitlist	20	10.3
Usual care / treatment-as-usual	13	6.7

* Each intervention or sham intervention was counted, irrespective of whether the trial was a single-arm cross-over trial. ± So-called attention controls were not counted as active comparator, only experimental conditions that were clearly assessed because they were deemed potentially effective alternatives (comparative effectiveness intention). + Intervention complexity: Single-step or single-technique interventions were judged as 'simple', irrespective of how often these were applied, and others as complex.

Also included were eight validation studies of sham control interventions tested in patients with pain (Hall et al., 2008; Kawchuk et al., 2009; Curtis et al., 2011; Vernon et al., 2012; Brose et al., 2013; Michener et al., 2015; Chaibi et al., 2015; González et al., 2021). Two (Vernon et al., 2012; González et al., 2021) of these studies were eligible for primary meta-analysis of pain-related outcomes. Three validation studies with more than 20 participants per group (Chaibi et al., 2015; Curtis et al., 2011; Michener et al., 2015) did not provide pain-related outcome data and were thus not included into respective meta-analyses. They were, however, included into the meta-analysis of trials reporting on blinding success.

Placebo and sham control intervention designs

The CONSORT statement (CONsolidated Standards of Reporting Trials) asks researchers to describe "*[t]he interventions for each group with sufficient details to allow replication, including how and when they were actually administered*." (Schulz et al., 2010; Boutron et al., 2017). In our sample of 198 sham control interventions, 67% complied with this reporting item

and provided a description of *the control intervention* whilst 77% did so for the experimental treatment. Compliance with the more relevant TIDieR-Placebo checklist (Howick et al., 2020) was not examined, as this tool had not yet been published when data extraction was performed.

Table 6 provides an overview of the main features of all reviewed control interventions, categorised by therapy type.

Table 6: Overview of employed control interventions per therapy type. N = 198.

Therapy type	Total N	'Placebo' / 'Sham' control interventions	n	%
Manual therapy with spinal manipulation	49	Manual, simulated manoeuvre	30	61.2
		Manual, soft touch	9	18.4
		Manual, technique to different area	1	2.0
		Disabled device	6	12.2
		Rest time-control	1	2.0
		Other (therapist attention, general anaesthesia)	2	4.1
Craniosacral therapy and gentle myofascial release	22	Disabled device	10	45.5
		Manual, soft touch	8	36.4
		Manual, simulated manoeuvre	2	9.1
		Other (low-strength static magnets, active joint movement)	2	9.1
Other manual therapy	63	Manual, simulated manoeuvre	21	33.3
		Manual, soft touch	19	30.2
		Disabled device	16	25.4
		Manual, technique to different area	1	1.6
		Manual, simulated manoeuvre & to different area	1	1.6
		Other (time-attention control, therapist attention, active joint movement, 2x low-pressure algometry)	5	7.9
Rehabilitation / Physiotherapy	22	Exercise, nonspecific	8	36.4
		Disabled device	5	22.7
		Educational attention control	5	22.7
		Exercise, key components altered	1	4.5
		Manual, simulated manoeuvre	1	4.5
		Manual, soft touch	1	4.5

		Other (nonspecific visualisation, rest + therapist attention)	2	9.1
Cognitive-behavioural and other psychotherapy	27	Multicomponent therapist interaction	11	40.7
		Educational attention control	7	25.9
		Cognitive task, nonspecific	2	7.4
		Writing attention control	2	7.4
		Other (relaxing music, nonspecific visualisation, open-label saline injection, nonhypnotic relaxation suggestions and white noise, nonspecific video and sound)	5	18.5
Spiritual / energetic / esoteric healing	8	Simulated hands-off manoeuvre (actor)	7	87.5
		Simulated hands-on manoeuvre (actor)	1	12.5
Self-management	5	Rest time-control	2	40.0
		Educational attention control	1	20.0
		Website, nonspecific	1	20.0
		Other (white noise)	1	20.0
Other	2	Other (white noise, headphones without sound)	2	100.0

Similarity between experimental and sham interventions

Conceptually, 29% of all studies explicitly reported matching or controlling for certain intervention components but the degree to which sham control interventions resembled the tested intervention varied widely.

The average similarity between experimental and sham intervention per trial was 0.88 (SD ± 0.66) across all rated features (on a scale of -2 to 2, with 2 indicating maximum similarity). Assessment of individual features showed that some items were frequently designed to match the active intervention, whilst this was rare for others (Figure 3, a table with additional statistical detail is provided in the supplemental digital content of the respective publication). For most items, however, confidence intervals were large. Overall ratings were different between simple and complex intervention trials ($t(1,195) = 4.67$, $p < 0.0001$), with comparisons between simple interventions and their shams being on average 0.4 points more similar (0.24 – 0.6 95% CI).

Important for the primary meta-analyses which included only large trials, average similarity ratings were not significantly different between large and small trials ($F(1,197) = 3.56$, $p = 0.061$).

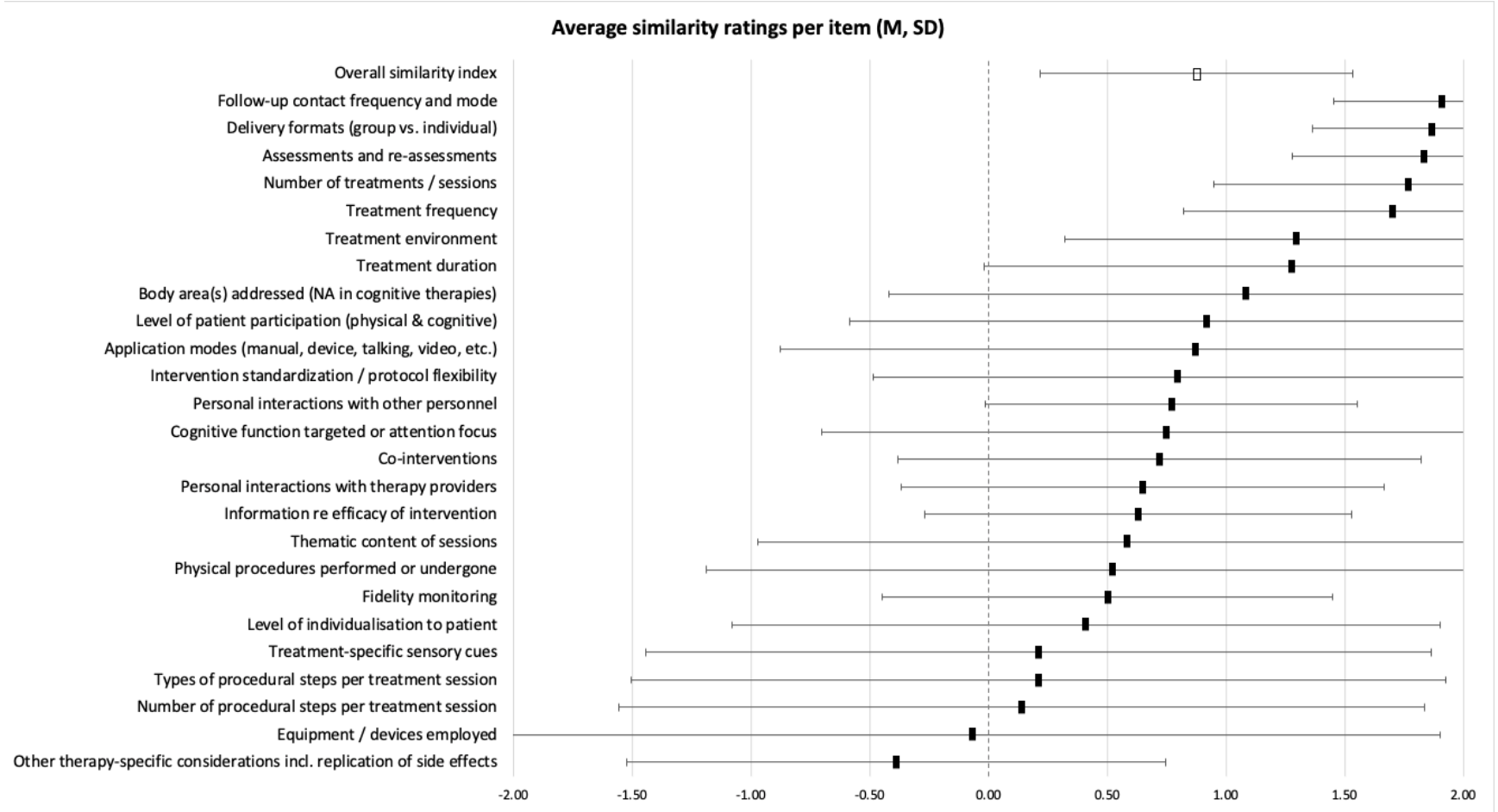


Figure 3: Similarity between tested intervention and placebo control intervention for all included studies. Items were rated if applicable to individual studies (N provided in Supplemental Digital Content 3), with the following possible ratings and corresponding numerical values: Yes = 2, Probably Yes = 1, Not reported = 0, Probably No = -1, No = -2. Squares indicate Means and Standard Deviations (SD) are provided as whiskers. The empty square represents the overall mean across all items and comparisons. N = 198.

Notable therapy-specific differences existed for overall similarity ratings ($F(7,190) = 5.28$, $p < 0.001$), with physiotherapy/rehabilitation trials having significantly lower average ratings (0.37 ± 0.77) than spinal manipulation trials (1.07 ± 0.54, $p < 0.001$), other manual therapies (excluding craniosacral therapy) (0.93 ± 0.6, $p = 0.008$), and trials of spiritual or energetic therapies (1.52 ± 0.42, $p < 0.001$). Figure 4 provides therapy-specific similarity ratings across all assessed features. A table with statistical detail is provided as supplement of the respective publication (Hohenschurz-Schmidt et al., 2023b).

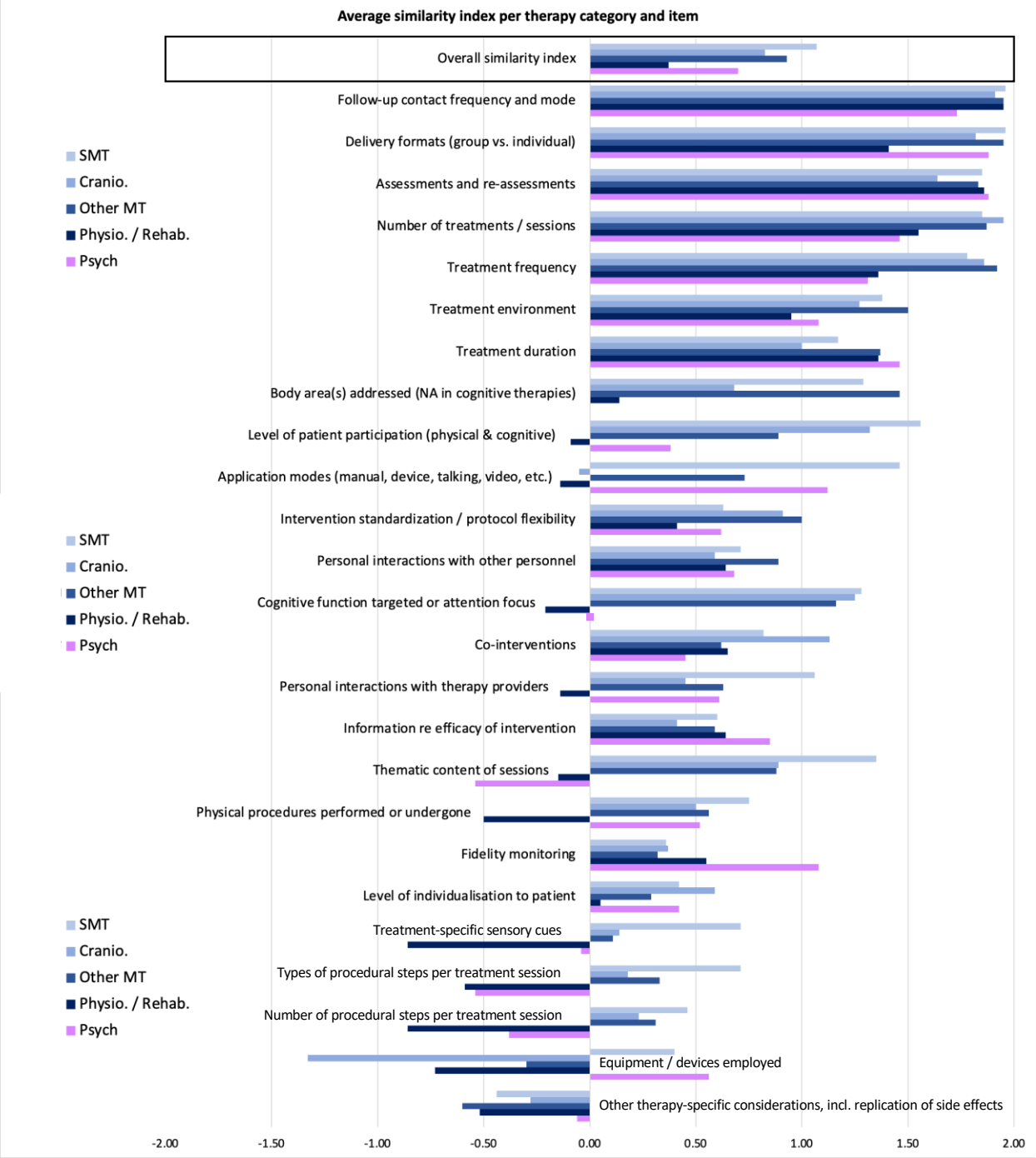


Figure 4: Average similarity ratings categorized by therapy type, comparing active and control interventions across 25 features. The overall mean across all items is provided in the top row. The end of each bar indicates the average rating. Measures of variability are provided in a supplementary table of the respective publication (Hohenschurz-Schmidt et al., 2022b) and the number of trials per group in the overview Table 5.

Interventions in the test and control groups were delivered by the same (set of) providers in at least 120 (59%) of all trials (clearly reported). Different providers were used in at least 32 (16%) trials, and this could not be ascertained or was not relevant due to treatment automation in a further 46 (23%).

In trials where it was clearly reported that different providers were employed, we further assessed if these were matched for expertise (e.g., educational background), experience (e.g., years in practice), behaviour, and if trial-specific training had been similar between groups (Table 7).

Table 7: Matching of provider-related factors. Assessed in trials for which it could be ascertained that different providers were employed for active and control interventions. Highest values highlighted by bold font. N = 32.

	Provider expertise		**Provider experience**		**Provider behaviour**		**Trial-specific provider training**	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Matched	7	21.9	3	9.4	12	37.5	9	28.1
Not matched	**18**	**56.3**	**16**	**50.0**	5	15.6	9	28.1
Not reported	7	21.9	13	40.6	**14**	**43.8**	**11**	**34.4**

Additional features of placebo control interventions

Reported methods of enhancing patients' expectations of a therapeutic effect included describing potential benefits of the sham intervention. Relatedly, some trials informed patients that only effective treatments were studied, either directly or by naming the control intervention differently (e.g., a "physical modality" (Lehtola et al., 2010)). Providers' positive expectations were modified, for example, by not informing them that the sham device was disabled or telling them that simple touch could have beneficial effects (Table 8).

Table 8: Additional features of control intervention design and implementation. Total n = 198, unless otherwise indicated. Note that the provided assessments are dependent on author reporting.

Feature	Reported in		
	n	%	
(Mis)Informing patients that only efficacious treatments are investigated	32	16.2	
Only enrolling treatment-naïve patients	30	15.2	
Deliberately enhancing patients' expectations of therapeutic benefit (other means)	16	8.1	
Providing specific training to the therapists to deliver the control intervention	43	22.8	Not applicable in 10 automated interventions (n = 189)
Enhancing control *provider* expectations of therapeutic benefit	9	4.8	Not applicable in 10 automated interventions (n = 189)

Control intervention development and theory

The review examined the reporting of processes and theoretical considerations underpinning the development of each control intervention. Where reported, information on development processes or theoretical considerations were brief, often no more than a half-sentence. Theoretical considerations included justifying the choice or omission of certain

elements of a control intervention. Overall, many studies provided no indication how the design of the control intervention was informed (Table 9).

Table 9: Reporting of control intervention development processes and theoretical considerations. Total n = 198, unless indicated otherwise.

'Sham' control intervention development	Reported in		Comment
	n	%	
Processes involved in the development of the sham control intervention	90	45.5	29 (14.6%) in the text and with some explanation 61 (30.8%) provided a reference to previous work
Specific testing of the control intervention (incl. feasibility or validation phase, either as part of the trial or externally)	22	11.6	Not applicable in 8 validation studies (n = 190)
Consulting with patient, practitioners, and / or public involvement groups	4	2.0	
Consideration of ethical aspects of respective sham intervention	14	7.1	
Theory	**n**	**%**	
Theoretical considerations in the design of the sham control intervention	59	29.8	
Mention of any hypothesized active (or specific) components of the treatment under investigation	160	80.8	
Mention of any hypothesized non-specific (or common) components of the treatment under investigation	58	32.6	
Mention of matching or controlling of the non-specific (or common) factors between therapy and sham	58	29.3	

Blinding

Assessing compliance with a relevant CONSORT reporting item, 53% of all included studies reported the blinding status of *all* involved stakeholders (patients, providers, outcome assessors and statisticians). An additional 36% reported the blinding status for some of the above. Trials were designed to blind patients to group allocation in 75% of cases, but information on patient blinding was not provided in 13% of reports, and 12% of sham-controlled RCTs reported that the trial was not designed to blind participants to the nature of the intervention received or the group allocated to. Whilst trial reports were often ambiguous on the specific circumstances, it appears that in these instances patients may have been aware of the group they had been allocated to, often because the two interventions were very dissimilar. They were, however, in most instances likely not aware that the control intervention was an intervention with no supposed effect on outcomes (Arcos-Carmona et al., 2011; Albert and Manniche, 2012; Buttagat et al., 2016; Castro-Sanchez et al., 2011b, 2011a; Lehtola et al., 2010; Schwerla et al., 2008; Kumar, 2011; Izgu et al., 2020; Andreae et al., 2020; Messier et al., 2021). Consequently, these trials were sham-controlled but employed deception as to the nature of the comparator intervention. In other instances (Allen et al., 2010; Ashar et al., 2021; Davis and Zautra, 2013; Garland et al., 2014; Ilgen et al., 2016; Khodneva et al., 2020; Schmidt et al., 2011; Vitiello et al., 2009; Wang et al., 2010, 2009), control interventions were employed that were not thought to be entirely inert but have circumstantial effects on outcomes. Almost exclusively, these latter were so-called attention controls for cognitive-behavioural interventions.

Providers were blinded in a minority of 3% of trials, but the methods to achieve double-blinding are noteworthy: Ajimsha et al. (2014) and Moraska et al. (2017) did not inform practitioners that the employed ultrasound machine was non-functional. In another case, the control intervention was provided by family members who read to the patients, essentially providing an attention control without knowing about its rationale in the trial context (Collinge et al., 2013). A similar strategy was applied to practitioners by Vitiello et al. (2013), with providers delivering an educational attention control not knowing that it was the trial's control condition.

In a further 7% of trials, provider blinding was of no concern, as, for example, automated or pre-recorded interventions were studied.

Blinded outcome assessment was reported for 58% of studies. A further 24% exclusively employed patient-reported outcome measures, thus ensuring blinded outcome reporting where patient blinding was successful. Unblinded outcome assessment was reported in 6% of trials and information on blinding status of outcome assessors was not reported in 11% of trials. The separation of treatment provider and outcome assessor roles was another common method to enhance internal trial validity, reported in 67% of trials (not performed in 6% and not reported in 28%).

Whether or not the statistical analysis was blinded was rarely reported (69% not reported), with 21% of trials reporting blinded, and 10% unblinded, statistical analysis.

Reports of blinding effectiveness and patient expectancy

Out of 198 control interventions, 150 (76%) were most likely designed to blind participants to the allocated intervention. Only in 35 (23%) of these cases did researchers evaluate whether participants blinding had been successful, which included all but one of the eight control intervention validation studies. Blinding was mainly assessed by patients guessing their group allocation, and occasionally via treatment credibility as proxy. The methods to analyse and interpret blinding success were highly variable.

In 18 reports, blinding indices were provided, or data were reported in a manner that allowed for calculating Bang's index. These studies were included for meta-analysis, presented below. In this subgroup, there were nine spinal manipulation, seven other manual therapy trials and one trial each of physiotherapy/rehabilitation and spiritual healing (Reiki). The average sample size at randomization was 64.4 patients (SD 37.4, range 10 to 154).

Only 4 studies reported unsuccessful participant blinding as per their own criteria; all others reported successful blinding or provided descriptive data without judgement. Trial-level detail is reported in a supplementary table of the respective publication (Hohenschurz-Schmidt et al., 2023b).

As per Colagiuri et al. (2019), the blinding indices of each trial's two groups were combined as a ratio using Hedge's g, with values larger than 0 indicating that participants in the active group were more likely to correctly guess their allocation to the active group than those in the control group, and values below 0 indicating that patients in the sham control group were more likely to wrongly guess that they received an active treatment compared to those in the active group. Having excluded one outlier where blinding was fully unsuccessful (Walker et al., 2013), the combined Hedge's g was 1.31 favouring participants in the experimental groups guessing correctly over control participants believing to have received active treatment (unsuccessful or unbalanced blinding; 0.2 to 2.43 95% CI, $p = 0.02$, $I^2 = 99.7\%$, participants = 1.013, studies = 17). Data are presented as a forest plot below (Figure 5).

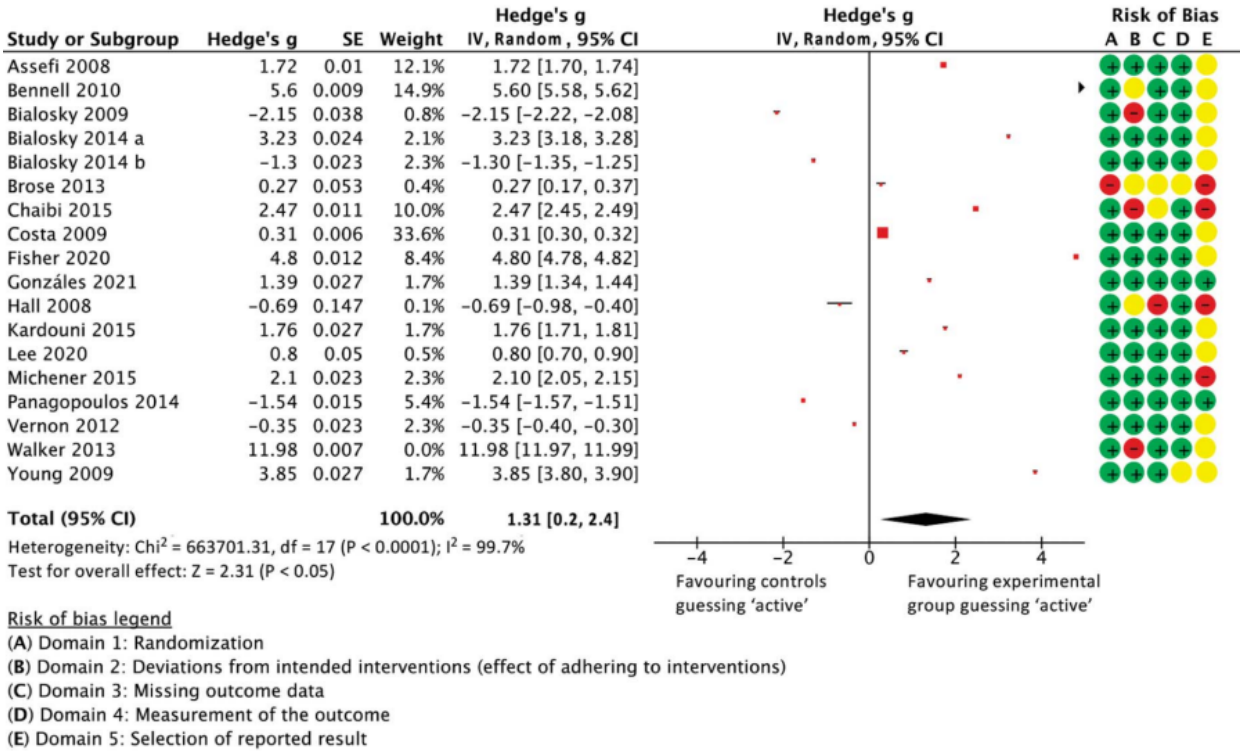


Figure 5: Forest plot showing all studies for which Bang's blinding index (BI) could be calculated (n = 18). The ratio between BI in active and control groups is presented as Hedge's g, with values near 0 indicating that participants in both groups were likely to make similar guesses as to which treatment they received (i.e., were adequately blinded). Values below 0 indicate that patients in the control intervention arm thought disproportionately more that they were in the active arm than those in the actual active group (indicating higher control intervention credibility), and values to the right indicating that the active treatment made more patients think that they had a real treatment than the control treatment in the same study (in other words, fewer control-arm patients thought that theirs was a real treatment than active-arm patients did about their intervention, indicating unsuccessful blinding). An extreme outlier (Walker et al., 2013) is shown in the forest plot but does not feed into the meta-analysis (weight = 0%).

Additionally, two cross-over studies assessed blinding effectiveness (Hall et al., 2008; Teys et al., 2008). In cross-over designs, patients can directly compare experimental and sham treatments, arguably making it easier to correctly guess group allocation. However, there was no indication of less successful blinding in the second phase of the Hall et al. trial (2008). Teys et al. (2008) indicated successful blinding but did not provide useful data for independent assessment.

The time points of blinding assessment differed, with most trials obtaining ratings after the first session or after the end of the treatment. Few studies monitored blinding throughout the course of a longer trial (except for Chaibi et al., 2017 & Lee and Lewis, 2020). Notably,

however, 22 other studies (11%) reported that their control intervention had been tested previously.

Apart from reporting on blinding success, 29 studies (14.6%) assessed the patients' expectation of treatment benefit, albeit in a very heterogeneous manner, or their satisfaction with the received interventions (9 studies overlapping with those reporting on blinding success). Occasionally, this was reported as a proxy for successful blinding but more commonly to study potential influences of patient expectancy on clinical outcomes. Only six trials (with seven control interventions) reported expectancy data in a manner that allowed for data pooling (Bialosky et al., 2009; Curtis et al., 2011; Nicholas et al., 2013; Bialosky et al., 2014; Michener et al., 2015; Coste et al., 2021). Apart from being an excessively small sample for meta-analysis (Borenstein et al., 2021), none of those studies reported significant differences in expectations between groups after an initial exposure to trial interventions. Meta-analysis and regression testing were thus not deemed promising and were not performed. In the remaining 23 trials with some mention of expectation-related assessments, the reasons for non-comparability were: 1) Inappropriate timepoint of assessments (either unexposed at baseline, or retrospectively after completing multi-session treatment programme and thus likely confounded by satisfaction), 2) compound assessment with treatment credibility without individually reporting data of expectancy-only questions (Constantino et al., 2018), 3) confounding expectancy and satisfaction, and 4) insufficient outcome data reporting.

Study outcomes

Extracted outcome measures

The most commonly extracted unidimensional outcome measures were pain intensity rating scales (92%). Where such ratings were not available, pressure pain thresholds were extracted (8% of the 168 pairs of active and control interventions with extractable unidimensional outcome measures). Multidimensional outcome measures were available for 130 comparisons. The extracted measures were mainly disability questionnaires (68%), followed by multidimensional pain questionnaires (19%), functional tests (8%), disease activity or symptom scores (5%), and general health questionnaires (2%).

Relatively more objective and supposedly pain-related outcome measures were employed in 37% of all trials, including functional tests, disease markers, nerve conduction, autonomic nervous system parameters, brain imaging, work absenteeism or medication use (data not extracted).

Attrition

The difference in percentual attrition between active and control intervention groups (i.e., the differential attrition) was a mean of 0.4 % more in active treatment groups (-.74 to 1.5 95% CI, Q(136) = 6054, $p = .51$, $T^2 = 43.52$, $I^2 = 97.8$ %), as estimated with a random-effects model with all studies weighted equally and applied to all large studies irrespective of therapy type (participants = 13,150, studies = 137, not reported in 6 studies). Longer studies reported more percentual attrition (r_s's(135) = 0.43, $p < 0.001$; 0.42, $p < 0.001$, respectively for active and control groups). Drop-out reasons related to the nature of the control intervention were reported in 9 studies (6.3%).

Risk of bias

In studies used for the primary meta-analysis, the overall risk of bias was judged low in 17% of studies, high in 44% and some concerns existed for 38% (Figure 6). In the smaller studies, 15% were rated as low risk, 82% high risk, and there were some concerns for 4% (not illustrated).

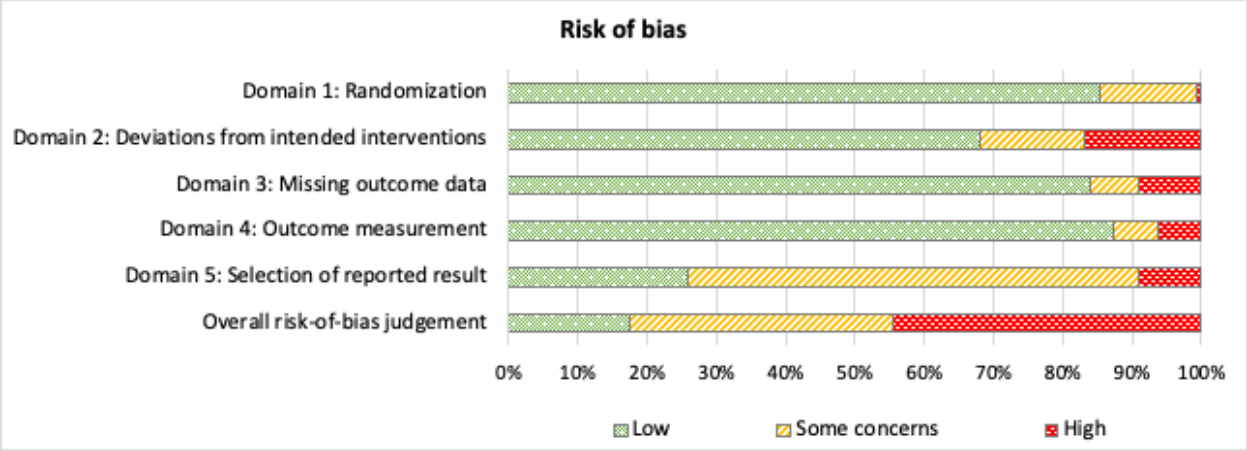


Figure 6: Risk of Bias assessment for all comparisons with 20 or more participants per group (eligible for primary analyses, n = 128). The overall risk of bias was judged "High" if at least one domain had a rating of high risk of bias or if there were concerns in more than one domain. The overall rating of "Some concerns" was given if there were concerns in any one domain, and "Low" if all domains were rated as of low risk of bias, as per the Cochrane Risk of Bias tool 2 (RoB 2) (Sterne et al., 2019).

Pain-related outcome measures

Effect sizes (as standardised mean changes, SMC) could be calculated for 166 intervention-vs-control comparisons for unidimensional outcomes and 125 comparisons using multidimensional outcome measures. For the entire sample, between-group differences were not significantly different when studies were grouped by high, some or low overall risk of bias (unidimensional outcomes: $F(2,164) = 1.217$, $p = 0.3$.; multidimensional: $F(2,123) = 0.231$, $p = 0.79$).

Meta-analyses and meta-regression

Out of 198 included comparisons, 120 were included in the primary meta-analyses, with a further 49 smaller studies used for secondary sensitivity analyses. Reasons for non-pooling included classification as outlier or apparent data errors (n = 5), patient populations in which no comparable improvements in pain report were expected (n = 8), and therapy subgroups too small for pooling (n = 16) (Figure 2 – PRISMA flowchart). Since not all studies provided data for each outcome, study numbers in the following analyses vary.

Sham control within-group effects (Placebo responses)

Across all studies included in the primary meta-analysis, changes from baseline to earliest follow-up within control intervention groups had an average effect size of -0.46 for unidimensional outcomes (-0.53 to -0.38 95% CI, $p < 0.0001$, $T^2 = 0.12$, $I^2 = 69.9$ %, participants = 10.557, studies = 112) and -0.32 for multidimensional measures (-0.39 to -0.24 95% CI, $p < 0.0001$, $T^2 = 0.08$, $I^2 = 62.06$ %, participants = 9.447, studies = 95), indicating small to moderate placebo responses.

Between-group differences in clinical outcomes (treatment effects)

The review was not designed to make judgements about the efficacy of the studied interventions. Mainly, to be a clinically useful review, outcomes other than the most short-term pain-related outcomes would have had to be extracted, and more attention paid to the homogeneity of study designs and populations. Accordingly, the resulting publications only presented forest plots and risk-of-bias assessments as a background to the subsequent meta-regression modelling. Similarly, summary effects per subgroup are presented in the meta-analysis overview table below (Table 11). Forest plots and risk-of-bias illustrations can be found in the respective publication (Hohenschurz-Schmidt et al., 2023c). In brief, small to moderate effects over control interventions were seen for all therapy subgroups. Effects were largest in the

subgroup of 'other manual therapies', i.e., massage and articulation-based interventions, and smallest for psychological interventions and spinal manipulation trials. Unidimensional outcome measures (mainly pain intensity ratings) showed generally larger effects than disability questionnaires and other multidimensional measures. Importantly, in most of the studied intervention subgroups, there was considerable unexplained heterogeneity in trial results between studies (Table 10).

Table 10: Meta-analyses results of test over control interventions per therapy subgroup and type of outcome measure. Primary analyses only. Between-group effects significant at 0.05 level are marked *. SMC = Standardised mean change.

Subgroup (therapy type, n of studies)	**SMC**	**95% CI**	***p* (two-sided)**	**Heterogeneity (I^2)**
Spinal manipulation				
Unidimensional (n = 35)	-0.36	-0.51 to -0.21	< 0.0001 *	71.1%
Multidimensional (n =24)	-0.26	-0.37 to -0.15	0.01 *	28%
Craniosacral and subtle myofascial techniques				
Unidimensional (n = 11)	-0.52	-0.84 to -0.2	0.001 *	78.1%
Multidimensional (n =13)	-0.47	-0.81 to -0.12	< 0.0001 *	83.8%
Other manual therapies				
Unidimensional (n = 31)	-0.72	-1.02 to -0.42	< 0.0001 *	90%
Multidimensional (n =22)	-0.45	-0.68 to -0.22	< 0.0001 *	79%
All manual therapy trials combined				
Unidimensional (n = 77)	-0.53	-0.68 to -0.39	< 0.0001 *	85.3%
Multidimensional (n =59)	-0.35	-0.47 to -0.23	< 0.0001 *	74.3%
Physical therapy interventions, rehabilitation, and exercise				
Unidimensional (n = 13)	-0.6	-0.89 to -0.3	< 0.0001 *	80%

Multidimensional (n =13)	-0.49	-0.74 to -0.24	< 0.0001 *	73%
Psychological interventions				
Unidimensional (n = 19)	-0.34	-.05 to -0.19	< 0.0001 *	62%
Multidimensional (n =18)	-0.2	-0.32 to -0.08	= 0.001 *	39%

Trial methods and study results (meta-regression)

In the following, an overview of meta-regression models is presented which examined the role of different variables in predicting heterogeneity in trial outcomes per intervention subgroup (Table 11). In most models, it was tested whether the degree of similarity between control and test interventions could predict variance in pain-related outcomes. Detailed regression tables of individual models are presented in the respective publication (Hohenschurz-Schmidt et al., 2023c).

In short, the regression models provided first-time empirical evidence to show that the heterogeneity in pain-related outcomes was frequently predicted by the degree of similarity between tested and control interventions. This was the case in all therapy types, and for ten out of twelve examined outcomes (classed as uni- or multidimensional outcomes). The proportion of unexplained outcome variance that could be predicted ranged from 15 to 73 percent. While regression models suggested an effect of six different similarity features, there was no similarity item that had a consistently predictive value (although some where significantly involved in several models). Also, the direction of effect varied, although it was mostly in the direction of less similarity predicting larger between-group differences. Apart from heterogeneity in pain outcomes, we examined if blinding effectiveness or differential attrition could be predicted by similarity ratings. This was the case for a significant proportion of unequal blinding and differential attrition. Conversely, however, blinding effectiveness or attrition did not predict variance in pain-related outcomes.

Table 11: Regression models with similarity ratings. Covariates that contributed statistically significantly to the respective model and overall significant models are highlighted (*) ($p < 0.05$). Estimated intervention effects may differ slightly from Table 9 due to different software employed in respective calculations. However, both estimates were calculated using random-effects models and the same weights were applied.

Subgroup and Meta-analysis of clinical outcomes (n, mean effect size and heterogeneity)	**Covariates in regression model** (Mean similarity rating and SD)	**Regression Model**						
		Q	**df**	**R^2**	**B**	**95% CI**	***p***	**Res. I^2**
Spinal manipulation								
Unidimensional (n = 35) -0.33 (I^2 = 94%)	- Number of sessions (1.91, 0.37) * - Information about efficacy (0.6, 1.03) * - Environment (1.43, 0.74)	50.33	3	0.59	-2.64	-3.35 to -1.93	0.000 *	10.1%
Multidimensional (n = 24) -0.26 (I^2 = 28%)	- Number of sessions (1.9, 0.45) - Information about efficacy (0.5, 1.02) - Environment (1.42, 0.72)	3.59	3	0.15	-0.74	-1.3 to -0.18	0.31	0%
Craniosacral and subtle myofascial techniques								
Unidimensional (n = 11) -0.52 (I^2 = 84.4%)	- Body areas (0.18, 1.9) * - Efficacy information (0.55, 0.93) * - Application mode (0.18, 2.1)	19.26	3	0.73	-0.37	-0.6 to -0.14	0.0015 *	0%

Multidimensional (n = 13) -0.47 (I^2 = 83.8%)	- Body areas (0.38, 1.9) - Efficacy information (0.46, 0.88) - Application mode (0.46, 2.03) *	7.0	3	0.38	-0.42	-0.79 to -0.06	0.02 *	21.6%
Other manual therapies								
Unidimensional (n = 31) -0.72 (I^2 = 90%)	- Number of sessions (1.81, 0.75) - Efficacy information (0.58, 0.85) - Environment (1.61, 0.62) *	5.35	3	0.16	-1.83	-3.07 to -0.59	0.004 *	1.6%
Multidimensional (n = 22) -0.37 (I^2 = 79%)	- Number of sessions (1.7, 0.88) * - Information about efficacy (0.5, 0.8) - Environment (1.6, 0.7)	9.97	3	0.37	-0.27	-1.13 to -0.6	0.019 *	0%
All Manual therapy trials combined								
Unidimensional (n = 77) -0.53 (I^2 = 85.3%)	- Number of sessions (1.88, 0.5) * Information about efficacy (1.1, 1.3) - Application mode (0.34, 1.7) - Body areas (1.4, 0.8)	16.94	5	.17	-1.57	-2.18 to -0.95	0.007 *	9.3%

	- Environment (1.47, 0.8)							
Multidimensional (n = 59) -0.35 (I^2 = 74.3%)	- Number of sessions (1.83, 0.62) Information about efficacy (0.49, 0.9) - Application mode (0.81, 1.81) - Body areas (1.19, 1.38) - Environment (1.42, 0.81) *	17.22	5	0.23	-0.39	-0.84 to 0.05	0.004 *	8.72%
Physical therapy interventions, rehabilitation, and exercise								
Unidimensional (n = 12) -0.46 (I^2 = 83.7%)	- Individuali-sation (-0.17, 1.3) * - Fidelity monitoring (1.0, 1.1)	12.28	2	0.62	-0.73	-1.07 to -0.39	.0000 *	0%
Multidimensional (n = 14) -0.54 (I^2 = 79%)	- Individuali-sation (0.59, 1.6) - Fidelity monitoring (0.33, 0.72)	1.1	2	0.1	-.056	-0.85 to -0.27	0.57	25.1%
Psychological interventions								
Unidimensional (n = 19) -0.33 (I^2 = 82.6%)	- Application mode (0.84, 1.7) * - Attention focus (-0.26, 1.6) *	11.48	2	0.41	-0.4	-0.55 to -0.24	0.003 *	1.8%
Multidimensional (n = 19) -0.3 (I^2 = 72.1%)	- Application mode (0.84, 1.7) *	7.18	2	0.21	-0.49	-0.73 to -0.24	0.0001 *	41.9%

	- Attention focus (-0.26, 1.6)							
<u>Other regression analyses</u>								
Blinding index subgroup								
Bang's blinding index (*g*) (n = 17) 1.32 (99.7%)	- Sensations (1.24, 1.15) * - Application mode (1.71, 0.99)	8.92	2	0.35	3.0	1.27 to 4.7	0.0006 *	15.1%
BI and pain-related outcomes								
Studies with BI and unidimensional pain-related outcomes (n = 15) -0.21 (43.7%)	Bang's BI (2.1, 3.55)	0.01	1	0.00	-0.21	-0.43 to 0.01	0.93	10%
Differential attrition and similarity between active and sham control interventions								
Differential attrition (n = 111) 0.31 (60.2%)	- Number of sessions (1.69, 0.9) - Standardi-sation (0.97, 1.2) - Fidelity Monitoring (0.61, 1.0) * - Individuali-sation (0.5, 1.5) - Application mode (0.8, 1.8) - Level of participation (0.9, 1.5) - Environment (1.2, 1.1)	15.1	6	0.13	-1.13	-4.1 to 1.83	0.45	0%
Differential attrition and trial outcomes								

Differential attrition; unidimensional (n = 121; 1.26, 7.0) and multidimensional measures (117; 0.91, 7.62)	No model significant							

Summary of review findings

The systematic review analysed 198 control interventions and compared them to respective experimental treatments, identifying a range of common control intervention designs. The review also found notable gaps in reporting important information about the development, rationale, and validation of employed control interventions, complicating the assessment of control intervention quality as well as the replication of methods by future researchers. Blinding effectiveness and participants' expectations were also rarely reported and, if so, in a variety of ways. In a subgroup where blinding effectiveness could be meta-analysed, control intervention participants guessed their group allocation correctly disproportionately often, indicating that blinding is often unsuccessful.

Based on the concepts of 'structural equivalence' and 'indistinguishability' (Baskin et al., 2003; Machado et al., 2008; Puhl et al., 2017), the review provided a detailed assessment of the similarity between control and experimental interventions. In this sample of trials, similarity was prioritised for features concerned with the extent and timing of treatments and outcome assessments and the delivery format. The environments in which control and experimental interventions took place were also similar on average, but the variability was larger and non-reporting contributed to lower ratings. Many other compared features were less commonly matched between groups. These concerned the patient experience (e.g., treatment-specific sensory cues such as touch or sound, attention focus during interventions, personal interactions

with providers and staff), procedural aspects of interventions (individualisation to patients, similarity and complexity of physical procedures, devices used in the application of control but not experimental treatments, use of co-interventions), and research-related aspects (e.g., differences in fidelity monitoring). Further, developing closely matched control interventions is less common for complex intervention studies and in physiotherapy trials.

The meta-analyses of trial outcomes found moderate placebo responses across PPS therapies for pain, with patients in control intervention groups improving with an average effect size of 0.46 from pre-treatment to the first post-treatment assessment. The tested interventions conveyed benefits over control interventions in all therapy subgroups; with effect sizes ranging from 0.34 in psychological interventions (number needed to treat, NNT ≈ 9) to 0.72 in some manual therapies (NNT ≈ 4) (NNT calculated according to Magnusson, 2021). However, only 17% of trials were judged as low risk of bias. Effect sizes were smaller when multidimensional measures such as disability questionnaires were employed, as opposed to simple pain intensity scales.

In most of the studied intervention subgroups, there was considerable unexplained heterogeneity in trial results between studies. Assessing and rating the level of similarity between experimental and control interventions, however, explained some of that heterogeneity. For example, a proportion of variance in trial outcomes was explained by how different tested and control interventions were in terms of the number of treatment sessions, application modes or treatment environment. Further, trials were at a higher risk of differential attrition when there were differences in monitoring of the groups' treatment adherence. In a subgroup of studies where the blinding success of control interventions had been reported, meta-regression analyses were also able to partially predict the risk of unbalanced blinding.

Finally, this large review was only made feasible thanks to a dedicated reviewer crowd, supervised by the thesis author.

Discussion of review findings

The review's findings largely confirmed that blinding and so-called sham control methods in nonpharmacological pain research vary considerably. Providing a comprehensive overview of methods in physical, psychological, and self-management trials, the review also allows for a detailed comparison of research practice between different therapy fields. Other available reviews have painted a similar picture of methodological diversity, but were often limited to specific therapies (Puhl et al., 2017; Braithwaite et al., 2019; D'Alessandro et al., 2022) (with few exceptions (e.g., Faltinsen et al., 2022)). Beyond describing blinding methods, the review also showed that the reporting of control intervention development and underlying rationales is mostly rudimentary or non-existing, although this is important information for the appropriate interpretation of trials. This finding mirrors a recent review of the reporting of control interventions (Webster et al., 2019).

By including trials from several nonpharmacological disciplines, the present review has demonstrated large differences in how established the aspiration of indistinguishability is between therapy fields. More importantly, and despite suggestions to the contrary (NICE, 2016; Williams et al., 2021), the review showed that the design and implementation of largely *similar* control interventions is feasible in any of the reviewed therapies. This suggests that there is a potential for some research areas to learn from the experience of others, and indeed that a unifying guideline across physical, psychological, and self-management interventions may be useful. While challenges may differ according to the studied intervention's complexity, interactivity, theory, and available mechanistic evidence, the review suggests the presence of a set of transferable best-practice considerations. To identify these principles is the objective of the next project stage.

The review's meta-regression analysis provided novel insights about the importance of resemblance between control and tested interventions for bias control, also adding empirical evidence to an ongoing debate about the value of blinding in trials. Meta-epidemiological studies of the role of blinding typically rely on the reported blinding status of participants (Armijo-Olivo

et al., 2020, 2017; Hróbjartsson et al., 2007; Moustgaard et al., 2020). Advancing the field, the present review's analysis was independent of reported blinding status; such reporting being poor (Alvarez et al., 2020; Boutron et al., 2007; Villamar et al., 2013; Webster et al., 2019) and recognising that successful blinding cannot be assumed unless specifically assessed (Beedie et al., 2018). The review circumvented these problems by instead assessing design features of control interventions in supposedly blinded trials, showing that the degree of similarity between control and test interventions alone influences trial outcomes. It remains unclear whether this effect is mediated through discrepant expectations and behaviour from unsuccessfully blinded participants or through differential exposure to contextual factors. On one hand, successful blinding could be considered as secondary based on the review's findings in a subgroup of studies that reported on blinding effectiveness: This analysis showed no influences of unbalanced blinding on pain-related trial results. On the other hand, this analysis sample of trials had little heterogeneity in pain outcomes between them, which limited the scope for meta-regression. Regarding the question of how important blinding is, the review was thus unable to add much clarity to an already discordant research field: Several evidence synthesis projects have now provided supportive (Hróbjartsson et al., 2007, 2014; Savovic et al., 2012, 2018; Colagiuri et al., 2019; Freed et al., 2021; Karjalainen et al., 2022), contradicting (Moustgaard et al., 2020), or inconclusive (Armijo-Olivo et al., 2017, 2020; Braithwaite et al., 2018) evidence regarding the role of blinding in preventing biased outcomes. However, by not solely testing blinding status or reports, this review added weight to the hypothesis that another biasing mechanism is important: differential exposure to contextual effects through dissimilar control interventions. With the largest analysis to date, and confirming an early study by Baskin et al. (2003), this review provides empirical data to support the assumption that control interventions that resemble the index treatment are preferable to dissimilar ones. High-similarity control interventions match known and unknown contextual factors between trial arms, reducing the risk of biased trial outcomes. While this was demonstrated here by synthesising relevant studies, research about the placebo effect has long suggested this to be true: A large body of experimental studies and reviews supports the effect of contextual factors on health outcomes, including the effects of healthcare settings, provider characteristics and interactions with patients, and amount of

exposure (Dunlop et al., 2012; Beedie et al., 2015; Bishop et al., 2017; Chen et al., 2019; Colloca, 2019a; Daniali and Flaten, 2019; Davis et al., 2020; Faltinsen et al., 2022; Hutchinson and Moerman, 2018; Meissner and Linde, 2018; Lene Vase et al., 2015a; Rossettini et al., 2020; Sherriff et al., 2022). In addition, this review identified features as influential that are more related to trial conduct rather than treatment context, such as similarity in fidelity monitoring. This mirrors findings from a meta-analysis of psychotherapy trials for depression (Mohr et al., 2014) and may be mediated by behavioural rather than expectancy effects.

Limitations of the review

There are general challenges for meta-epidemiological studies. Mainly, evidence synthesis relies on the assumption that the effects of methodological differences and random variations between primary studies become negligible when the number of synthesised studies is large enough (Hempel et al., 2011). Nonetheless, the included trials need to be sufficiently similar in design, patient population, and outcomes to be comparable. While all these points were levered as criticism against a recent meta-epidemiological review of blinding in trials (Moustgaard et al., 2020; Tack, 2021), the present review was arguably stronger, including only trials for pain, with pain-related outcome measures, and subgroups of ten or more trials. Still, the influence of chance and unknown confounders cannot be ruled out entirely. Notably, this may have influenced the direction of effect in some findings: (Dis)similarity between control and index interventions predicted variance in outcomes for numerous assessed features – but the direction of effect was inconsistent. Another potential explanation for this finding is that the direction of bias from insufficient blinding or patient expectations may vary, largely depending on study context (Tack, 2021). Kleijnen (2020) illustrates this with hypothetical homeopathy trials, where bias from unblinded outcome assessors may increase observed effects if the research staff are supportive of homeopathy, and vice versa if a trial is conducted by sceptics. In the present review, (dis)similarity may have biased trial outcomes in several different ways, including through the highly variable and context-dependent placebo effect or through behavioural alterations of non-blinded or uncertain trial participants.

This review can further be criticised for not having included studies published before 2008. The decision to search from 2008 onwards was based on the hope that reporting would be more complete after the first CONSORT extension for nonpharmacological trials had become available (Boutron et al., 2008). There is, however, conflicting evidence to this effect (Kane et al., 2007; Hoffmann et al., 2013; Nagendran et al., 2013; Ma et al., 2016; Alvarez et al., 2020). For example, Alvarez et al. (2020) only report small positive trends in manual therapy trial reporting since the introduction of the reporting guidance. Nonetheless, not including older studies had other benefits: First, while reporting may not have improved, the average quality of trials likely has (May, 2019), possibly also owing to the increasing number of publications in the PPS field focusing on internal validity (e.g., Draper-Rodi et al., 2022). Second, the 2008 cut-off allowed for the critique of relatively current methods. Third, a longitudinal analysis, for which a longer timeframe had been preferable, was beyond the scope of this work. Lastly, the shorter inclusion period made the project somewhat more feasible by reducing the number of included trials. Given the small number of pre-2008 trials (see Figure 1), this latter effect was probably negligible, and so was the potential for bias from this time limit.

Another method to discuss critically is the employed rating process of similarity between control and test interventions. Reviewers rated those from dissimilar (-2) to similar (2), also assigning values of -1 and 1 to 'Probably dissimilar' and 'Probably similar', respectively, 1 to 'Partly similar', and 0 to 'Not reported' items (illustrated in Table 12 for ease of understanding). No value was assigned if similarity assessments for particular items were not relevant.

Table 12: Method of rating the resemblance between control and experimental interventions. Assessments were performed based on trial reports and published protocols and supplements during data extraction and by two independent reviewers. These assessments were performed for each of the 25 design features examined per comparison.

Assessment of similarity	Dissimilar	Probably dissimilar	Not reported	Probably similar Or: partly similar	Similar	Not applicable
Value assigned	-2	-1	0	1	2	Not rated

This rating approach addressed the following challenges: First, resemblance could only be assessed if sufficient information was reported. Explicit reporting, however, could not be expected but induction often appeared possible, so that 'probably' answer options were introduced. An example is the statement that a study was a single-centre trial, but no explicit information is reported whether test and control interventions were delivered in the same room. In this case, 'Probably Yes' would be selected in response to the question if treatment environments were similar. Introducing an element of informed guessing on behalf of the reviewers was deemed preferable to loosing these data completely. Also, reviewers were instructed to use this option conservatively and choose 'Not reported' if needed. Further, there were scenarios for which a 'Partly' answer was required. For example, a meditation study may have provided home-based practice exercises only for participants in the tested intervention arm. Even if both the test and the control participants received instructions in the same clinic rooms, 'Partly' would have been selected due to the partial deviation in treatment environments for patients performing additional exercises at home. Another challenge was how to incorporate 'Not reported' and 'Not applicable' ratings. In effect, by assigning any value to relevant but not reported information, trials' similarity ratings would be influenced by information that did not derive from actual similarity considerations. On the other hand, 'not reported' and 'not relevant' were clearly different types of information. In the end, 'not relevant' items were excluded at study-level, and relevant but 'not reported' items were assigned the value 0.

Finally, the synthesis of similarity ratings was challenging, as these derived from up to 25 individual design features. Initially, the intention was to employ a single 'similarity index' value

for analysis (e.g., the average rating across all items per study). However, the validity of this approach appeared uncertain for numerous reasons: many items appeared interlinked (and ratings were indeed often statistically correlated); for example, trials with the same number of treatment sessions between control and intervention groups also often employed treatment sessions with the same duration. In this example, trial authors probably strove to make the amount of intervention exposure the same, rather than considering session number and duration independently. When contemplating meta-regression analyses, asking whether these ratings predicted pain-related outcomes, blinding effectiveness, or control intervention acceptability, it seemed inappropriate to weigh all items equally due to this partial lack of independence. Further, dissimilarity for some items may act as a 'give-away' clue for unblinding. For example, when only asymptomatic body parts are 'treated' in the control group, participants are likely to develop doubts (Hohenschurz-Schmidt et al., 2021). Other features, such as the extent of exposure to context factors (through similar amounts of provider contact, similar environments etc.) may have disproportionate influences on placebo effects. For these reasons, the average across all items was only presented descriptively, and its interpretation qualified as outlined. Similarly, the selection of individual similarity ratings for meta-regression modelling focussed on items that were easily rated, well-reported, conceptually distinct, and statistically independent from one another.

Overall, the approach to rate the similarity between control and test interventions proved an effective way of describing and analysing complex control interventions but will require further development and validation. If anything, this review laid the groundwork to identify a smaller number of features that are easily rated retrospectively, independent from one another, and capture influential aspects of control intervention design. Despite some limitations, the approach resulted in findings that were consistent with existing evidence, partly recurring across different therapy subgroups, and that were confirmed in sensitivity analyses. As will be seen, the resulting methodological guideline incorporated such considerations.

Equipped with this comprehensive overview of current methods, and reassured of the importance to standardise control intervention development and implementation, the PhD project then embarked on a consensus process to develop a guideline for this purpose.

Chapter 3: Guideline Development for Control Intervention Design, Implementation, and Reporting

Background

So far, the project confirmed the heterogeneity of control and blinding-related methods which underlies much methodological criticism of efficacy trials of PPS interventions. The review in chapter 2 provided a comprehensive overview of control methods in trials of the past 13 years and demonstrated the impact of (dis)similarity between control and tested interventions on trial outcomes. Despite the heterogeneity in methods, the review also suggested the existence of a unifying principle of best-practice control design in efficacy and mechanistic RCTs, namely the attempt to replicate as many of the tested intervention's components as possible in the control intervention. Importantly, the review also included trials which showed that such controls were possible in multiple, different therapies, and uncovered approaches to further enhance the success of control interventions.

This was deemed a good basis to create a best-practice recommendation framework for control intervention design, conduct, and reporting, applicable across PPS interventions. The thesis author was aware that recommendations for high-quality control interventions had to go beyond the design process and include multiple related considerations. For example, guidance might have to propose how to prepare trial therapists for the delivery of a control intervention or how to assess the control's quality. The existence of various stakeholder groups with potentially conflicting viewpoints was also acknowledged. For example, researchers might consider blinding methods that involve deception as acceptable, but such methods may be ethically questionable or not acceptable to patients. To reflect such challenges, this project stage was designed to draw on the practical experience of trial researchers, the insights of researchers of the placebo effect, and the lived experience perspective of potential trial participants. To achieve this, the thesis author adapted the Delphi consensus technique for the project's purposes and embedded it in a broader guideline development process.

Before describing the methodology, this paragraph will reflect on existing guidance for nonpharmacological control interventions and its limitations. For the interventions discussed in this thesis, such documents are geared towards individual therapies, such as psychotherapy (Baskin et al., 2003), behavioural interventions (Aycock et al., 2018), rehabilitation (Hart and Bagiella, 2012), sports and exercise (Beedie et al., 2018), physiotherapy (Braithwaite et al., 2020b), and manual therapy (D'Alessandro et al., 2021). These guidelines provide no quality checklist, few generalisable principles, and often disregard problems of intervention complexity (Skivington et al., 2021). Consequently, fundamental questions of control intervention design, such as how closely the control should resemble the study intervention, are approached differently in various fields. Whilst there is a trend towards high-similarity controls in some research areas (notably spinal manipulation), in others (e.g., other physical therapies and psychological therapies) control interventions often do not resemble the study treatments or efficacy trials are avoided altogether (Hohenschurz-Schmidt et al., 2023b). A lack of consensus on relevant issues exposes the field to justifiable criticism due to concerns over bias (Bueno and Harrison, 2019; Furukawa et al., 2014; Pagoto et al., 2013; Hohenschurz-Schmidt et al., 2021, 2022a) and leaves questions of treatment mechanism unanswered. Similarly, 'sham' control interventions for nonpharmacological interventions are different to placebos for drugs, but the concept of 'placebo controls' originated from and is still influenced by drug research. Existing guidance for nonpharmacological trials often acknowledges potential implications for control intervention design and implementation, but rarely for the interpretation of respective trials (Walach, 2001). Finally, the only reporting checklist for control interventions covers both drug and non-drug interventions (TIDieR-Placebo) (Howick et al., 2020). It may thus not be sufficiently comprehensive to reflect the challenges of complex control interventions in all types of PPS intervention efficacy trials.

A positive example of guidance for nonpharmacological 'placebo controls' are the ASPIRE guidelines for surgery trials (Beard et al., 2020), published in *The Lancet* in March 2020 and cited over 45 times since (as of 23 March 2023). Although the development of the ASPIRE guidelines

themselves is not transparent (involving only a "workshop" of undisclosed methodology with subject experts and patients), ASPIRE does provide a generalisable framework for the design of such control interventions. ASPIRE also considers questions of practical implementation, acceptability to patients and providers, ethics, and trial interpretation, providing a template for guidance considerations in other research fields. Separate efforts for PPS interventions were nonetheless required because the ASPIRE does not translate well to non-surgical trials, mainly due to surgery's distinct risk profile and the opportunity to use general anaesthesia for blinding.

The objectives of this project part were thus to produce a comprehensive recommendation statement for the development, conduct, reporting, and interpretation of control interventions in efficacy and mechanistic trials of PPS interventions. This statement was to contain generalisable principles for all PPS interventions, while providing sufficient nuance to reflect the complexity of these interventions and the challenge of blinding in such trials. Finally, differences to drug research had to be acknowledged.

Methods

To develop the guideline, the Delphi technique was embedded within other common methods of guideline formulation, i.e., the systematic review discussed in the previous chapter, stakeholder interviews, and consensus meetings (Figure 7). The thesis author adapted this process from instructions for developers of health research reporting guidelines (Moher et al., 2010) and clinical practice guidelines (Brouwers et al., 2010). Deemed best practice, these methods are also a prerequisite for the uptake of reporting guidelines into the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) database (personal communication with EQUATOR). Although the present guideline focuses on development and implementation aspects, rather than reporting, these methods represent the highest available standard for guideline creation. They also enabled the author to draw on input not just from subject experts (i.e., Delphi panellists), but also on empirical evidence from the systematic review, on methodological guidance developed by others previously, and on the views of potential trial

participants through stakeholder interviews. Employing a Delphi method ensured that the core recommendations would be consensus-based and not influenced by certain social biases. Conversely, employing live consensus meetings in addition to the Delphi method facilitated a nuanced discussion of draft recommendations, necessary to reflect the complexity of the challenge of control interventions in nonpharmacological trials. In the following, the guidance development process is positioned philosophically before individual methodological steps are detailed.

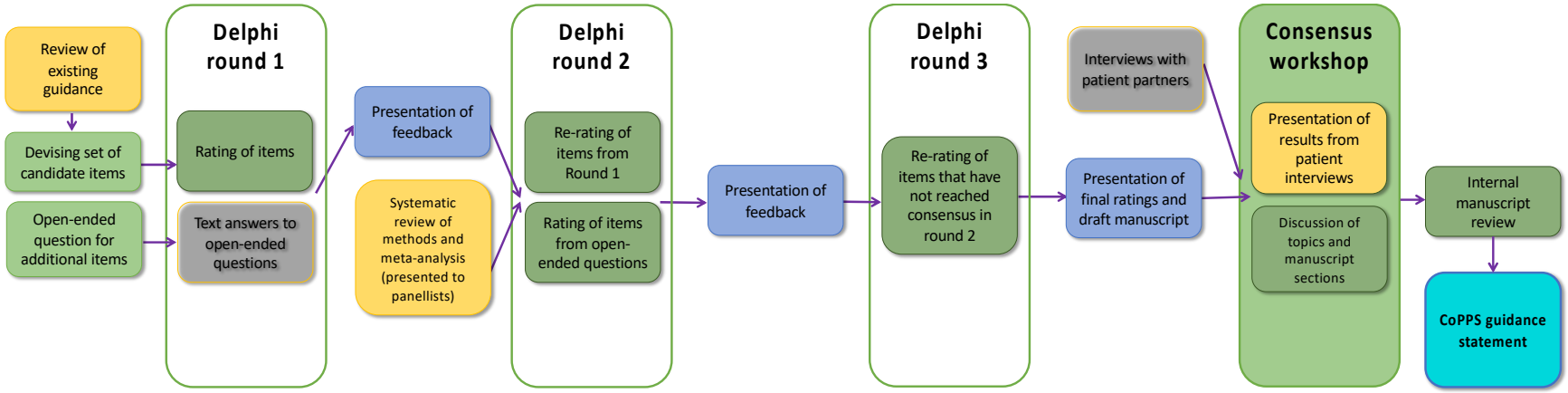


Figure 7: Flowchart of guidance generation process as per pre-registered study protocol. Colour legend: Yellow, external input into the expert-based consensus process; Dark green, expert input; Bright green, main methodological components; Turquoise, final guidance statement.

Philosophical positioning of the guideline development process

The guideline development process was influenced by several, arguably partly conflicting, philosophical stances. On one hand, systematic review and Delphi methods employ quantification and objectification of information. These methods are typically thought to draw on positivist or post-positivists paradigms, seeking to uncover or approach an objective 'truth' in a single reality (Guba and Lincoln, 1994; Shariff, 2015). Confidence in such scientific 'truth' certainly informed my attempt to ground the Delphi process in empirical data from a systematic review (and a meta-analysis which sought to quantify influences of trial methods on outcomes). On the other hand, influences from constructivist philosophy are apparent. At ontological level, this paradigm presumes multiple perspectives on what *is*; epistemologically, constructivism postulates that knowledge is socially constructed (Alvesson and Sköldberg, 2017). I was conscious of my own dominant role in the project, for example leading the design of the research process, conducting and interpreting interviews with stakeholders, moderating consensus discussions, and writing the final manuscript. In doing so, I influenced the direction of the project, and my input likely changed as my own understanding of the subject evolved. Acknowledging my own influence is consistent with a constructivist perspective on research, which acknowledges the "*meaning-making activity of the individual mind*" (Schwandt, 1998, p. 240, cited in Lee, 2012). Rather than negating any researcher bias, I instead critically reflect on the influence of my own philosophical stance in the reflection sections of this thesis (Andersen et al., 2019). Again embracing a constructivist stance, I actively promoted the involvement of subject experts and lay people so that their experiences and personal perspectives affected the research output. No longer a purely post-positivist approach, the project design acknowledged (and indeed valued) the existence of a multitude of 'realities' in the "*intersubjectively shared, social constructions of meaning and knowledge*", best termed 'social constructivism' (Schwandt, 1998, p. 240, cited in Lee, 2012). At the same time, post-positivist elements remain strongly represented: Throughout the process, my research design limited the ability of individuals (including myself) to unilaterally determine the research output, for example through the anonymous nature of the Delphi method and through multiple checks and balances during the latter parts of the consensus process (described below). In negotiating the research output in this manner, I drew on

Hegelian dialectic (Shariff, 2015) and post-positivism, which assumes that reality can only be known probabilistically (Guba and Lincoln, 1994).

Interestingly, the observed tension between (post-)positivism and constructivism is also seen in clinical trial and medical research at large: While clinical trials are traditionally understood as deeply rooted in the positivist belief that some fundamental truths can be discovered, medical researchers are incorporating multiple methods that acknowledge subjective and inter-subjective experience as important sources of knowledge. In clinical trial research, examples include patient and public involvement initiatives (Greenhalgh et al., 2019; Locock and Boaz, 2019), patient-reported or even individualised outcome measures (Sales, 2017), and qualitative evaluations of stakeholder experiences (Corry et al., 2019). Indeed, it might be the combination and triangulation of different methods which facilitate an improved understanding of the complexity of health and disease, and the human experience at large (Turyahikayo, 2014; Rocca and Anjum, 2020); an endeavour and pragmatic ontology to which I am committed, too.

Relating to the present guideline development, I "*organize[d] the research project as a collaborative learning community of scholars and practitioners with diverse perspectives*" (Van De Ven and Johnson, 2006, p. 815). The feedback from panellists shows that this endeavour was successful (see Appendix 2: Consensus methods, additional content').

I now describe the individual methodological steps of our guideline development process.

Delphi methods

The Delphi technique is commonly used in health research to develop "*consensual guidance on best practice*" (p. 685, Jünger et al., 2017). The method can be tailored to serve the objectives of a given project, although it is characterised by four main features: the consultation of experts (so-called 'panellists'), anonymity to avoid undue influences of hierarchies and social desirability, iterative rounds of questions or surveys, and a mechanism by which answers from previous rounds influence subsequent rounds.

The primary purpose of the Delphi process in this project was to derive a consensus-based set of recommendations. To do so, a Delphi technique was used with experts from the fields of PPS clinical trial design and placebo science. The Delphi technique was modified to allow for the elicitation of panellists' opinions but also for the consideration of the results from the preparatory systematic review of methods (Hohenschurz-Schmidt et al., 2023b, 2023c) (Figure 7). The Delphi consensus provided the basis for consultation research with people with persistent pain, live workshops, and guideline writing, the methods of which are described below. The Delphi technique was deemed appropriate for this part of the study because it allowed for consensus formation in a highly variable research field and facilitated correlation of judgements across several research areas.

The Delphi methods and reporting were mainly based on the Recommendations for the Conducting and REporting of DElphi Studies (CREDES) (Jünger et al., 2017) and personal communication with the CREDES lead author Professor Saskia Jünger (Bochum University). Other methodological guidance included Hasson et al. (2000), Iqbal and Pipon-Young (2009), McPherson et al. (2018).

The Delphi was restricted to three rounds to limit research burden and attrition risk (McPherson et al., 2018). The first Delphi round started with a survey instrument informed by a topical review of the literature, identifying previously recommended considerations for control interventions, and a set of open-ended questions (Sinha et al., 2011). The aims of the first round were to obtain an initial level of agreement with potential recommendation items for control interventions in PPS trials, and to elicit additional considerations. Prior to round 2, panellists were then presented with the results from the systematic review (Hohenschurz-Schmidt et al., 2023b, 2023c) and a summary of panellist ratings from round 1. Rounds 2 and 3 aimed at identifying the level of consensus for individual items when presented with current evidence and the views of peers.

Development of the initial Delphi questionnaire

The round 1 Delphi survey was developed based on pertinent literature, discussing or developing 'placebo', 'sham', or 'attention' control interventions in nonpharmacological

therapies or making recommendations for their reporting. Specifically, the articles used and available at the time were: (Turner et al., 1994; Baskin et al., 2003; Safer and Hugo, 2006; Tough et al., 2009; Moher et al., 2010; Campbell et al., 2011; Hart and Bagiella, 2012; Gooberman-Hill et al., 2013; Hróbjartsson et al., 2014; Chaibi et al., 2015; Felson et al., 2016; Boutron et al., 2017; Puhl et al., 2017; Aycock et al., 2018; Braithwaite et al., 2019, 2020). The systematic review of methods extracted data for the same considerations of control and blinding methods. The review was presented to panellists only *after* round 1 to obtain uninfluenced responses to the initial set of items and open-ended questions. Please see Appendix 2: Consensus methods, additional content for an overview of initial survey items.

As per recommendation (Sinha et al., 2011; Jünger et al., 2017), the first round survey included a set of open-ended questions (see Appendix 2: Consensus methods, additional content), "*[s]o that researchers do not impose their views on participants and thus introduce bias into the study*" (p.3). Specifically, these questions were aligned with the overarching research objective, intended to elicit additional considerations not captured by the systematic review and not commonly discussed in the literature.

The initial survey instrument was formulated with help of the supervisory / steering group. The survey was piloted and refined based on feedback from two manual therapy and one psychology researcher, as part of the protocol development. Doing so, the risk of employing poorly worded questions was reduced, which can lead to erroneous answers, research waste, and poor use of participants' time (Bautista, 2012). Apart from further detail in Appendix 2: Consensus methods, additional content, the full survey instrument and results from round 3 are provided on the project's webpage at https://osf.io/jmyhq/.

Rating process

Panellists rated the Delphi items on a 9-point rating scale, anchored at Strongly Disagree (1) and Strongly Agree (9) (Table 13). Panellist were asked to rate each item, answering the following question: *Should researchers planning or conducting a placebo controlled RCT in a PPS intervention for pain consider the following item?*

Table 13: Rating scale employed during Delphi rounds. This is the original RAND UCLA rating scale, developed for Delphi forecasting methods, employing a 9-point Likert instrument (Rescher, 1998).

Strongly Disagree	Disagree	Moderately Disagree	Mildly Disagree	Undecided	Mildly Agree	Moderately Agree	Agree	Strongly Agree
1	2	3	4	5	6	7	8	9

Definition of consensus

Consensus was defined as 80% of the *entire* panel rating an item as not worth considering (rated as 1-3) or worth considering (7-9). Eighty percent agreement is a common threshold for consensus definition in Delphi studies (Jünger et al., 2017). Once an item had reached consensus, panellists did not have to rate it again in subsequent rounds (in which case the panellist's rating was set to the average rating from the previous round), unless desired (in which case the new rating was used to inform the overall rating). In this way, time requirements for panellists were reduced but panellists retained the ability to influence consensus levels.

Items that reached consensus in individual panel groups (e.g., only amongst placebo researchers) but not on average across the entire panel were further discussed at the live workshops; these items were mostly not included in the final guidance and if so, only in a qualified form.

Feedback and information between rounds

Between rounds, each panellist was provided with a descriptive summary of the panel's ratings, separated into four groups: The entire sample, experts in clinical trials, placebo researchers, and experts in both research fields. For each item and group, average ratings were presented and the number of panellists per rating score (i.e., the overall distribution). Items for which agreement or disagreement had been reached in a previous round (as defined above), were visually highlighted.

Between round 1 and 2, participants were presented with two sets of information: the ratings for each survey item, and the methods and results of the preparatory systematic review of control intervention methods in trials of PPS interventions for pain (Hohenschurz-

Schmidt et al., 2023b, 2023c). Further, responses to open-ended questions from round 1 were converted into new items that were rated in subsequent rounds. The process of structuring the responses to open-ended questions from round 1 entailed elements of thematic analysis (Braun and Clarke, 2006). Coding and identification of themes was performed by two independent researchers (the thesis author and Jules Phalip), reconciled by the author, and cross-checked by an expert in qualitative research methodology (Dr Oliver Thomson, University College of Osteopathy). The study's steering group (consisting of the PhD supervisors) accepted the resulting statements to represent the panellists' views accurately and truthfully, accepting an element of researcher influence as rooted in constructivism (Alvesson and Sköldberg, 2017).

After round 3, panellists were provided with an initial draft guidance which included all items for which consensus had been reached. This information prepared panellists for the subsequent guidance development workshop (see below). The draft guidance also was the basis of interviews with eight non-researchers living with persistent pain (methods and results below).

Selection of stakeholder groups and nomination of panellists

Trial researchers from all PPS interventions were invited to participate in the guidance development process, beginning with the Delphi study. As a second panellist group, experts in placebo science were recruited, because control interventions in efficacy trials are mainly employed to control for the placebo effect and related phenomena (such as expectations). The author also hoped to reduce the risk of circularity by including placebo experts and not focussing on a single PPS intervention: This approach avoided simply replicating current practice of a specific research field without introducing new ideas for communal consideration.

Formally, Delphi panellists were included if they consented and fulfilled the eligibility criteria below (Table 14). Broadly, the sample consisted of people with practical or theoretical experience with a) 'placebo' or 'sham'-controlled efficacy or mechanistic clinical trials of a PPS intervention and / or b) the placebo effect.

Table 14: **Eligibility criteria for panellists**, recruiting into either of two expert groups.

Inclusion criteria	
Clinical trial researchers	Either: • At least one peer-reviewed publication related to 'placebo' or 'sham'-controlled clinical trials of PPS interventions (irrespective of author position and including systematic reviews or methods papers concerned with such trials) Or: • Academic or professional training related to clinical trial methods, i.e., at least an undergraduate module at a registered teaching institution o *and* professional background in a PPS intervention, i.e., at least undergraduate degree from a registered institution *or* research interest in clinical trials of PPS interventions
Placebo researchers	At least one peer-reviewed publication related to placebo effects or mechanisms (irrespective of author position)
Exclusion criteria	
Inability to understand and write English	
Inability or unwillingness to complete up to four online surveys within the expected timeframe	

Sample size of Delphi panel

Formal sample size calculations are uncommon in Delphi designs. Most studies recruit between 10 and 20 experts (Hasson et al., 2000; McPherson et al., 2018). In this study, the decision to stop recruitment was influenced by whether we recruited leading experts in the field, because the success of a Delphi depends on both adequate panel size and the qualification of panellists (Powell, 2003; McPherson et al., 2018). Given that two distinct groups were recruited and that the accepted minimum number of panellists is ten, at least 20 individuals had to be recruited, and more if leading experts declined participation. Due to the broad scope of therapies addressed under the umbrella of PPS, the aim was to recruit significantly more panellists for this group, aiming for a total panel size of 30-50. A recent comparable Delphi study recruited two groups of experts (in research methodology and in deceptive or hypnotic techniques) but was limited to physical interventions only, resulting in an overall sample of 38 panellists (Braithwaite et al., 2020b).

Sampling strategy and recruitment rate

Purposive sampling was used, sending a standardised email invitation to corresponding authors of relevant publications. Corresponding authors of all eligible trials of the prior systematic literature search were contacted, which covered all trials in the pain field of the past 13 years and should have resulted in a representative list of invitees. Snowball sampling was also used. Further, the author attempted to recruit a diverse sample regarding geographic location, therapeutic area of interest, gender, and career stage, for example searching for specific academic profiles and research groups.

Adverts for study participation were distributed on social media and via the newsletter of the Society for Interdisciplinary Placebo Studies (SIPS). The latter was the main mechanism to recruit experts in placebo research.

No financial incentives for participation were offered. However, participants were informed that they could decide to co-author the guidance document upon their completion of the consensus process. Designated co-authors also had to revise and approve the guidance manuscript and agree to be accountable for its content, as per ICMJE criteria (International Committee of Medical Journal Editors, n.d.).

Participant retention and response rate to survey rounds

Losing panellists over the course of the study is an important threat to the internal validity of a Delphi technique, next to insufficient or biased recruitment (McPherson et al., 2018). Possible reasons for panellist attrition are survey workload, disinterest, discontent with the direction of the study or its management, and external reasons (Sherwood et al., 2006). To ensure validity of findings, retention of panellists should be at least 70%, as per recommendations (Sumsion, 1998; Iqbal and Pipon-Young, 2009). To increase response rates and improve retention, up to two reminder emails were sent to non-responding enrolled panellists, deadlines for round completion were extended if needed, and individualised thank-you messages sent upon round completion. Some authors have noted that response

rates drop with increasing iterations, so that this Delphi was concluded after three rounds even as not all items had reached consensus (McPherson et al., 2018).

Administration of the survey and data management

Survey data were collected and managed using the Research Electronic Data Capture (REDCap) tool hosted at Imperial College London (Harris et al., 2009, 2019). While surveys included identifiable information for practical reasons, data were made unidentifiable for analysis purposes. Between-rounds feedback provided to the entire group was anonymous.

All data were stored within the Imperial College IT infrastructure, using encryption and password protection as per Imperial College recommendations. Prior to commencing the project, a Data Privacy Impact Assessment (DPIA) was conducted by an Imperial College data protection advisor, ensuring all processes were compliant with GDPR / DPA 2018.

Informed consent

Panellists partook in the consensus process in their role as researchers, based on their voluntary and informed decision. Participants were free to withdraw at any point by informing a member of the research team and without giving a reason.

Analysis

Panellist retention was analysed as the percentage of initially recruited individuals completing each Delphi round and attending a consensus workshop. Final-round ratings were analysed descriptively. Consistency of ratings between rounds 2 and 3 was analysed as intraclass correlation coefficient (ICC) for average measures, as an indicator of whether another Delphi round would have been beneficial. This analysis was added retrospectively, consistency in ratings not being the only reason to determine the number of Delphi rounds (McPherson et al., 2018).

Protocol registration and deviations

A protocol was registered and time-stamped for the complete consensus process prior to commencing Delphi recruitment, using the Open Science Framework (OSF) registry (available at https://osf.io/jmyhq/). The only notable protocol deviations were that the systematic review was presented to the Delphi panel twice, once as planned after round 1 and again after round 2 following extensive methodological feedback on the systematic review methods from the expert group, resulting in a re-analysis. This did, however, not change the main conclusions of the review. Secondly, panellists were *not* reminded of their own previous rating during subsequent Delphi rounds, because this individualisation of surveys proved unfeasible with the employed data collection platform. Finally, the live consensus meeting was planned to employ methodology of the James Lind Alliance. After noting that such a prioritising-procedure would not be appropriate, focus group methodology was used instead (described below).

Methods of interviews with people with persistent pain about the draft guidance

Between the final Delphi round and the consensus workshop, a series of qualitative interviews was performed with eight people with persistent pain (Figure 7), henceforth called 'patients' or 'patient partners'. The results informed the consensus workshops, presented at the workshops as a 5-minute video summary. Two patient partners also attended three of the four workshops.

Involvement of lay people with persistent pain was deemed desirable to ensure the real-world applicability of the guideline (Brett et al., 2014), notably regarding the acceptability of proposed methods to potential trial participants. For these interviews, the author compiled the methodological recommendations that had achieved consensus during the Delphi study into 6 vignettes. These vignettes were designed to capture how different research methods would impact the patient experience during a hypothetical clinical trial.

Scientists and patients were consulted separately prior to consensus workshops; a decision aimed at enhancing the patients' voice (Rashid et al., 2017): The Delphi was concerned with the technicalities of controlled trials and hypothetical best-practice scenarios from a largely scientific perspective. Adding the qualitative interview element after the Delphi provided patient representatives with the opportunity to judge the real-world implications of this scientific consensus. The manner of involvement (qualitative interviews rather than inclusion of patient partners into the Delphi or the conduct of a separate survey) allowed for the necessary depth of information to nuance the final guidance during the subsequent workshop discussions.

This patient involvement part of the study was reviewed and approved by the Imperial College institutional ethics review board together with the entire consensus-finding process. The interview process is reported based on the Guidance for Reporting Involvement of Patients and the Public (GRIPP) 2 (Staniszewska et al., 2017).

Patient partner eligibility

Adults living with or having experienced persistent pain (of three months or more) were recruited for patient consultation. Apart from the presence or history of persistent pain, patients required access to an internet-supporting device, and had to speak conversational English, German, French or Spanish. The sample size was not pre-specified (Boddy, 2016). Given the in-depth nature of the planned interviews and complexity of the topic, the author expected to require 5-10 participants to gain an initial understanding of the subject matter.

Patient partner recruitment

Interview participants were recruited via the Global Alliance of Partners for Pain Advocacy (GAPPA), and social media posts. The invitation was also available in an accessible audio-visual format (https://youtu.be/64IltL5m8Gw). Upon registering interest, a participant information sheet was provided. Consent forms included the explicit agreement to use anonymised direct quotes from participants. Participants were not offered reimbursement or material incentives.

Interviews and interview script

Individual one-to-one interviews were conducted via Microsoft Teams in a semi-scripted manner, reading out the dedicated vignettes, and then using open-ended questions to elicit participants' views and opinions about the vignettes. The script and vignettes are provided in Appendix 2: Consensus methods, additional content.

Interview analysis

Basic thematic analysis methodology was employed to analyse interview transcripts. This included the repeated reading of transcripts, coding of text, creation of themes, and mind-mapping of themes and supportive quotes (Braun and Clarke, 2006). Coded themes were cross-checked with the original transcripts by a third-party research assistant. Any quotes or themes emerging from the interviews were anonymised before being employed to inform subsequent consensus workshops or for published material.

Further patient consultation

Interview participants were invited to also attend the live online workshops. It was made clear the interview findings would be presented there and that it would be ensured that their voice was heard during the consensus discussion. Similarly, all patient partners who attended a workshop were invited to comment on draft guidance and could then co-author the publication.

Consensus meeting methods

The results of the Delphi and of patient interviews were discussed in a series of consensus workshops. Invited were all Delphi panellists who had completed at least two Delphi rounds, and all interviewed patient partners. Meetings were held online on four occasions to facilitate attendance across time zones and schedules. This iterative approach

also ensured smaller groups at each meeting but provided challenges for the coordination between individual meetings.

Prior to the meeting dates, the thesis author drafted a manuscript which outlined the recommendations derived from the Delphi. This draft was circulated via email several weeks prior to the workshops. Panellists were invited to notify the author of any proposed amendments. Panellists were also informed that the focus of the workshop would be to gather additional nuance on individual recommendations to make context-specific or otherwise qualified recommendations, not to change the outcomes of the Delphi study.

The workshops had several aims:

- To discuss topics highlighted by panellists in the pre-meeting phase as part of the draft manuscript review. Collecting panellists' comments prior to the meetings ensured that any controversial or complex topics were discussed at all meetings and gaps in the guidance addressed.
- To represent the views of potential trial participants, achieved by inviting patient partners to the meetings and presenting results of stakeholder interviews at the beginning of each meeting.
- To determine how the guidelines could best be presented. This included discussions with attendees around a core recommendation checklist, infographic, and the manuscript's structure.
- To elicit additional considerations from panellists by including open-ended questions and providing time for panel discussions.

At the beginning of each meeting, findings from the patient interviews were presented as a pre-recorded video (available at https://youtu.be/1-64iLPL0zM; not publicly listed). This 8-minute video ensured identical input to each meeting iteration, and subsequent discussions sought to integrate patient and stakeholder perspectives with the opinions of subject experts.

Procedures at the workshops were aligned with best-practice focus group methods, including a list of pre-defined questions and opportunities for moderated discussion amongst participants (Ritchie et al., 2013; Holloway and Galvin, 2016). The thesis author compiled the

agenda, structuring the panellists' pre-collected comments into overarching topics and writing brief descriptions of the problems or questions to be discussed. Workshops were facilitated by the author and a supervisor (JDR). These facilitators introduced the ground rules for the meeting, moderated discussions, and made sure that all agenda items were addressed. Facilitators also aimed to promote engagement of all participants, including patient partners, and ensure that discussions were not dominated by a few individuals. When conclusions were reached, the author summarised those and asked for final comments to ensure that meeting notes reflected the panel's views accurately. Arguments from earlier workshop iterations were briefly summarised when starting a new topic to ensure congruency between workshop iterations, and arguments cross-checked using an evolving agenda/analysis document. Workshops were conducted, recorded and transcribed in Microsoft Teams videoconferencing software, and analysed using elements of thematic analysis immediately after each workshop iteration (Braun and Clarke, 2006).

Ratings from the Delphi stage were respected in the workshop discussions and in draft iterations of the guidance. Instead, live discussions and subsequent manuscript reviews added nuance and explanations where required and ensured that the consensus items were compiled in a coherent manner.

Methods of guidance manuscript writing

The thesis author wrote a first draft manuscript prior to consensus meetings. In doing so, an introductory section was added, laying out background and rationale of the guidance. At the live workshops, this enabled meeting attendees to evaluate the framing and how the guidance was situated within the literature.

After the consensus workshops, additional explanatory sections were added to better contextualise the guidance and to reflect the workshops' nuanced discussions. As agreed during workshops, a checklist was compiled of recommendation items that a) were deemed applicable under most, if not all, circumstances, and b) were uncontroversial during the Delphi rating process and consensus discussions. The manuscript was then shared again with all meeting attendants, who were invited to make editorial changes.

Given the large number and extensive engagement of co-authors, two iterations of collaborative online revision were required. The thesis author negotiated co-author contributions largely through in-document comment threads, and intermittently emailed individual authors to seek clarification or propose solutions. Co-authors tracked their changes in the document, while the author monitored fidelity to previously agreed recommendations. After all co-authors had approved the final version, the manuscript was submitted to a journal for consideration of publication.

Dissemination strategy

To enhance dissemination and usefulness to readers, the guideline materials included an infographic, a designated checklist for control intervention development and implementation, and multiple supplements containing detailed methods and results of individual project stages. A graphic designer was recruited for the development of the infographic. Together with the design checklist, this infographic was hoped to make the recommendations accessible for immediate understanding and use.

Co-authors agreed to submit the manuscript for publication in a high-impact journal with a general medical readership, due to the recommendations' supposed interdisciplinary importance, including for clinicians and policymakers. During the peer-review process, the thesis author made several modifications in the presentation of the guideline. Notably, the manuscript was divided into a shortened main guideline, containing the main messages, and an 'Explanations and Elaborations' (E&E) document. The latter contained background and additional considerations for each manuscript section, and the thesis author added examples from published trials.

Now accepted for publication in *The BMJ* (Hohenschurz-Schmidt et al., 2023d), the recommendation statement will be fed back to the advocacy groups involved in the recruitment of patient partners, and the Society for Interdisciplinary Placebo Studies (SIPS). It will be submitted to EQUATOR for consideration for their reporting guidelines library (accepted by EQUATOR after initial thesis submission), and to relevant clinical trial design societies for dissemination amongst members. Importantly, the author will approach major

research funders and committees of journal editors, exploring options for the presentation of the guidelines and how they could influence policy (Calster et al., 2021). Already during development, elements of the recommendation development process were presented at relevant scientific conferences, and several further presentations are scheduled (See Appendix Table 44). Such endeavours for dissemination will continue after publication of the final guideline, and merge into presentations of a project that seeks to practically apply the guidance statement (Chapter 6).

Results

This thesis will only present the main outcomes of the individual stages of guideline development and summarise the resulting guidelines. The complete guidelines were accepted for publication on 04 March 2023 (Hohenschurz-Schmidt et al., 2023d).

Delphi results

Panellist demographics

Sixty-eight panellists were initially recruited and received the round 1 Delphi survey. Sixty-two completed the first Delphi round and provided demographic data, presented in Table 15.

Table 15: Baseline demographic data of initial panel. Data are presented for all participants completing the first Delphi round (n = 62) and thus providing baseline data.

Total sample after round 1	N = 62
Area of expertise (as per eligibility criteria)	N (%), unless otherwise indicated
Clinical trials	24 (38.7)
Placebo science	10 (16.1)
Both	28 (45.2)
Experience	
Years in clinical practice (where applicable) (M, IQR)	18.5 (10 – 25)

Years in academia (M, IQR)	19.9 (13 – 25)
Current academic position	
Postgraduate (Master's or PhD student)	7 (11.3)
Postdoctoral researcher	14 (22.6)
Associate or assistant professor	12 (19.4)
Full professor	25 (40.3)
Professor emeritus	1 (1.6)
Independent clinician researcher	3 (4.8)
Number of peer-reviewed publications	
< 5	3 (4.8)
6 to 30	16 (15.8)
>30	43 (69.4)
Clinical profession (if any)	
None	18 (29.0)
Clinical Psychology	16 (25.5)
Medicine	9 (14.5)
Physiotherapy	8 (12.9)
Chiropractic	8 (12.9)
Osteopathic manual therapy	3 (4.8)
Country of residence	
USA	17 (27.4)
Australia	6 (9.7)
UK	6 (9.7)
Denmark	5 (8.1)
Italy	5 (8.1)
Switzerland	5 (8.1)
Germany	4 (6.5)
Norway	4 (6.5)
Canada	3 (4.8)
Netherlands	3 (4.8)
France	2 (3.2)
Belgium	1 (1.6)

Sweden	1 (1.6)

Panellist retention

The third and final Delphi round was completed by 70.6% (n = 48) of the initially recruited panellists (n = 68). Experts in only placebo science were less likely to complete all rounds, compared to experts in clinical trials or experts in both placebo and trials ($\chi(3) = 9.45$, $p = 0.024$). Individual completion rates per subgroup were: Placebo researchers = 58.3% (n = 9), triallists = 80% (n = 14), experts in both = 78.6% (n = 25). Attrition was either due to unknown reasons (unsuccessful follow-up, n = 12), conflicting work commitments (n = 4), or personal circumstances (n = 2).

Rating results

Starting with 66 survey items and five open-ended questions, the first Delphi round generated an additional 44 items, evaluated in rounds 2 and 3. Out of those items, 68 reached the consensus threshold *for* recommendation (>80% of the entire sample rating agreement), and none against. Detailed round 3 ratings per item and expert group are provided as supplement to the published article (Hohenschurz-Schmidt et al., 2023d) and on https://osf.io/jmyhq/.

Average ratings between rounds 2 and 3 were highly consistent (ICC: 0.986 [0.978 - 0.991 95% CI; $F(99) = 76.892$, $p < 0.001$]), suggesting that further Delphi rounds would not have changed ratings much and that the decision to terminate the Delphi after three rounds was appropriate.

Consensus meeting results

Forty-four Delphi panellists attended a consensus workshop (64.7% of the initially recruited sample). All eight patient partners were invited to the meeting, but only two chose to attend, possibly highlighting the value of having given them space to speak openly beforehand.

At the live consensus workshops, Delphi ratings were not overruled. However, the workshop discussions led to contextualising paragraphs, changes in terminology, and the rewording of some consensus items in the final guidance manuscript. Modifications mainly concerned conceptual issues (especially the question of 'inertness' of control interventions in PPS trials) and resulting terminology (whether it was appropriate to call these control interventions 'placebo controls'). Further, after the presentation and discussion of interviews with people with persistent pain, considerations for the involvement of stakeholders and the communication with lay people were added or rephrased. The meeting discussions' main topics and content are summarised in Table 37 in Appendix 2: Consensus methods, additional content.

Results from patient partner interviews

Sample details

Eight individuals were recruited for patient consultation and individual online interviews were conducted, lasting from 20 to 40 minutes. Sample details are provided in Table 16. Recruitment ceased after eight interviews because no new themes had emerged for two subsequent interviews.

Table 16: Participant characteristics of patient consultation interviews.

Participant characteristic	Total n = 8
Age (Mean, range)	43.3 (30-63)
Gender self-identification	Female (5), Male (3)
Country of birth (n)	UK (3), Germany (2), Nigeria (1), Australia (1), Ireland (1)
Ethnic group (n)	White (4), White British (2), Celtic (1), Yoruba (1)
Country of residence (n)	UK (3), Germany (2), Nigeria & South Africa (1), Australia (1), Canada (1)
Profession	Project management (2), Healthcare (1), Healthcare management (1), Academia (1), Education (1), Humanities (1), Civil Service (1)

English language capacity (n)	Mother tongue (5), Very well (1), Conversational (2)
General health (n)	Critical (1), Poor (2), Reasonable or OK (2), Good or very good (3)
Living with chronic physical or mental health conditions (n)	Yes (7), No (1)
Pain experience (n)	Lifetime (2), >10 years (4), Intermittent (2)
Employment status (n)	Employed full-time (3), Employed part-time (2), Self-employed (1), Unemployed but active in advocacy (1), Retired (1)
Previous clinical trial experience (n)	Yes (3), No (5)

Qualitative analysis

Thematic analysis of interview transcripts resulted in numerous themes grouped into five categories: 'enablers' (to trial participation), 'barriers' (to trial participation), 'unblinding', 'blinding assessment', and 'end of trial'. Together with the researcher during interviews and afterwards during consensus meetings, interviewees also developed solutions to the identified challenges. All categories and most themes were supported by interview material from multiple participants. In Appendix 2: Consensus methods, additional content (Figure 15), an analysis mind map is presented which includes supporting quotes.

Barriers to trial participation included fear of negative effects from assignment to control interventions, the overall burden of taking part in trials, and a lack of understanding of key terms and the concept of control interventions. People felt enabled by an understanding of the placebo effect, research purposes, and when they understood the concept of control interventions. Knowing that patients were involved in the design of a trial and being able to trust the research team were also perceived as positive. Enablers and barriers to trial participation, and their potential implications for trials, are detailed in Appendix 2: Consensus methods, additional content (Table 38), alongside supporting quotes.

Further, under the category 'blinding assessment', participants agreed that they would not mind being asked to guess their group allocation. However, different preferences as to the timepoint of blinding assessment existed, with five interviewees voicing concerns that such blinding assessment during an ongoing trial might affect the care they received or

provoke doubts about their participation in the study (Quote: "*I don't wanna be constantly reminded of it because I don't want my mindset to influence how I react to the treatment […].").* The three remaining participants had no preference regarding the timepoint of assessment.

After a hypothetical RCT, all interviewees noted that they would like to find out which group they were in (category 'unblinding').

Summarised as the category 'end of trial', interviewees also discussed the desire for 'aftercare', for example, being offered the opportunity to discuss their experience with a professional or even receiving (some amount of) the tested treatment if shown to be effective. Offering additional treatment to participants in the control group could be a recruitment incentive. As demanded by many funding agencies (Bagley et al., 2016), interview participants expected to be presented with the results of the trial, ideally in a manner understandable to lay members of the public.

Involvement of patient partners in consensus workshops

Two interview participants chose to also attend a live consensus workshop. Held on four occasions, three of the workshop dates were attended by a patient partner (one person attended two workshop dates). At the workshop, one participant made no contributions in speaking despite invitations from the moderators but approved of the discussion results when consulted via email afterwards. The other person did contribute actively to the discussion at various points throughout two workshop dates.

Final guideline

The consensus workshops were followed by two rounds of online collaborative manuscript revision, and an extensive revision following reviewer feedback at *The BMJ*. The final guideline title was agreed during this process:

Recommendations for the Development, Implementation, and Reporting of **Co**ntrol Interventions in Efficacy and Mechanistic Trials of **P**hysical, **P**sychological, and **S**elf-Management Therapies - The **CoPPS** Statement.

This title reflects the changes in terminology as agreed during consensus meetings as well as the scope of the guideline, speaking of 'control interventions' (not 'sham' or 'placebo') and making explicit the exclusive relevance to efficacy and mechanistic RCTs. Title and scope omit a reference to 'pain'. Although treatment for people with pain was the focus of the expert recruitment and the Delphi study, all authors agreed that the final guideline proposed best-practice principles that were generalisable to other patient populations and trials, especially to trials exploring complex, persistent, or recurring health problems and symptoms that are amenable to expectancy effects. This includes the mental health field. Given this potential transferability, it was agreed to deemphasise the topic of pain to increase the guideline's potential audience. Nonetheless, the final manuscript acknowledges the focus on pain research during the guideline development process as a potential limitation and states that "*the CoPPS Statement must be considered in the context of the given patient population and research question*" (Hohenschurz-Schmidt et al., 2023d).

Guideline content

The final guideline consisted of a main publication, a checklist of core recommendations for the development and implementation of control interventions in efficacy and mechanistic PPS trials, an infographic, and an explanations and elaborations (E&E) document. A hybrid reporting guideline between TIDieR-Placebo (Howick et al., 2020) and additional reporting items identified during the Delphi was provided as a further supplement.

Principal considerations for the development and implementation of control interventions in efficacy and mechanistic trials of PPS interventions are presented in Table 17 (adapted from the CoPPS checklist of core recommendations; Hohenschurz-Schmidt et al., 2023d). These core recommendations are elaborated on in the CoPPS Statement and E&E document.

Table 17: Checklist of recommendations for the development and implementation of a control intervention in efficacy and mechanistic trials of physical, psychological, and self-management interventions (CoPPS Checklist). These items represent a core of best-practice recommendations that apply in any controlled efficacy or mechanistic trial of PPS interventions. The checklist complements the broader decision-making framework presented in the original publication and its supplementary Explanations & Elaborations document.

Essential recommendation item
Design and development stage
Ethical arguments for and against performing an efficacy or mechanistic trial with a control intervention should be considered, including from a trial participant's perspective.
The objectives of the control intervention in the context of the research question at hand should be defined.
A literature review of comparable control interventions and their available blinding data should be performed.
The mechanism(s) of interest of the test intervention should be defined.
The components of the test intervention thought to act on the above mechanism(s) should be specified.
The control intervention should replicate as many components of the test treatment as possible, other than the component(s) being tested.
It should be ensured that the control intervention is inert for the studied mechanism(s) and does not include the component(s) of interest.
The control intervention should be tested in a feasibility or validation step, ensuring that certain quality criteria are met (see CoPPS Statement for details of such criteria).

The risk of group contamination should be considered and mitigated if required.
Providers should be specifically trained to deliver the control intervention (if applicable).
Staff (not just treatment providers) should be educated to recognise the importance of maintaining effective blinding (if applicable).
Outcome assessors should be blinded.
The roles of treatment providers and outcome assessors should be separate (applicable if providers cannot be blinded).
Statistical analyses should be blinded.
Trial conduct stage
Providers' fidelity to intervention protocols and scripts should be monitored (if applicable).
Participants' adherence to and compliance with intervention protocols should be monitored (if applicable).
Provider expectations of benefit from the control versus the test treatment should be evaluated (if applicable).
Participants' expectations of treatment benefit should be assessed at baseline and after starting treatment sessions.
Participant blinding should be assessed (if applicable).
Reasons for participants' withdrawal from the study should be documented.

In the following, individual sections of the CoPPS Statement are briefly presented, highlighting not so much their application to trials but rather their relationship to the overall objective of this thesis, which was to enhance trial standards in the PPS field. For a more nuanced presentation of individual recommendations and examples of their application, please see the CoPPS Statement and its supplements (Hohenschurz-Schmidt et al., 2023d).

Terminology and communication (CoPPS section 1)

The CoPPS Statement aims to improve communication amongst researchers and with trial stakeholders by advocating for simplified, descriptive language around control interventions. CoPPS further recommends avoiding the terms 'sham', 'placebo control', or 'attention control'. Instead, control interventions can be described in terms of their objectives as part of a clinical trial and relation to the study treatment. For instance, it could be stated that "the control is the same as the tested treatment, except that one component has been removed. In this trial, we test the effects of this component." Clarifying that the objective is to study treatment mechanisms or efficacy may be required to avoid confusion with other comparator conditions in trials.

This recommendation was borne out of the terminological heterogeneity and lack of descriptive detail around control interventions documented by the systematic review (Hohenschurz-Schmidt et al., 2023b), and concerns from research experts and patient partners regarding potential misinterpretation of terms. Further, speaking of 'control interventions in efficacy trials' rather than 'placebo controls' was hoped to caution against the uncritical transfer of the placebo pill concept from drug trials to complex intervention studies (see CoPPS E&E document for details).

Trying to change terminology is ambitious, and unlikely to fully succeed. Nonetheless, the status quo undermines the interpretability and quality of efficacy trials, with study authors able to 'hide' behind the terms 'sham' or 'placebo' without describing the content or objectives of their control interventions. Together with endeavours to improve reporting of control interventions (TIDieR-Placebo; Howick et al., 2020), the proposed change in terminology may yet lead to improved communication. As identified in the patient involvement component of this guideline development process, clearer communication may also address some fears of potential trial participants and assist recruitment.

General considerations for the design of control interventions (CoPPS section 2)

Objectives of control interventions and the similarity principle of control design (CoPPS section 2.1.)

Balancing expectations of trial participants between study arms is the main objective of control interventions in efficacy trials. Relatedly, control interventions are used to minimise the potential confounding effect of known and unknown differences between tested treatment (other than those studied) and its comparison, influencing outcomes either directly or through the placebo effect.

Having identified these objectives during the Delphi study, the thesis author translated them into a design principle for control interventions in efficacy or mechanistic trials of PPS interventions, called 'the similarity principle':

> *Control interventions should replicate as many components of the investigated treatment as possible, apart from the components whose effect the trial aims to study.*

Based on the preparatory literature review and reaching consensus in the Delphi, the CoPPS Statement then specifies 21 contextual and treatment characteristics for which similarity ought to be considered between tested and control treatment, and four provider-related characteristics which should be matched between groups.

Blinding is widely discussed in the literature about control interventions (e.g., Boutron and Ravaud, 2011). However, blinding is mainly a means to achieve balanced expectations, as attendees clarified during consensus meetings. Whether balancing expectations is possible without blinding remains to be studied. Irrespectively, participant blinding likely also prevents behavioural differences between study arms (e.g., concomitant treatment use).

Highlighting the matching of expectations and contextual factors as the main objective of control interventions may orientate trial designers away from 'deception' and towards control interventions that are informed by the contemporary study of the placebo effect. This was also deemed desirable because of the apparent challenges to perfect participant blinding in complex intervention trials (and the difficulty in defining 'perfect blinding'). Further, the above *similarity principle* was hoped to enhance acceptability of control interventions to trial providers and participants. As the thesis author noted in a commentary on an osteopathic

trial (Hohenschurz-Schmidt et al., 2021), highly dissimilar control interventions may be illogical for study participants in the control arm, for example when they fail to address then symptomatic area. This project's meta-analysis also suggested implications for differential attrition (Hohenschurz-Schmidt et al., 2023b). Apart from patients withdrawing from the study or not complying with protocols, trial providers could feel more at ease delivering control interventions that resemble the test treatment than when enacting largely artificial protocols. This notion is further discussed in the CoPPS sections on provider training below.

Further considerations for trial and control intervention design (CoPPS section 2.2.)

Almost 30% of the PPS trials included in the systematic review had more than two study arms; 15% had an active comparator in addition to a specifically designed control intervention, 10% had no-treatment or waitlist controls, and 7% had usual care arms (Hohenschurz-Schmidt et al., 2023b). During consensus discussions for CoPPS, the value of such additional comparators was largely perceived to lie in the possibility to answer additional (comparative effectiveness or pragmatic) research questions in the same trial, and to "*help contextualise observed effect sizes between test and control interventions*" (Hohenschurz-Schmidt et al., 2023d).[6] A relevant recent example comes from the surgery field, where Beard et al.'s (2018) placebo-controlled shoulder arthroscopy trial showed no benefit of real surgery versus placebo surgery, and statistically significant but likely clinically irrelevant benefits of both interventions over no treatment. Also based on this powerful direct comparisons, the trial immediately changed clinical practice guidelines (Vandvik et al., 2019). Amongst CoPPS co-authors, however, there were contrasting opinions regarding the clinical usefulness of certain three-armed designs, and experts were alive to the risk of promoting impossibly complex or expensive trials. This paragraph was thus cautioned in that "*decisions for such [three-armed] designs depend on feasibility and the research questions under investigation*." (Hohenschurz-Schmidt et al., 2023d)

[6] No page numbers are provided for direct quotes from the CoPPS Statement because the article had not yet been typeset at the time of thesis submission.

The choice of outcome measures was the second topic presented in this 'additional considerations' section. Again, this was a topic where experts were hesitant to make definite recommendations but felt a discussion of the topic could not be avoided due to its prominence in the literature on placebo effects and control interventions. For example, Kluzek et al. (2021) argue that objective outcome measures are less susceptible to expectancy effects and thus preferable in efficacy trials. Conversely, most CoPPS panellists felt that the evidence for this statement was ambiguous, further arguing that objective outcomes were not always available or clinically relevant. CoPPS authors were also hesitant to influence which research questions readers felt they should study, so that the final CoPPS Statement holds that "*a trial's hypothesis will dictate the choice of outcome measures." and that "[if] available and appropriate, both patient-reported and more objective outcome measures can be used*." (Hohenschurz-Schmidt et al., 2023d).

Ethical considerations (CoPPS section 2.3.)

CoPPS recommends to "*[c]onsider ethical arguments for and against performing an efficacy or mechanistic trial with a control intervention, including from the perspective of a trial participant*." (Hohenschurz-Schmidt et al., 2023d). There is ample literature on the ethical basis for employing placebos in clinical trials (Miller and Crouch, 2008; Annoni, 2018), so that this section mainly referred to such existing essays. To address a common argument against efficacy trials and to reflect the current situations in PPS research and practice, CoPPS further notes that the need for high-quality efficacy data also holds in scenarios where interventions are already widely available in clinical practice or established treatment alternatives exist.

Ethical discussions of placebo-controlled trials often focus on risk/benefit considerations (e.g., Millum and Grady, 2013), the 'risk' of a placebo pill mainly being to withholding effective treatments from potentially seriously ill patients. In PPS intervention research, however, target clinical populations are rarely acutely ill nor efficacious interventions widely available. For this reason, risk/benefit considerations in CoPPS focus on the best-possible use of resources and trial participants' time.

Finally, the thesis author co-developed a discussion of the concept of 'deceit' with Dr Marco Annoni, CoPPS co-author and bioethicist from Rome

(https://www.marcoannoni.com/). This was then reviewed and added by all other co-authors during the internal manuscript review process (Hohenschurz-Schmidt et al., 2023d).

Following this section of fundamental considerations for designers of efficacy and mechanistic RCTs in PPS research, the next CoPPS section aimed at providing practical guidance for control intervention development and testing.

Control intervention development and testing (CoPPS section 3)

Conceptual development (CoPPS section 3.1.)

The conceptual development of control interventions is a focus of the CoPPS Statement, mainly to prevent researchers from thoughtlessly adopting generic control interventions from earlier trials and to facilitate conceptual clarity during the reporting of control interventions. The following recommendations encourage researchers to clearly define the objectives of their control intervention and the supposed mechanistic theory of the tested intervention:

- Clearly define the objectives of the control intervention in the context of the research question at hand.
- Define the mechanism(s) of interest of the test intervention.
- Specify the components of the test intervention thought to act on the above mechanism(s).
- Ensure that the control intervention is inert for the studied mechanism(s) and does not include the component(s) of interest.

The CoPPS E&E document discusses the common scenarios in PPS trials that treatment mechanisms are unclear or debated, or that multiple treatment components or mechanisms may be of interest.

In addition, CoPPS recommends to *"[p]erform a literature review of comparable control interventions and their available blinding data*" (Hohenschurz-Schmidt et al., 2023d) to ensure that triallists learn from previous researchers' experiences as much as possible.

Practical development and validation (CoPPS Section 3.2.)

Recommendations for practical pre-trial testing and validation of control interventions aim to prevent the implementation of non-credible or otherwise unsuitable control interventions. Although CoPPS collaborators recognised that pre-trial testing might add considerable work for researchers, this was deemed a useful employment of resources considering the alternative scenario of a potentially biased and difficult-to-interpret full-scale trial, adding to problems of research waste (Pirosca et al., 2022). At the same time, pre-trial testing is a good opportunity for impactful involvement of key stakeholders, such as members of a trial's target population and trial providers.

CoPPS guides pre-trial testing by specifying the following quality criteria, and providing references for suitable measurement instruments:

- The control intervention must be credible as a treatment (believability of the control as an intervention that will provide benefit) (Devilly and Borkovec, 2000; Hall et al., 2008).
- If blinding is an objective, the control should successfully blind participants to group allocation (similar proportions of participants in both groups believe they have received the test or control treatment or do not know) (Colagiuri, 2010).
- The control intervention should elicit a similar expectation of benefit as the test intervention (Devilly and Borkovec, 2000; Younger et al., 2012; Barth et al., 2019).

Another important concept, albeit more difficult to measure, is how 'engaging' the control intervention is for trial participants and providers. Again, this can be evaluated prior to a trial by engaging stakeholders.

Finally, this section also discusses the plausibility of information provided to participants, for example as part of the informed consent process, and its potential to modify treatment expectations.

Provider training: protocol fidelity, blinding, and equipoise (CoPPS section 3.3.)

Already in the Delphi survey, trial providers were recognised as pivotal in the successful delivery of a control intervention. For example, there was consensus that providers should receive trial-specific training according to a protocolised structure. During the live workshops, attendees added valuable practical tips from their own experience as trial designers. Tips included the use of suggested language for conversations with patients and videorecording for fidelity monitoring and on-trial supervision.

Blinding providers is usually impossible in PPS trials. This being a major difference to drug studies and common criticism of PPS trials, CoPPS extensively discusses the matter and provides advice for mitigation. For example, CoPPS recommends educating providers about the need to maintain participant blinding and comply with intervention protocols. Provider equipoise and allegiance to test interventions should also be examined and mitigated if possible. Again, the consensus workshops led to a list of practical suggestions to do so.

Blinding of other parties (CoPPS section 3.4.)

CoPPS considers the blinding of outcome assessors as essential, and recommends blinding of other involved staff. Again trying to mitigate the bias risk from non-blinded providers in PPS trials, CoPPS states that unblinded treatment providers must not act as outcome assessors.

Patient involvement and patient communication (CoPPS section 3.5.)

There was no consensus in the Delphi survey to recommend engagement of patients, practitioners, and / or public involvement groups in the development of control interventions. Concerns of panellists related mainly to the overly generic nature of the recommendation and the anticipated additional workload for researchers, as discussed during live meetings. Nonetheless, after the Delphi survey, the potential for PPI to inform control interventions was demonstrated by the interviews with patient partners, and panellists at meetings agreed that this potential benefit should be highlighted. CoPPS therefore states that *"[…] the expected benefit of involving potential participants in the development of a control intervention and*

planning of a trial is large"; the E&E document outlining potential areas for exploration during trial-specific PPI (Hohenschurz-Schmidt et al., 2023d).

Conducting a controlled trial (CoPPS section 4)

Methodological quality of control interventions is largely ensured through adequate conceptual and practical development prior to trial commencement. During a trial, CoPPS only aims to ensure that protocols are followed, and that certain quality criteria are met. These brief considerations were summarised under the topics 'Fidelity monitoring and participant adherence' (CoPPS section 4.1.), 'Measuring participant expectation and blinding effectiveness' (section 4.2.), and 'Attrition' (section 4.3.). CoPPS is, however, the first guideline to explicitly recommend the assessment of participants' blinding status and treatment expectancy in every trial. Finally, attrition may be linked to aspects of control intervention design (Hohenschurz-Schmidt et al., 2023b), so that it is recommended to collect reasons for participant withdrawal and non-adherence as much as possible, and to examine these data for potential problems with the control intervention.

Reporting a controlled trial (CoPPS section 5)

Without adequate reporting, trials remain difficult to interpret and replicate. Albeit not a reporting guideline, the CoPPS statement dedicates a section to control intervention-specific reporting. The recent publication of TIDieR-Placebo (Howick et al., 2020) also meant that control intervention reporting had made it into the EQUATOR library of reporting guidelines, but panellists felt additional detail was required. Adding to TIDieR-Placebo, this concerned the following:

Reporting of control intervention development:

- Sources and processes that informed the development of the control intervention.
- Theoretical considerations underlying the control intervention.

Reporting of control intervention content:

- A highly detailed description of the content of the control intervention (covering all components listed in Table 3 (of the CoPPS Statement) and including resemblance or differences to the test intervention).
- Provider characteristics and additional provider-related information.

Reporting of indicators of control intervention quality:

- Whether any reasons for loss to follow-up (participant attrition) were related to the control intervention.
- Results of a blinding effectiveness assessment.
- Results of a participant expectation assessment.

For ease of use, a hybrid TIDieR-Placebo/CoPPS reporting checklist was created.

Interpreting efficacy and mechanistic RCTs of PPS interventions (CoPPS section 6)

The final CoPPS section resumes the notion that PPS trials are distinct from drug trials, thus requiring methodological adaptations. Consequently, efficacy and mechanistic trials of PPS interventions may have to be interpreted differently than drug studies. Highlighting this was important to the panellists, who anticipated misinterpretations with undesirable consequences for future research agendas and uptake of interventions into clinical guidelines and healthcare policy. Because quantitative evidence for the discussed considerations is scarce, this section provides general guidance and highlights areas where results from such trials need to be interpreted with caution.

First, CoPPS underscores the complexity of most PPS interventions and warns against the assumption of linear additivity of intervention components. To prevent results from efficacy trials from devaluating entire complex interventions, CoPPS highlights that trials in which only the effects of individual components are studied "*usually cannot directly answer questions of real-world effectiveness or reflect (lack of) efficacy of the intervention as a whole.*" (Hohenschurz-Schmidt et al., 2023d)

Secondly, it may be that the real treatment mechanisms of an intervention differ from those studied in an efficacy or mechanistic trial (i.e., that the control intervention was designed based on an erroneous mechanistic theory). Inspired by the controversy around

acupuncture and certain NICE guidelines (Molassiotis, 2013; Birch et al., 2016; Musial, 2019; Birch et al., 2020), the E&E document also warns against 'comparing apples and oranges' by pooling efficacy trials with differentially conceptualised control interventions.

Thirdly, and probably the area with the most apparent lack of empirical data, it remains unclear whether effect sizes can be interpreted in a manner similar to drug trials, and how effect sizes from high-quality efficacy trials relate to an intervention's effectiveness under real-world circumstances. Drawing on hypothesis and logical reasoning, CoPPS states that "*trials that control for all but one component of a complex intervention cannot be expected to show effect sizes comparable to those of drug trials, where a single pharmacological ingredient may primarily account for the benefit over placebo. Similarly, effect sizes may not reflect the effects that can be obtained in real-world practice or the benefits over structurally different 'usual care*.'" (Hohenschurz-Schmidt et al., 2023d).

Ultimately, however, adequate interpretation requires conceptual clarity during the development of control interventions and comprehensive reporting. The CoPPS statement was developed to promote that, with the aim to facilitate high-quality trials that answer important questions of intervention efficacy or mechanism.

Dissemination results

Apart from the successful publication of the CoPPS Statement in *The BMJ*, further interim outcomes of the dissemination strategy include the presentation of the parts of the project at scientific conferences, such as the Australian Trials Methodology conference (virtual lecture, 07 December 2021), the annual conference of the Society for Back Pain Research (plenary lecture, 30 May 2022, New Investigator's Award), the world congress of the International Association for the Study of Pain (poster, 21 September 2022), and the national conference of the German Pain Research Association (symposium, 20 October 2022) (Appendix 6: Academic development and awards).

Efficacy Section: Discussion and Limitations

Strengths and limitations of the CoPPS statement

If implemented, the CoPPS statement could remedy several systematic shortcomings of efficacy and mechanistic trials in PPS research. In pain trials, only 2% report involving patients in the development of the control intervention, only 12% report validation or pre-trial testing, 23% state that providers were trained to deliver control interventions, less than a third report any theoretical consideration regarding the control intervention, there is insufficient reporting of blinding status, and less than a quarter of trial reports provide information on the effectiveness of participant blinding (Hohenschurz-Schmidt et al., 2023b). The CoPPS statement is a comprehensive guide to enhanced reporting, but importantly it is a conceptually thorough and evidence-based guideline for the development of control interventions in efficacy and mechanistic trials of PPS interventions, and their rigorous implementation and reporting.

The strengths of the Delphi technique are that it can capture the views held by experts in a given field and explore the scope for consensus, while reducing authority bias and effects of social desirability. Further, unpublished but valuable experience of practitioners (trial researchers in this case) can be drawn on by Delphi techniques (Jünger et al., 2017; McPherson et al., 2018). However, the Delphi technique can only provide data based on supposed experts' opinions. Whilst these opinions are undoubtedly influenced by quantitative evidence, they are also necessarily limited and liable to personal preference and bias. As in this guideline development process, these effects may be partly mitigated by presenting experts with current empirical evidence as part of the Delphi, such as a systematic review.

The main limitations of the present guideline development process are its geographical and cultural focus on Europe, North America, and Australia, its interventional bias towards Western therapies, and the design focus on pain research. Despite efforts to produce generalisable principles, it will have to be investigated how these guidelines apply to other cultural, clinical, and healthcare system contexts.

To answer the present research question – which standards should a control intervention in efficacy and mechanistic trials fulfil – expert opinion and a systematic overview of current practice were valuable contributors, but further research is required. Mainly, the recommended features of control interventions need to be assessed using 'hard' quality criteria, particularly assessing whether they lead to the successful blinding of trial participants and balanced expectations between study arms. Further, the methods recommended in the CoPPS statement need to be practically feasible, often within normal clinical settings. They need to be cost-effective, ethical, and acceptable to participants, providers, and researchers. Such considerations featured strongly in the decision-making of Delphi panellists during the consensus workshop and were reinforced by patient interviews about the draft guidance and patient partners' involvement in consensus meetings. Following the publication of the guidelines, further research needs to investigate these questions of blinding effectiveness, balanced expectations, and usability and acceptability. The development of a protocol and control intervention for a feasibility trial presented in chapter 6, provides initial insights on the applicability and usability in the context of an efficacy-focused RCT of a multimodal, manual therapy-based intervention. Other projects examining the practical usability of the guideline from a researcher perspective are ongoing, namely two MSc projects at the University of Health Friborg (Haute école de santé Fribourg, HEdS-FR) in Switzerland (supervised by the thesis author). Once the CoPPS statement is more widely used, independent reports will hopefully follow.

Overall, the development of the CoPPS statement incorporated all best-practice methods for guideline development (Moher et al., 2010b), except for practical testing (Brouwers et al., 2010), exceeding the rigour of comparable projects by some margin (e.g., Beard et al., 2020). A mixed-methods approach allowed for the nuanced discussion of a complex problem in a manner that hopefully accurately reflects the practical constraints of research practice and the lived experience of people participating in trials, while addressing current methodological shortcomings in RCTs.

Barriers to implementation for the CoPPS statement

Methodological change in clinical trial research is slow at best. For example, the evidence is mixed whether reporting has improved at all after the publication of milestone

reporting guidelines (Ladd et al., 2010; Alvarez et al., 2020; Pirosca et al., 2022). This is despite publications like the 2010 CONSORT statement being amongst the most-cited medical articles (https://www.scopus.com/record/pubmetrics.uri?eid=2-s2.0-77950189829). Risk of bias has likely reduced in some areas, although over long timeframes and with blinding remaining an unaddressed challenge in many fields (Reveiz et al., 2015; Saltaji et al., 2017; Long et al., 2020). To improve research methodology, recent commentaries have called for more top-down efforts, appealing to journals, universities, funders and governments to unite existing methodological initiatives and to "*put methodology first*" (p. 219, Calster et al., 2021). Presenting this work to these gatekeepers is thus an important part of the dissemination strategy for the CoPPS Statement.

Evaluating the impact and uptake of the CoPPS Statement will form part of its evaluation (Brouwers et al., 2010). A priori criteria to measure this may include the rate of authors reporting the guideline's adoption in 5- or 10-years' time, compliance rates with the CoPPS reporting guideline, the proportion of highly similar control interventions compared to the present systematic review, as well as comparison of other design features (Hohenschurz-Schmidt et al., 2023b), and finally, the impact of the new guideline on the outcomes reported in PPS trials. Further, feedback from researchers, journal editors, and other stakeholders on the usefulness and applicability of the CoPPS Statement should be collected systematically, including any barriers to implementation or areas for improvement. As discussed by Calster et al., (2019), changes in funding requirements or regulatory guidance may also be required and could be assessed.

Ideally, CoPPS enhances the quality and reporting of efficacy and mechanistic trials in PPS. Even if this should succeed, however, there remain structural challenges for the appropriate consideration of complex intervention trials for policy and clinical guidelines. These barriers may largely stem from the question whether nonpharmacological trials are comparable to drug trials. On one hand, some authors have justifiably argued for the application of the same research standards for the evaluation of non-drug studies as for drug studies (e.g., Cramer, 2022). On the other, there are undeniable differences between these intervention classes, leading to differential challenges in trials and requiring specific methods for bias prevention (as detailed in the CoPPS statement (Hohenschurz-Schmidt et al., 2023d)).

Common risk-of-bias assessment tools for evidence synthesis automatically downgrade trials when providers are aware of participants' group allocation (Yamato et al., 2017) and they do not allow for the mitigating effects of methods other than provider blinding (Sterne et al., 2019). If, for example, trial designers follow CoPPS recommendations, demonstrating to have achieved provider equipoise and that no protocol deviations occurred, the risk of bias should be regarded as low also in the absence of provider blinding. A similar scenario is imaginable in the absence of participant blinding but with balanced expectations between groups. While the current Cochrane risk-of-bias tool (RoB-2) allows for such scenarios in theory, prominent researchers consider it a possibility that the tool is applied "*over-zealously*" (Pirosca and Treweek, 2022). Such inappropriate application of RoB tools may contribute to higher risk-of-bias evaluations in non-drug studies compared to drug studies (as, e.g., found by Pirosca et al., 2022). Potentially, RoB tools require adaptation for non-drug studies or at least specific annotation to guide systematic reviewers. CoPPS' respective section on the interpretation of efficacy and mechanistic trials with high-similarity control interventions is another effort to recognise the potential importance of different characteristics of PPS interventions.

The broader context of the present work on efficacy and mechanistic trials

Controversies about the importance of blinding

An underlying premise of this work on improving efficacy trials in PPS is that participant blinding matters; Control interventions ideally balance expectancy and learning effects between study groups and limit response bias through blinding. While our own meta-analytical study supports the notion that such matching is indeed important (because 'un-matched' controls bias study outcomes), there is conflicting evidence regarding the importance of blinding from other meta-analyses. On one hand, a large meta-epidemiological study of 2,443 trials from any area of health research attested a 13% exaggeration of intervention effect estimates in trials with high or unclear risk of bias for blinding, which included blinding status of providers, participants, and outcome assessors (Savovic et al., 2018). This finding appears aligned with earlier analyses from the same group (Savovic et al., 2012) and others who assessed the impact of participant blinding (Hróbjartsson et al., 2014). Conversely, Moustgaard et al. (2020) found no evidence that treatment effects differ

between trials with or without blinded participants. These authors combined 1,153 trials from various fields: pharmacological (67%), psychological (12%), surgical (2%), and 'other' (19%). Similarly, and focusing on surgery trials (n = 100), Karjalainen et al. (2022) were surprised (personal communication) to find little difference in effect sizes between placebo-controlled trials and such with open-label nonsurgical controls, concluding that there was no "*evidence supporting the hypothesis that surgery is associated with a large placebo effect*." (p. 6). Instead, they argued that most benefits in placebo groups were due to other nonspecific effects such as symptom regression and natural disease course.

Caution is required when transferring findings from the above works to the field of PPS, due to several noteworthy limitations and differences: First, Savovic et al. (2018, 2012) as well as Moustgaard et al. (2020) assessed blinding across a very heterogeneous (albeit large) group of interventions, not providing any subgroup analyses. This would appear necessary given different placebo effect magnitudes observed between interventions and conditions, with pain studies being particularly susceptible (Hróbjartsson and Gøtzsche, 2010, p. 201). Besides, the Moustgaard meta-epidemiological study has received considerable methodological criticism since. This criticism includes relying on very small samples in many individual comparisons, large heterogeneity in the design, study population, and settings of primary studies, and using primary trials without direct medical interventions (Tack, 2021). Criticism can also be levered against the Karjalainen study: These authors assumed the difference between 'nonsurgical' comparator treatments and surgical placebos to indicate the 'placebo effect'; However, the included comparators were other active interventions such as exercise or education programmes which exposed the trial participants to both active therapeutic elements and, potentially, different contextual cues for the placebo effect. Because they did not focus on true no-treatment groups, their analysis of the placebo effect in surgery is likely confounded by these processes in their comparator groups. Another important concern is that surgery is often a one-off intervention with no repeated exposure to contextual factors. Also, short-term outcomes are rarely reported and the median timepoint of extracted outcome data was 12 months (IQR 3-12). By then, any placebo effect may have been 'washed-out' (Karjalainen et al., 2022) (as happens elsewhere; Koog and Jung, 2013), despite potentially large initial expectations for benefit. These aspects likely limit the transferability of Karjalainen's findings to the PPS field; also representing a difference to our

analysis where we provided a nuanced structural evaluation of control interventions, confirmed findings across various PPS subgroups, and only looked at short-term outcomes.[7]

What can be learnt from Karjalainen et al. (2022) is certainly the meta-epidemiological approach to compare ('placebo'-controlled) efficacy trials with other comparator interventions. Ideally, however, only true no-treatment (probably more aptly named 'nonexposure') groups are used to elucidate the magnitude of placebo effects and ideally these placebo and nonexposure data stem from the same trials to avoid the unreliability of indirect comparisons (Hróbjartsson and Gøtzsche, 2010) (p. 20). In the Karjalainen review, only two trials had such a third arm and previous attempts from manual therapy were inconclusive (Lavazza et al., 2021) (likely due to overly restrictive inclusion criteria resulting in only four included studies for this analysis). The thesis author has since identified a subgroup of such potentially relevant three-armed RCT from the PPS review (Hohenschurz-Schmidt et al., 2023b); a search update and a project to elucidate the magnitude of the placebo effect in these trials are ongoing.

Although widely-read editorials question the importance of blinding (Anand, et al., 2020; Drucker and Chan, 2020) (also see Yonis et al., 2020), methodological concerns about this conclusion exist (Tack, 2021) and blinding importance likely depends on several preliminary considerations (Anand, et al., 2020). Such considerations include whether blinding is indeed needed for a scientifically sound result (Will the intervention have a placebo effect which needs to be separated from its true effect?); how likely it is that patients or clinicians will behave differently if they know the intervention and if this would bias the results; the potential harms to patients of using blinding in relation to the anticipated social value of the study; and the financial cost of blinding which may compromise spending on other methodological aspects of trial integrity. These criteria suggest that blinding may be important in PPS studies for pain after all, because of the susceptibility of pain to placebo effects, and the likely behaviour change of patients when receiving an intervention perceived

[7] Arguably, this methodological difference regarding timepoints of assessments is appropriate as it represent clinical reality: Patients expect ongoing relief when opting for invasive procedures such as arthroplasties (McHugh and Luker, 2012; Nyvang et al., 2016), while short-term benefit, outcomes related to understanding and management rather than pain relief, and repeated consultations appear more acceptable to patients when seeking out, for example, manual interventions for pain (Plank et al., 2021) or physiotherapy (Bernhardsson et al., 2017; Calner et al., 2021).

as less desirable (See Table 39 in Appendix 2: Consensus methods, additional content for a comprehensive discussion of these criteria in the PPS context).

Importantly, the relevance of blinding is interwoven with the behaviour of the placebo effect. Since placebo effects appear largely unpredictable (Vase et al., 2015b), it can be argued that standardisation is the most reliable way forward: Best-practice approaches need to be agreed upon and followed to reduce methodological heterogeneity. At the same time, best-practice recommendations need to be based on what is known about the behaviour of the placebo effect and trials' susceptibility to bias *in individual fields of clinical research*, as these factors are unlikely to translate (as discussed above). These steps might increase comparability between studies *within research fields* but not necessarily across therapeutic modalities. The impactions of this are stark: evidence synthesis for clinical guidelines or policy decision-making needs to consider these field-specific challenges. At this point, little can be said about, for example, the comparability of effect sizes in 'sham'-controlled efficacy trials of surgery versus physical or manual therapy. At the very least, as more and more best-practice frameworks for efficacy trials are published (Braithwaite et al., 2020a; Beard et al., 2020; Hohenschurz-Schmidt et al., 2023d), stakeholders have access to field-specific quality markers that should inform decision-making.

A further factor that may have a bearing on the importance of effective participant blinding but went largely unexplored in this thesis, is the nature of the outcome measures (Ford and Norrie, 2016). Here, distinctions are commonly drawn between 'objective' and 'subjective', or 'observer-reported' and 'patient-reported' measures. As per Hróbjartsson and Gøtzsche (2010), the latter are more susceptible to placebo effects, albeit many other factors contribute (Vase et al., 2015b).

Further challenges to blinding are participant recruitment (patients may be hesitant to enrol when they might receive a placebo), participant retention, efforts and expenses related to the design of control interventions, motivation of participants and study staff to break blinding, and even risks to patient safety (which, according to Anand et al., 2020, includes "*muscular problems from sham physiotherapy*", p. 2) (Anand, et al., 2020). The CoPPS Statement addresses most of these concerns (Hohenschurz-Schmidt et al., 2023d). Other arguments can be refuted or need to be contextualised (Table 39 in Appendix 2: Consensus

methods, additional content): Costs attached to control intervention design are necessary to not produce trial findings that are biased or uninterpretable, and direct risks from PPS control interventions are likely harmless – as they are in the tested interventions (Carnes et al., 2010; Niemeijer et al., 2020). Omission of important medical interventions is also unlikely, since PPS interventions are rarely used independently for severe pathology and because participant blinding does not interfere with medical triaging.

A major criticism of blinding in trials, put forward by Anand et al. (2020), is its potential to reduce trial 'pragmatism' by creating a scenario that differs from routine clinical practice. Such criticism, however, fails to recognise that a) efficacy-related information may be an important prerequisite for more pragmatic research designs (Keefe et al., 2022) and b) the intention of pragmatic trials is to 'inform clinical decision-making' with high-quality methods (Ford and Norrie, 2016). Therefore, any such criticism requires contextualising and must not omit important caveats regarding potential pitfalls of unblinded effectiveness trials. As will be discussed in the next thesis section, pragmatism in RCTs is not a dichotomous decision and several methods for bias control can be considered.

Uncertainty about the comparability of effects in drug trials and non-drug trials

Pharmacological and nonpharmacological efficacy trials are different due to the differences in the studied interventions. While the present work is hoped to contribute to the strengthening of research methods in the nonpharmacological field, one fundamental question remains to be answered: If these trials are so different from drug trials and if CoPPS promotes high-similarity control interventions for complex therapies, can we compare effect sizes across pharmacological and nonpharmacological studies? Again, there is little empirical evidence directly addressing this problem. At the very least, there is the concept of minimal clinically important differences (MCIDs) to fall back on, but these may not be reached if the hypothesised therapeutically active component is not very powerful in itself, if only the effect of a single component of a multimodal intervention is studied in an efficacy trial, or if many of the control intervention components are not 'inert'. Comparisons against well-designed high-similarity control interventions could be made more meaningful be the inclusion of a third, non-exposure study arm, also facilitating discussions about the extent to which various stakeholders are willing to accept that treatment effects are due to contextual factors. On the

other hand, even if effect sizes remain below MICDs, perhaps one needs to remember that appraising and synthesising trials should be informed by more than effect sizes: risk-of-bias assessments, economic data, and ideally effects on additional outcomes measures such as quality of life or function all ought to contribute to such assessments. In the end, the CoPPS statement has certainly potential to help advance the mechanistic study of PPS interventions. To interpret their efficacy remains challenging.

Questioning the relative importance of efficacy research compared to pragmatic trials

Apart from methodological considerations, the CoPPS statement may impact PPS research agendas more broadly. Arguably, CoPPS could promote efficacy and mechanistic trials at the expense of comparative effectiveness or pragmatic trials.

In this context, it needs to be noted that there is considerable debate around the value of efficacy versus comparative effectiveness trials, with some authors arguing to focus on the latter in nonpharmacological intervention research (prominently, for example, the EU project CAMbrella's consensus statement by Fischer et al., 2014) (also see Fønnebø et al., 2007; Walach and Loef, 2015). Individual researchers of the CoPPS group have also conducted and advocated primarily for pragmatic and comparative effectiveness trials, but the need for efficacy research was undebated. A dual approach, using more explanatory as opposed to more pragmatic attitudes to trial design to study a broad range of research questions, is aligned with current research strategies in the pain field (Becker et al., 2018; Bushnell et al., 2021). It was also agreed at consensus workshops that CoPPS should not attempt to resolve the debate of which type of research question (and thus trial design) should be prioritised, and instead focus on improving efficacy and mechanistic trial methods where authors decided that such trials are ethically justified and required by their respective research questions. Therefore, the CoPPS statement largely avoids the question of how efficacy trials relate to effectiveness trials; for example, it is not discussed whether efficacy of intervention components must be demonstrated before pragmatic trials are conducted (in other words, whether a 'pipeline'-type approach to PPS intervention testing should be implemented as during drug development). Instead, cautious formulations were used, such as: "*positive signs from an efficacy trial […] should increase end-users' confidence in an intervention under real-world conditions*" (Hohenschurz-Schmidt et al., 2023d), or generic statements made such as

"*well-designed pragmatic trials […] are useful for further elucidating questions of comparative and real-world effectiveness*.". This approach was chosen for several reasons. First, to reflect the fact that many PPS interventions are widely used despite ambivalent efficacy evidence, making a 'pipeline' approach unrealistic; Second, due to uncertain evidence about how efficacy and effectiveness trials relate to one another in terms of effect sizes (Mbuagbaw and Aves, 2022); and thirdly, to not devaluate comparative effectiveness or pragmatic trials. Nonetheless, the status quo of PPS intervention research is thoroughly criticised by this PhD work. High-quality control intervention design and participant blinding *are* possible in all areas of PPS research. The systematic review demonstrated this with several examples of high-quality trials and CoPPS provided a comprehensive guideline for the entire field. Therefore, the argument that 'efficacy trials are not possible' is refuted, and a lack of high-quality efficacy and mechanistic RCTs in many areas of PPS research problematised.

The common argument that only pragmatic trials produce 'relevant' insights (Fischer et al., 2014) also ignores two important considerations: First, trial designs need to be chosen according to pertinent research questions, there being a need for several types of studies for each intervention (Jonas, 2005; Skivington et al., 2021). And secondly, the simple replication of 'real-world practice' without consideration of internal trial validity may hold the risk of producing pragmatic trials that are uninterpretable and thus irrelevant. To prevent this from happening, the next section of this thesis focuses on pragmatic trials.

Effectiveness Section: Pragmatic and Comparative Effectiveness Trials

Section Summary

Pragmatic trials have been defined as trials that address research questions directly relevant for clinical or health policy decision-making (Schwartz and Lellouch, 1967; Ford and Norrie, 2016). In studying these questions, pragmatic trials are often embedded within or resemble clinical practice, and many self-declared pragmatic trials compare two or more available treatment options. In pain research, pragmatic trials are increasingly common because they are thought to address clinical problems more adequately than traditional efficacy trials; for example, by reflecting the multimorbidity of people with persistent pain or the variability of treatment delivery and adherence in clinical practice. By means of a systematic review of methods, this chapter describes how the concept of pragmatism is implemented in trials of pain therapies, identifying common themes and challenges. Through an expert and stakeholder collaboration project under the auspices of the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT), recommendations were made on how to ideally apply the concept to trial design and balance it with requirements of internal validity. The latter consideration represents an important addition to current debate, since external validity ambitions of pragmatic trials are not usually discussed in the context of how they might undermine a trial's fundamental ability to produce valid and reliable results. In a second IMMPACT statement, considerations specific to pragmatic trials in pain research are presented.

Section background

Preliminary considerations of concepts and terminology

The most common definition of pragmatic trials focuses on a trial's research question, stating that pragmatic trials address research questions directly relevant for clinical or health policy decision-making (Schwartz and Lellouch, 1967; Ford and Norrie, 2016). In practice, however, the term 'pragmatic' is also used to relate to individual features of trial design, such as the 'pragmatic' delivery of a treatment (Taljaard et al., 2022). Here, 'pragmatic' indicates that something is done in a manner similar to routine clinical practice or is embedded within it. The PRECIS-2 tool (pragmatic–explanatory continuum indicator summary) allows the rating of nine design domains from 'explanatory' to 'pragmatic' (Thorpe et al., 2009; Loudon et al., 2015b), illustrating that a trial's ability to inform clinical decision-making can depend on design choices across a range of RCT features. PRECIS also shows that the labelling of a trial as either pragmatic or explanatory likely misses the complexity of trial design: A trial can be very pragmatic for some aspects but highly controlled for others (i.e., explanatory) (Treweek and Zwarenstein, 2009). Unfortunately, the strong focus of PRECIS (and its successor PRECIS-2) and other available tools (Wieland et al., 2017; Zuidgeest et al., 2017) on how a trial is designed also obscures an important underlying consideration, namely which research question a trial set out to answer (Nicholls et al., 2020). The fact that PRECIS-2 is increasingly used to retrospectively assess how 'pragmatic' a trial is (Dal-Ré et al., 2020; Zwarenstein et al., 2020) reinforces the notion that pragmatic trial methods are inherently desirable, rather than allowing for a discussion of the trial's research question, how the question informed design, and whether the design allowed the question to be answered in a valid manner (Nicholls et al., 2020). This latter conceptualisation is also better aligned with the originators of pragmatic trials; Schwartz and Lellouch (1967) "*characterised pragmatism as an attitude to trial design rather than a characteristic of the trial itself*" (p. 2, Treweek and Zwarenstein, 2009) (also see Zwarenstein et al., 2020).

Additional terms relevant to pragmatic trials are 'effectiveness' and 'generalisability'. As outlined earlier (also see the glossary in Table 1), 'effectiveness' refers to the evidence

produced by studying intervention effects under usual circumstances of medical practice (Wieland et al., 2017), asking the question "*Does it work in practice?*" (p. 652, Haynes, 1999; Skivington et al., 2021). With regards to the above-described lack of clarity in the term 'pragmatism', effectiveness clearly relates to the intention of a trial, not individual design features. Relatedly, 'comparative effectiveness' research has been variably defined as "*compar[ing] the effects of a number of treatments in current use on clinical outcomes in order to guide decision making*" (p. 2152, Fiore and Lavori, 2016) or as "*a strategy that focuses on the practical comparison of two or more health interventions to discern what works best for which patients and populations*" (p. xiii, Institute of Medicine, 2009)[8]. In both definitions, the pragmatic intention to inform clinical practice resonates; and indeed, some authors contend that pragmatic trials are often comparative effectiveness trials since deciding between treatment alternatives is a common clinical problem (Chalkidou et al., 2012; Dal-Ré et al., 2020).

Finally, 'generalisability' refers to whether trial results apply to populations, settings, and contexts outside those of a respective trial. More specifically, Rothwell (2005) speaks of results that are "*relevant to a definable group of patients in a particular clinical setting*" (p. 82), or "*reasonably likely to be replicated when applied to a definable group of patients in a particular clinical setting*." (p. 92, Rothwell, 2010). The terms 'generalisability', 'external validity' and 'applicability' are broadly synonymous for the extent to which trial results can be extrapolated in this manner (Rothwell, 2005a, 2010)[9].

Producing generalisable results is often a declared aim of pragmatic trials. Surprisingly, in the literature there is very little discussion of the notion that, on examination, generalisability can only ever be an educated guess: Trial results only apply with high certainty to the circumstances of a given trial (and even here problems with methods and bias may put in question their validity) or to circumstances very similar to those of a trial. Typical pragmatic trials broaden the range of populations, settings, and circumstances to which trial results may apply by increasing the within-trial heterogeneity of these features. Nonetheless, when

[8] The more commonly cited and more extensive Institute of Medicine definition for Comparative Effectiveness Research is: "*the generation and synthesis of evidence that compares the benefits and harms of alternative methods to prevent, diagnose, treat, and monitor a clinical condition, or to improve the delivery of care. The purpose of CER is to assist consumers, clinicians, purchasers, and policy makers to make informed decisions that will improve health care at both the individual and population levels*." (p.203) (Sox and Greenfield, 2009)

[9] However, see Dekkers et al. (2010) for a discussion of potential nuances between these terms.

judging the relevance of trial results to a clinical scenario *outside* the trial, inferential errors likely ensue from unknown differences between trial and external settings and conditions (Dekkers et al., 2010). Practically, Rothwell notes that making judgements about generalisability "*requires clinical rather than statistical expertise and usually depends on a detailed understanding of the particular clinical condition under study and its management in routine clinical practice*" (p. 95, Rothwell, 2010). A single trial can also impossibly be representative of all or even just multiple external settings, questioning the notion of 'high generalisability' itself (Pawson, 2019). Further, the term 'applicability' probably better encapsulates the necessary assumptions and limitations of such judgements (Treweek and Zwarenstein, 2009; Dekkers et al., 2010). PRECIS-2 helps judge how applicable trial results may be to a population outside of a trial by consideration of several relevant domains (Loudon et al., 2015a), but the above limitations remain (Pawson, 2019). Other, more sophisticated methods are available that draw on comparative data from epidemiological studies to judge the representativeness of trial features or examine how variations in relevant trial parameters may affect outcomes (Greenhouse et al., 2008; Treweek and Zwarenstein, 2009). Finally, retrospective assessments of external validity depend crucially on the completeness of reporting in study publications; which has been found wanting (Bornhöft et al., 2006; Boutron et al., 2008; Rothwell et al., 2009; Dworkin et al., 2012; Loudon et al., 2015a; Boutron et al., 2017). To promote the reporting of relevant information, an extension of the CONSORT statement exists for pragmatic trials (Zwarenstein et al., 2008).

While supporting trial designers to develop studies with increased chances of being widely applicable, the PRECIS-2 statement contains the following caveat: "*The PRECIS-2 tool is focused exclusively on the issue of applicability, and readers should use other tools to assess the internal validity or any other aspect of their design choices*." (p. 3, Loudon et al., 2015a). This section will return to the relationship between applicability and internal validity repeatedly.

Having outlined conceptual and terminological challenges with regards to so-called pragmatic trials, the next paragraphs situate this thesis in current developments of pragmatic trials in the pain field and identify research needs.

The need for more pragmatic trials

Throughout medicine, explanatory RCTs are increasingly criticised for the supposedly limited generalisability of their results, deemed due to key aspects of their design (Rothwell, 2005b). Common features that make RCTs less representative of normal clinical practice are excessive exclusion criteria for participants (Rothwell, 2005b; Spall et al., 2007; Barnett et al., 2012; Rowbotham et al., 2013; Whitty et al., 2020) and thus trial populations which differ from the general patient population (Masoudi et al., 2003; Movsas et al., 2007), and unrealistic requirements for treatment compliance (McCarthy et al., 2012; Coleman et al., 2013; Kipping et al., 2014). Maybe most importantly, what matters to a patient may not have been assessed in an RCT: to be relevant for clinical decision-making, statistical changes in outcome measures need to be reflected in clinically noticeable *and* personally valuable changes in symptoms, quality of life, or disease risk (Dworkin et al., 2005; Taylor et al., 2016; Coon and Cook, 2018; Weinfurt, 2019). Further, the time horizon of patients' decisions is rarely encapsulated by common RCT follow-up periods that are usually six months or less (Dworkin et al., 2010). Despite often limited external validity, RCTs form the basis of most health policies, medicines regulatory approval, and treatment guidelines (Persaud and Mamdani, 2006; Irving et al., 2017; Frieden, 2017; Gastaldon et al., 2019), as they allow for cause-and-effect evaluations and not least because they provide a situation in which most factors apart from the intervention are controlled as well as possible.

Pragmatic trials in pain

The above concerns about efficacy trials also exist in pain research. It has further been argued that the limited usefulness of many analgesic drugs in clinical practice may be due partly to the absence of pragmatic trials from the drug approval process (Watson, 2007). For example, limiting the applicability of many trials to clinical practice is the common exclusion of people with co-morbidity, polypharmacy, psychiatric illness, or substance use disorders (Whitty et al., 2020). Especially in the context of opioid research and research on people with persistent pain, the resulting trial populations differ from clinical target populations (Bruehl et al., 2013; Rowbotham et al., 2013). Similarly, drugs for neuropathic pain are most

commonly studied in relatively stable diseases such as painful diabetic polyneuropathy or post-herpetic neuralgia, but findings are then generalised to 'neuropathic pain' at large (Watson, 2007; Finnerup et al., 2015). Finally, explanatory trials provide little data relevant to tailoring of treatments or multidisciplinary approaches to pain management (Bruehl et al., 2013); both of which are key ambitions of current pain research agendas and national policy (Pizzo et al., 2011; Australian Government Department of Health, 2021).

Pragmatic trials in PPS interventions

For nonpharmacological trial research, pragmatic and comparative effectiveness research holds the promise of circumventing some challenges of efficacy RCTs, notably blinding (Boutron et al., 2004; Mason et al., 2002; Williams et al., 2021). Some prominent initiatives and authors in the CAM world (complementary and alternative medicine) even argue that pragmatic and comparative effectiveness trials are the only types of trials which can provide "*clinically meaningful*" insights (p. e10), for example because they consider the attempt to distinguish specific from unspecific effects in efficacy trials to be fundamentally misguided (Fischer et al., 2014).

Most contemporary authors would likely subscribe to integrative (Jonas, 2005), circular (Walach et al., 2006; Skivington et al., 2021), or mixed-methods (Fischer et al., 2012) approaches for the evaluation of nonpharmacological research, and pragmatic trials play an important role in all of these research models. Nonetheless, methodological challenges of pragmatic trials are rarely discussed within the PPS field: Critical methods-oriented publications are only recently beginning to emerge (e.g., Keefe et al., 2022). As in all research endeavours, scientifically and conceptually flawed trials produce 'uninformative' results (Zarin et al., 2019), and thus research waste (Yordanov et al., 2015; Buchbinder et al., 2020). The risk for nonpharmacological pain research is thus that focussing on pragmatic trials without giving thought to methodological rigour may further undermine the evidence base of PPS interventions and waste resources.

If well-conducted, however, pragmatic trials may indeed contribute valuable insights regarding the role of PPS interventions, either as stand-alone therapies in clinical practice or part of integrated models of pain care, and do justice to their complexity. A US Veterans Health Administration's policy conference has identified pertinent research questions to

advance nonpharmacological management of chronic musculoskeletal pain (Becker et al., 2018); Well-conducted pragmatic trials may address many of these questions, including the combination or sequencing of treatment components and modalities, the evaluation of optimal dosages and delivery modes, the identification of responder characteristics, the study of long-term interventions and effects, as well as including people with comorbid mental health or complex medical conditions. Becker et al. (2018) also introduce the idea that trials for these research questions may employ innovative designs that go beyond the traditional two-armed parallel-group RCT, such as adaptive or sequential multiple assignment randomised designs.

In sum, there is considerable interest in pragmatic trials in pain research in general, and in nonpharmacological intervention research in particular. However, few initiatives have balanced the promises of these trials with the promotion of methodological rigour to ensure informative results.

With this in mind, the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) dedicated its 2020 meeting to pragmatic trials in pain research. Through the long-standing association of Prof Rice with this group, IMMPACT's director Prof Robert H Dworkin commissioned the thesis author to lead a preparatory systematic review of current pragmatic trials in pain research, the meeting itself, and any subsequent publications.

Chapter 4: Systematic Review of Methods in Pragmatic and Comparative Effectiveness Trials

Introduction

As described in the introduction of the section, pragmatic trial designs have been proposed as a possible bridge between RCTs and clinical practice (Ford and Norrie, 2016), including in pain research (Gordon et al., 2020a). At the same time, a 'good' pragmatic trial is largely conceptualised as one that enhances generalisability by replicating normal clinical practice (Loudon et al., 2015a). By concept, for example as advanced by the PRECIS tool (Thorpe et al., 2009; Loudon et al., 2015a), pragmatic trials should be large in scale, embedded into ongoing clinical practice, and may frequently investigate complex interventions (for which explanatory trials are arguably more difficult to conduct [Boutron et al., 2004; Williams et al., 2021]). Whilst patients are still randomly assigned to treatment groups, they are likely rarely blinded to their allocation. Also, treatment protocols should be flexible, for example allowing clinicians to adjust therapy to individual patients, and adherence requirements for patients non-existent. Outcome measures should be more likely to reflect what is important in clinical practice, focusing on disability and function, risk-benefit analyses or even cost-effectiveness rather than average pain scores (Roland and Torgerson, 1998; Moore et al., 2010; Chalkidou et al., 2012; Williams et al., 2015; Ford and Norrie, 2016). At the extreme end of the pragmatic-explanatory spectrum, pragmatic trials assess outcome data sampled routinely in clinical practice and alter routine care minimally or not at all (Loudon et al., 2015a). Whether this ideal is reflected in current practice of self-declared pragmatic pain trials, was, however, unclear at the time of project planning. It can also be argued that this ideal conflicts with the practical requirements of conducting RCTs and may have undesirable effects on the validity and interpretability of trial results (Ford and Norrie, 2016; Keefe et al., 2022).

To understand the current practice of trial design, conduct and reporting in the field of pragmatic trials of pain treatments, a systematic review of methods was conducted. This

review then informed a consensus process that aimed to produce best-practice considerations for pragmatic trials of pain.

Specifically, the objectives of the systematic review were to:

- Survey the number of randomised controlled trials that are called ‘pragmatic’ or ‘comparative effectiveness’ and investigate any therapy aimed at pain-reduction in an adult human population experiencing clinical pain.
- Identify which therapeutic interventions have been assessed in such trials.
- Evaluate the prevalence of individual design features relating to the concept of ‘pragmatism’ amongst the included studies.
- Determine areas for future debate and research within the field of pragmatic trials of pain treatments.

Methods

Protocol registration

A protocol was formulated in accordance with the 2015 statement of Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) (Moher et al., 2015; Page et al., 2020) and pre-registered with PROSPERO prior to commencing data extraction (Booth et al., 2012) (registration ID: CRD42020178954).

Eligibility criteria

Inclusion and exclusion criteria are summarised in Table 18.

Table 18: Eligibility criteria. The 'PICOS' framework groups eligibility criteria into study populations, interventions, comparators, outcomes, and research designs (Higgins et al., 2019), and has been adapted here.

PICOS domains	Criteria applied
Inclusion criteria	
Study population	People with clinical pain (including procedure-related pain), irrespective of age, gender, underlying pathology, severity, and duration of patients' pain experience
Intervention	Any interventions aimed at reducing pain or at affecting an outcome relevant to the treatment or management of people in pain
Comparator	Any control or comparison group
Outcome measures	Primary analyses including on pain or pain-related measures, or the declared primary aim being to assess endpoints directly relevant to the treatment and clinical management of patients in pain
Study design	Randomised controlled trials
Timeframe and publication sources	Primary outcome reports published in peer-reviewed sources between January 2018 and March 2020
Other	Declared by the study authors to be 'pragmatic', 'practical', or 'comparative effectiveness'; Studies published in the languages English, German, Spanish, Italian, French, and Mandarin, and others if translations could be obtained
Exclusion criteria	
Study population	Experimental pain (i.e., healthy participants or patients with pain induced or aggravated for scientific purposes); Less than 40 participants per study arm
Intervention	-
Comparator	-
Outcome measures	Primary outcome measures relating to administrative processes or diagnosis
Study design	Retrospective and observational studies; studies exclusively using registry data; single-case designs; declared feasibility and pilot studies
Other	No full text obtainable

Information sources and timeframe

The following databases were searched from 01 January 2018 to 01 March 2020: MEDLINE®, Embase, and PsychInfo (through Ovid interface); the Cochrane Central Register of Controlled Trials (CENTRAL), NIH Clinicaltrials.gov, CINAHL (nursing and allied health, via EBSCO), and the Physiotherapy evidence database (pedro.org.au). Grey literature was not searched. Reference lists of included studies and retrieved systematic reviews were reviewed for additional eligible studies. When protocols were retrieved, trial registries were consulted or authors contacted to identify trial status.

Search

Database-specific indexing and free-text terms were used to search for the following concepts, using the following Boolean operators: pain OR painful conditions (i.e., specific disease names) AND (pragmatic trials OR practical trials OR comparative effectiveness). Complete search strings are provided as supplement to the main publication (Hohenschurz-Schmidt et al., 2022c).

Study selection

Screening was performed in Covidence (Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org). In a first step, screening was performed based on study titles and abstracts, and then again on full text publications, with each record screened by two independent reviewers at each step. Disagreements were resolved through discussion or, if not possible, by a third party.

Data extraction

Data extraction required two independent reviewers and was reconciled as above. For any included study, protocols were consulted for additional information. Missing data which

could not be retrieved in this way were recorded as not reported. Where reports were ambiguous, authors were contacted for clarification.

The domains of data extraction were source details, funding, trial methods, outcome measures, analysis methods, discussion and contextualisation of information, and reporting. The extraction table is available as supplement to the review's publication (Hohenschurz-Schmidt et al., 2022c).

Information on trial methods

The extraction of study methods had a triple focus. First, key aspects of pragmatic trial design, conduct, and analysis were extracted (such as eligibility criteria, treatment provision and statistical methods). Secondly, to assess how trialists handled the tension between external and internal trial validity, methods deemed to affect internal trial validity were extracted (such as randomization procedures, allocation concealment, and blinding of participants and personnel [Sterne et al., 2019]). Information on placebo control groups, blinding, and number of trial settings were extracted to address currently debated areas of pragmatic trials design (Dal-Ré et al., 2020; Zwarenstein et al., 2020); and baseline age data were extracted to identify potential shortcomings in randomisation by means of heterogeneity testing between trial groups (Clark et al., 2014). Thirdly, relevant information to examine how researchers dealt with the specific challenges of pragmatic trials was extracted (such as the discussion and methodological treatment of potential heterogeneity between study arms, differences between multiple study centres, differences in therapist expertise, setting resources, treatment flexibility and fidelity, lack of blinding, prolonged follow-up periods, differential attrition, and study cost) (Roehr, 2013; Zuidgeest et al., 2017; Keefe et al., 2022).

PRECIS-2 ratings

Descriptive data relating to the nine core domains of the PRECIS-2 tool were sampled (Loudon et al., 2015a) and trials rated for each of these domains. Where domains were not applicable or information insufficient to perform ratings, domains were left blank, reflecting

current practice (Dal-Ré et al., 2020; Zwarenstein et al., 2020). Data extractors were trained on example studies to perform ratings, and relevant reference material was made available (incl. https://www.precis-2.org/Trials). Ratings were performed in duplicate and inter-rater reliability was monitored.

Risk of bias within studies

No formal risk of bias assessment was performed as the purpose of this review was not to judge clinical or comparative effectiveness and because no clinical outcome data were extracted.

Reporting guideline compliance

Compliance with and data on recommended CONSORT reporting items for pragmatic trials were extracted (Zwarenstein et al., 2008).

Data synthesis

The main results of the review were qualitative, presented using descriptive statistics and appropriate graphs. Additionally, it was assessed if certain trial methods were more prevalent under certain circumstances, using appropriate statistical correlation methods and following a pre-specified analysis plan (see pre-registered protocol).

Methodological quality across studies

To assess trial quality, data on reporting quality were sampled, which mainly affects the readers' ability to judge the generalisability of results. Potential shortcomings in randomisation procedures were identified by extracting and analysing average and dispersion measures of study participants' age between groups with a heterogeneity meta-analysis (Clark et al., 2014; Hicks et al., 2018).

Additional Analyses

Sensitivity analyses examined the above correlations without the pre-identified covariates. The only deviation from the analysis plan was the addition of two subgroup analyses, investigating whether PRECIS-2 ratings differed between trials of pharmacological and nonpharmacological pain therapies as well as between trials of acute versus chronic pain.

Results

The review was published in PAIN in 2021 (print publication as Hohenschurz-Schmidt et al., 2022c). The below is an overview of the main findings.

Study selection

The results of search and screening procedures are illustrated in Figure 8. Meta-analyses in the sense of descriptive statistics and correlation analyses were performed on the entire sample of 57 included studies.

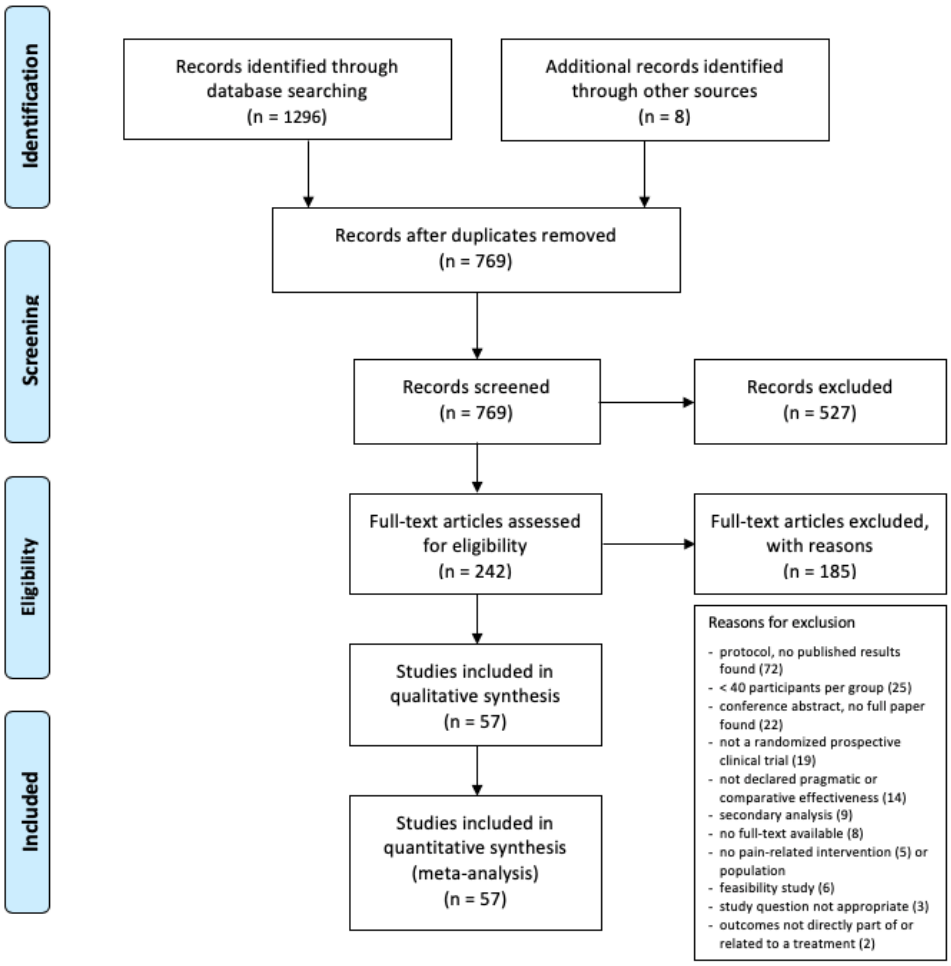


Figure 8: PRISMA flow diagram showing the identification, screening and selection process of records for the systematic review of pragmatic trials for pain. Reasons for exclusion at the full-text eligibility-screening phase are provided.

Study characteristics

Treatment modalities, participant demographics, and trial settings

All included trials investigated programmes of complex interventions, such as rehabilitation, manual therapy, cognitive-behavioural therapies, various forms of patient management, surgery or drug regimens, or treatment programmes of several modalities. Patients presented with pain in various locations and of diverse classifications, but chronic durations and musculoskeletal pain dominated. Trials were conducted in all kinds of settings (including public and private, and primary to tertiary), mostly in the US (35%), followed by the UK (14%), Australia (7%), and Norway (7%). Four studies (7%) were conducted in East Asia and one in South America. Only one study was conducted across multiple countries (Rigoard et al., 2019). Detailed information is provided in Table 19.

Table 19: Study characteristics regarding treatment modalities, participant demographics, and trial settings. All therapies studied across included trials as well as pain disorders, and the average duration of pain are presented. Note: Some pain descriptors have been applied twice, e.g., knee arthritis has been classified as 'Peripheral joint pain' and 'Arthritis'. Further, some samples included patients with musculoskeletal pain, which may have included 'Back or neck pain' and 'peripheral joint pain'. Depending on individual trial reporting, the average duration of the pain-related diagnosis or the time since the onset of pain was used. Acute pain was defined as pain lasting < 4 weeks, subacute as 4 weeks to 3 months, and chronic as 3 months. CFS, chronic fatigue syndrome; CRPS, chronic regional pain syndrome; FM, fibromyalgia; RA, rheumatoid arthritis; OA, osteoarthritis.

Component therapeutic modalities	n of trials	%
Pharmacological therapy	12	21.05
Cognitive-behavioural & other psychotherapy	9	15.79
Surgery	7	12.28
Acupuncture / acupressure	6	10.53
Manual Therapy	6	10.53
Physiotherapy	4	7.02
Multidisciplinary care (non-drug)	3	5.26
General Practice (non-drug)	2	3.51
Rehabilitation	2	3.51
Body-mind therapies	2	3.51
Education	1	1.75

Automated symptom and treatment side effect monitoring	1	1.75
Virtual reality	1	1.75
Dentistry	1	1.75
Pain disorder / descriptor	**n of trials**	**%**
Back or neck pain	19	33.33
Peripheral joint pain	10	17.54
Arthritis (RA and / or OA)	8	14.04
Pain (not further specified)	7	12.28
Post-medical intervention pain	5	8.77
Abdominal and other visceral pain	4	7.02
Neuropathic pain	3	5.26
Headaches	3	5.26
Leg pain	2	3.51
Post-injury pain	2	3.51
Tooth pain	1	1.75
Diffuse chronic pain (CFS, FM, CRPS)	1	1.75
Pain duration	**n of trials**	**%**
Acute	7	12.28
Subacute	2	3.51
Chronic	31	54.39
Mixed	1	1.75
Not reported	16	28.07
Type of setting	**n of trials**	**%**
Primary	25	43.86
Secondary	20	35.09
Tertiary	17	29.82
Community	5	8.77
Setting specification	**n of trials**	**%**
Public hospital	23	40.35
Private hospital	6	10.53
Patient home, phone, text messaging, mail, or online (entirely or predominantly)	6	10.53
Private practice	8	14.04
Military medical practice	3	5.26

Research institute	1	1.75
Rehabilitation centre	1	1.75
Emergency Dept.	1	1.75
University teaching clinic	1	1.75

Trial size

The median sample size at the point of randomization was 234 (Q1/3: 135.5; 363.5); the median number of trial centres or settings 5 (Q1/3: 1; 9.25), with one trial for gout flare-ups taking place across 100 general practice clinics across England (Roddy, 2019) and 12 (21%) single-centre studies. Seven studies did not report the number of treatment centres.

Whilst 37 studies (65%) did not report the number of providers *in the treatment group*, the median number of reported therapy providers was 11 (Q1/3: 6.25; 35.5), with Adams et al. (2019) involving 400 general practitioners. In control groups, the number of providers was only reported in 20 (35%) of included trials, making it impossible to provide an overall number of therapy providers across all trial groups.

Concomitant treatments

Concomitant pain treatments were disallowed in six trials (11% of applicable cases, N = 55), and two further trials discouraged patients from seeking treatment outside the trial. Ten trials did not report whether concomitant therapies were permissible.

Where allowed, concomitant treatments were unrestricted in 30 trials (68% of applicable cases) and six trials permitted any concomitant treatment other than those akin or similar to the study interventions. Six trials excluded individual unrelated interventions. Only one drug trial made specific allowances for medications and dosages, including some of the same class as study interventions and opioids (Emery et al., 2020). Another trial changed the regimen from allowing NSAIDs only in the first six weeks to applying no restrictions thereafter (Chesterton et al., 2018). Thirty trials (68% of applicable cases) reported detail on concomitant analgesic treatment actually received, four of which, however, only partly or for

specific treatments deemed to relate to the trial intervention, but not for others (Wibault et al., 2018).

Trial funding

Seven trials (12%) were fully industry-funded, 44 (77%) had public funders, and five (9%) were funded by mixed sources. Funding sources were not reported in one trial (Qi et al., 2018).

Trial protocols

Protocols were registered for all but one trial (Qi et al., 2018).

General trial methods

Trial design and comparators

The main design features of included trials are presented in Table 20. All but one of the included trials were parallel-group RCTs, with Berdal and colleagues (2018) employing a stepped-wedge design and another trial including a cross-over option after 12 weeks (Rothrock et al., 2019). The number and nature of groups differed between trials. Most trials (45; 79%) employed a two-group design, ten trials (18%) had three groups, and two trials (4%) had five groups. More than three-quarters of the trials in the sample were comparative effectiveness trials, comparing multiple specific interventions or using treatment-as-usual as the comparator (other comparators listed in Table 20).

Table 20: Methods of trial design. Employed comparators and recruitment and randomisation methods are presented. Notes: Multiple comparator groups were possible. The difference between a waitlist control group and a no-treatment control group is that patients expect treatment at a later point or know that they have been assigned to not receiving any treatment, respectively. Categories marked * are deemed part of Comparative Effectiveness Research (CER). Convenience sampling is the recruitment of patients who attend the trial-delivering service anyway, whilst targeted strategies seek to specifically contact populations of potentially eligible participants. The category of 'blocked randomization' includes various was of blocking, including a single fixed block size, regularly varying sizes, and randomly permuted block sizes. Blocking was occasionally stratified by site. Stratification was usually by trial centres (sites). 'Other stratification' includes stratification by gender, diagnosis or treating surgeon. Cluster randomization refers to trials where the unit of randomization were not patients but, for example, clinics or individual providers.

Comparator	**n of trials**	**% of sample**
Another active specific therapy (comparative effectiveness) *	29	50.88
Treatment / care as usual *	14	24.56
Placebo or sham intervention	5	8.77
no treatment group (explicitly assigned, i.e., patients know they will not get any treatment)	4	7.02
Treatment / care as usual plus something else (e.g., advice, education, etc.) *	2	3.51
Waitlist control	1	1.75
Advice only	1	1.75
Wait & see (not waitlist but monitoring)	1	1.75
No-treatment group (but unaware of trial)	1	1.75
Recruitment method		
Targeted recruitment (such as identification through records)	25	43.86
Convenience sampling	16	28.07
Not reported	9	15.79
Mixed (convenience and targeted)	8	14.04
Method of randomisation		
individually randomised	27	47.37
of which simple randomisation	15	26.32
of which blocked randomisation	12	21.05
stratified by site	15	26.32
other stratification	9	15.79
Cluster randomised	6	10.53

Blinding and informed consent

Participants were reported to be blinded in 13 trials (24% of the trials reporting on participant blinding, n = 54), providers in four (7%, n = 55), and outcome assessors in 45 trials (90%, n = 50). Only five studies reported unblinded assessment (Table 21).

Fifty-four of 57 trials (95%) reported that patients gave informed consent, with none of the trials stating clearly what this information and consent process entailed. In two trials consent for patients was waived. In a third trial, the unit of randomization were physicians for whom consent was not required; patients, however, did provide informed consent. Lastly, a single article did not report whether or not patients provided informed consent.

Outcome analysis and interpretation

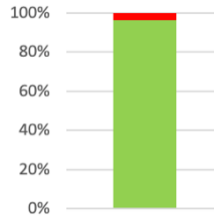
Trials were designed as superiority trials in 52 instances (91%), four (7%) were non-inferiority or equivalence trials, and a single study did not report whether or not the trial was designed to show a difference between groups. 'Unsuccessful' superiority trials cannot claim equivalence between interventions (Gewandter et al., 2015); Nonetheless, equivalence or comparative effectiveness was the reported in 9 out of 24 (38%) superiority trials where no significant difference between groups had been demonstrated. Despite the recommendation to include a third (placebo) control group in non-inferiority trials in order to account for the trial-specific possibility of no demonstrable effect beyond placebo in the control group (Food and Drug Administration, 2016), none of the four non-inferiority trials included a placebo control group.

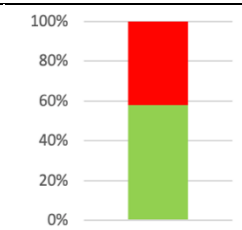
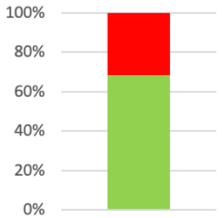
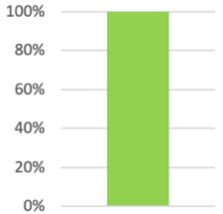
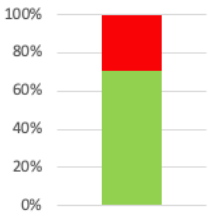
Out of fifteen trials with multiple outcome measures defined as primary outcomes, nine (60%) did not address the issue of multiplicity in their analysis.

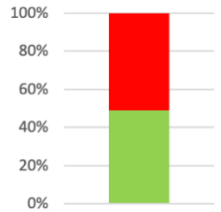
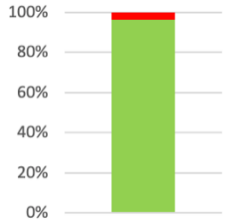
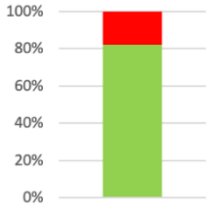
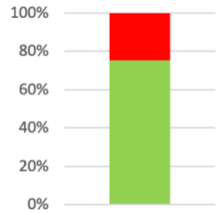
Adherence to reporting guidelines

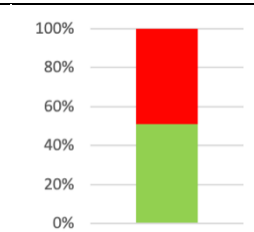
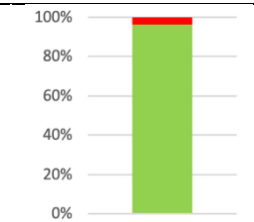
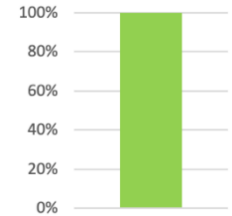
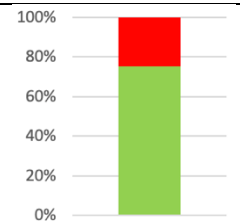
Adherence to relevant CONSORT reporting guidelines is presented in Table 21. Deficits in reporting information relevant for the interpretation of pragmatic trials were mainly found regarding reporting of: Background (not situating the tested intervention amongst available treatment alternatives), description of study settings and whether alterations were made to deliver the trial intervention, and outcome justification (not clarifying why the primary outcome was important).

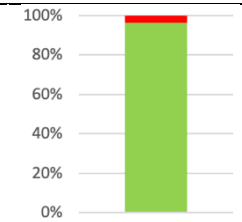
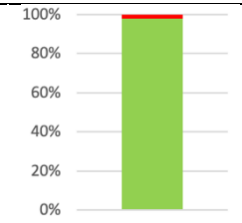
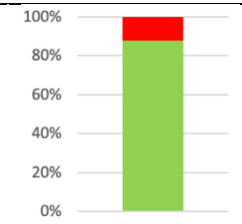
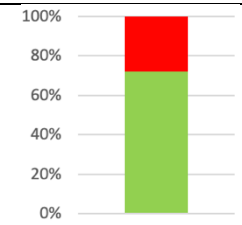
Table 21: Reporting of selected items of the 2010 update of the Consolidated Standards of Reporting Trials (CONSORT) statement (Schulz et al., 2010)[1] and all items of the extension for the reporting of pragmatic trials, as published in 2008 by Zwarenstein et al.[2]. Where indicated by colour or strikethrough font, items were modified to match the review question or split into several items in order to be extractable as individual data points. Bar charts indicate the percentage of studies complying with respective reporting items (green) and not complying even though applicable (red). If items were not applicable to the entire sample of 57 studies, the applicable number is stated in the respective row.

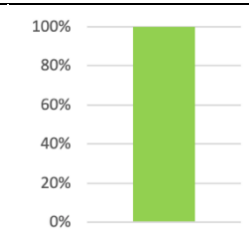
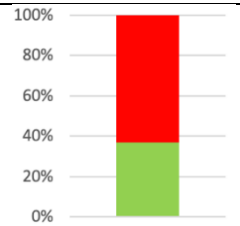
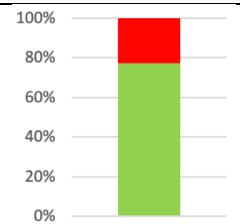
Item (number, section & CONSORT document (1 or 2))	Description of reporting item	Results: Number of studies which complied with respective reporting items (n, %)
2: Background[2] modified	Describe the health or health service problem that the intervention is intended to address	55 (96.49%) 

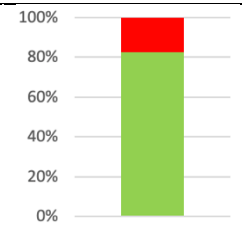
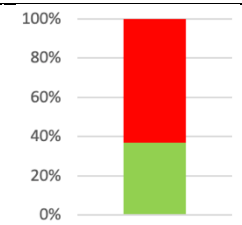
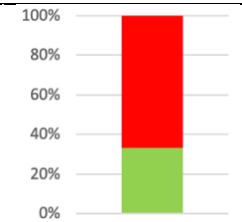
	and other interventions that may commonly be aimed at this problem	33 (57.89%) 
3: Participants[2]	Eligibility criteria should be explicitly framed to show the degree to which they include typical participants **and/or, where applicable,** typical providers (e.g., nurses), institutions (e.g., hospitals), communities (or localities e.g., towns) and settings of care (e.g., different healthcare financing systems)	39 (68.42%) 
5: Interventions[1] modified	Precise details of the interventions intended for (the intervention group) ~~each group and how and when they were actually administered~~ (Note to reader: Colour and strikethrough font indicate modification of original CONSORT items)	57 (100%) 
	If the 'treatment-as-usual' or 'usual care' was used as comparator, provide additional information as to the nature of the intervention(s) available as part of this	(applicable in n = 17) 12 (70.59%) 

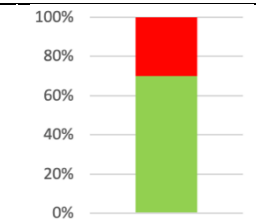
	If the 'treatment-as-usual' or 'usual care' was employed as comparator, collect and report data on care received by patients in this group	(applicable in n = 17) 10 (58.82%)
4 Interventions[2]	Describe extra resources added to (or resources removed from) usual settings in order to implement intervention	28 (49.12%) 
	Describe the health or health service problem that the intervention is intended to address	55 (96.49%) 
	Indicate if efforts were made to standardize the intervention or if the intervention and its delivery were allowed to vary between participants, practitioners, or study sites	Not applicable as intervention automated: 1; Reported: 46 (82.14% of 56); of those standardized: 35; Not standardized: 9. 
	Describe the comparator in similar detail to the intervention	43 (75.44%) 

6 Outcomes[2]	Explain why the chosen outcomes and, when relevant, the length of follow-up are considered important to those who will use the results of the trial	29 (50.88%) 
7a: Sample size[1]	Report how the sample size was determined.	55 (96.49%) 
7: Sample size[2]	If calculated using the smallest difference considered important by the target decision maker audience (the minimally important difference) then report where this difference was obtained	*Not extracted*
8b: Randomization[1]	Report the type of randomization & details of any restriction (such as blocking and block size).	57 (100%) 
10: Allocation concealment implementation	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions?	43 (75.44%) 

11a: Blinding / masking[1] modified	Whether participants,	Not applicable in one case as patients unaware of participating in a trial. Reported in 54 (96.43%) of 56.	
	those administering the interventions,	Not applicable in one case as intervention independent of providers. Reported in 55 (98.2%) of 56 relevant trials.	
	and those assessing the outcomes were blinded to group assignment	Reported in 50 cases (87.72%).	
11: Blinding / masking[2]	If blinding was not done, or was not possible, explain why.	31 (72%) of relevant studies reported reasons (n = 43).	

13a: Participant flow[1] modified	Flow of participants through each stage (a diagram is strongly recommended)—specifically, for each group, report the numbers of participants randomly assigned	57 (100%) 
	~~receiving intended treatment~~	*Not extracted*
	completing the study protocol	57 (100%)
	analysed for the primary outcome;	57 (100%)
	describe deviations from planned study protocol, together with reasons	21 (36.84%) complied. Of those: 10 reported following protocol; Deviations with reasons reported in 11; Deviated without providing reasons: 4. 
13: Participant flow[2]	The number of participants or units approached to take part in the trial, the number which were eligible, and reasons for non-participation should be reported.	44 (77.19%) 
16: Numbers analyzed[1] modified	Whether the analysis was by 'intention-to-treat'; For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	48 (84.21%) reported the primary analysis as 'Intention-to-treat'; Not discernible in one instance; all other did not explicitly report a primary Intention-to-treat analysis.

		Out of those 48 studies which called their primary analysis 'Intention-to-treat', ten trials (20.83%) excluded participants from the primary analysis who did not provide follow-up data or where data were missing.
19: Harms[1]	All important harms or unintended effects in each group.	Whether or not significant harms or unintended effects occurred was reported in 47 studies (82.46%). Harms did occur in 22 of those studies, and in nine of those cases there was a significant difference between groups. 
21: Generalizability[2] modified	Describe key aspects of the setting which determined the trial results.	21 (36.84%) 
	Discuss possible differences in other settings where clinical traditions, health service organisation, staffing, or resources may vary from those of the trial	19 (33.33%) 

25: Funding[1]	Sources of funding and other support (such as supply of drugs) **and** role of funders.	40 (70.18%)	

Pragmatic trial methods

Average PRECIS-2 ratings

The PRECIS-2 instrument has nine domains that are each rated from 1, indicating a very explanatory design, to 5, indicating a pragmatic approach to a design feature (Loudon et al., 2015a). Reviewers were able to rate 97% (496) out of 513 items, with the remainder not reported. In this sample of studies, the average rating across all domains was 3.8 (SD 0.62), indicating that, on average, 'pragmatic' and comparative effectiveness trials in pain employ more methods that belong to the pragmatic end of the spectrum. The lowest overall rating was an average of 1.86 points (Qi et al., 2018), the highest was 5/5 for all domains ((Griswold et al., 2018). Ratings for individual domains and factors influencing these ratings are presented below and in Figure 9.



Figure 9: Average PRECIS-2 scores per domain for all included trials. The thick blue line represents mean scores and the shaded area standard deviations. For more statistical detail see supplementary table of journal publication (Hohenschurz-Schmidt et al., 2022c).

Eligibility

Assessment of this domain was possible in all instances (M 3.97; SD 1.1; n = 57). The main reasons for low ratings in this domain was the exclusion of patients with comorbidities common for the specific trial population (15 studies; 26%). Similarly, common medications were a reason for non-eligibility in seven studies (12%).

Not informing PRECIS-2 ratings, but also relevant for generalisability of trial findings are entry requirements for trial providers and centres:

Eligibility criteria for providers, specifically a minimum number of years in practice, could be confirmed in six trials (11%). This information was not reported in 41 studies (72%) and not relevant for assessment in a further three trials (e.g., where the treatment was automated). A minimum amount of experience with the trial intervention (other than trial-specific training) was required in at least six trials (11%).

Entry criteria for trial centres existed in at least eight cases (14%), not in four cases (7%), and were not reported on in 42 trials (74%). Three further studies took place entirely in the patients' home or another community setting and were thus not relevant for this assessment.

Recruitment

Convenience sampling was deemed most in line with the principles of pragmatic trials for patient recruitment (Loudon et al., 2015a). Instead, however, almost half the trials (25 cases, 44%) employed targeted recruitment methods, (e.g., patient identification through records or adverts). A mixed approach was employed by 14%, and recruitment methods were not reported in another 16% of studies. More laborious recruitment strategies led to the lowest average PRECIS-2 ratings for this domain (M 3.03; SD 1.6; n = 47); it also points to patient recruitment as a major challenge in pragmatic trials, especially when sought to be performed 'pragmatically', i.e., in tune with every day-practice. In further support of this argument, out of 56 studies which reported a target recruitment number, 15 (27%) did not reach their aim.

Setting

The PRECIS-2 domain 'setting' asks how different the setting of the trial and the usual care settings are. Whilst reporting of relevant information was often poor (Table 21), extrapolation allowed reviewers to rate this PRECIS-2 domain in all but one case, averaging at 3.79 points (SD 1.4; n = 56).

Organization

The PRECIS-2 domain 'organization' compares the provider expertise, resources and the organisation of care delivery in the intervention arm of the trial to those available in usual care (Loudon et al., 2015a). Reporting of resource alteration and other relevant information was low (Table 21), and trial-specific training of providers was part of at least 23 trials (40%). PRECIS-2 ratings for this domain were lower than for most other domains, with a mean of 3.5 (SD 1.3; n = 51). Poor reporting meant that this domain could not be rated in six cases.

Flexibility (Delivery & adherence)

Aspects of the intervention delivery were standardised in 35 trials (61%), with 11 reports not indicating (19%). Out of those trials in which the delivery was reported to be standardised, about one third reported monitoring of the fidelity with which treatments were provided, for example by record checking or video taping of treatment sessions (n = 11, 31%). Where applicable, these features contributed to low PRECIS-2 ratings for this domain and the sample's average was 3.51 (SD 1.2; n = 56) for flexibility in treatment delivery.

For the patients' adherence to treatments and interventions, however, an average of 4.34 points was obtained (SD 1.0; n = 56), meaning that patients were flexible in how they had to follow intervention plans, adherence was encouraged little more than what would be expected

in usual practice, and non-adherence rarely meant exclusion from the trial analysis. In fact, post-randomisation exclusion criteria such as minimum compliance, absence of adverse events, or other outlier criteria, were only present in a small percentage of trials (n = 5; 9%).

Follow-up and attrition

The PRECIS-2 ratings for the domain 'follow-up' is concerned with the frequency and duration of follow-up appointments as well as the intensity of clinical assessments compared to usual care (Loudon et al., 2015a). Based on this, the average rating was 3.24 points (SD 1.3; n = 57), meaning that follow-up was often more elaborate than what would be expected from normal practice.

Relatedly, we examined the length of the follow-up period and differential attrition. The primary time point at which outcomes were measured was a median of 26 weeks post-randomization (Q1/3: 8; 52 weeks; n = 36) ranging from a single day to five years. However, there were many studies which assessed outcomes over a period (n = 7), at flexible time points (n = 3), studies which defined several time points as primary (n = 2), and several which did not specify a *primary* point of follow-up (n = 8). A better indicator for the length of pragmatic trials may thus be the longest time point of follow-up, for which the median was about one year after randomization (Median 50 weeks, Q1/3; 23.25; 52 weeks; n = 56; not reported in one case).

Over such potentially long follow-up periods, attrition of study participants can be expected. A third of studies reported differential attrition (19; 33%, n = 57). Where there was differential attrition, it was almost as often into the direction of the control group (n = 9) as it was into the direction of the intervention group (n = 11), with intervention groups losing an average of 14.9% (SD 12.9) and control groups losing an average of 14.8% (SD 12.8) of participants until the point of primary follow-up ($t(112) = 0.032$, $p = 0.97$), possibly reflecting that both groups were usually active interventions.

Outcomes

The choice of outcomes in pragmatic trials should reflect what 'matters' to the patient, choosing direct symptom reports or function-related measures over lab tests, surrogate markers, expert assessment or other external judgement (Loudon et al., 2015a). In this sample of analgesia trials, subjective pain ratings, certain condition-related questionnaires, or pain-related functional assessments were the obvious choice (Dworkin et al., 2005). Indeed, on average, trials had the highest rating for this domain (M 4.46; SD 1.0; n = 57). Forty-nine trials obtained a primary outcome through such patient report (86%; n = 57), five trials used a lab or other remote physiological assessment such as radiographs (9%), three used a physical or personal assessment (5%), and one trial each (2%) used data obtained from health records or an objective incident in the medical management of patients, namely the reoccurrence of a medical intervention.

As secondary outcomes, objective measures were far more common, being reported in 30 trials (54%) and including measures of healthcare utilisation, physical tests, lab markers, and absence from work. Only half the reports, however, complied with the CONSORT item to justify the chosen outcome and length of follow-up (29 cases, 51%) (Table 21, item 6). Whether or not significant harms or unintended effects occurred was reported in 47 studies (82%) (Table 21, item 19). Harms did occur in 22 of those studies, and in nine of those trials there was a significant difference between groups.

Primary analysis

The highest PRECIS-2 rating for this domain is obtained by trials which perform a true intention-to-treat (ITT) analysis for their primary outcome assessment (Loudon et al., 2015a), meaning that all patients randomized are analysed as if treated, irrespective of actual treatment compliance or a failure to attend follow-up assessments (essentially resulting in missing data). 'Pragmatism' in the primary analysis was high, averaging at 4.3 points (SD 1.3, n = 57). The distinction between a true intention-to-treat analysis and a modified ITT (Abraha and Montedori,

2010; Montedori et al., 2011; Gewandter et al., 2014; Higgins et al., 2019) is made clear by the following data: Whilst 48 studies reported to have performed an ITT as primary analysis (84%; not discernible in one instance), 10 of those (21%) excluded participants who did not provide follow up data or had missing data.

Multi-centre trials

At least 45 trials (79%) were multi-centre studies (insufficient reporting in 7 studies, n = 50). Only two studies reported having assessed differences between centres (for example in terms of case load, resources, and attending patient population. Those authors also discussed how such differences may have affected trial results, enabling readers to better judge generalisability. Randomization was stratified by site in 15 trials, another method to account for potential differences between study centres.

At least 21% of trials were single-centre trials, which have been highlighted as potentially 'unpragmatic' in recent debates due to supposed low generalisability (Dal-Ré et al., 2020).

Second-level analyses

Baseline heterogeneity

When testing for differences in mean age between intervention and control groups as an indicator of baseline heterogeneity, no significant difference was detected ($t(52) = 1.79$, $p = .079$), suggesting that randomisation was not systematically biased in this sample of studies.

Preliminary analysis: Testing for potential confounders

There were significant correlations between the overall trial size (total sample size at randomisation) and a number of other variables of interest: These included overall PRECIS-2 scores, driven by significant correlations with the domains 'setting', 'organisation', and 'analysis'.

Larger trials were also less likely to show differential attrition, irrespective of in which group most drop-outs occurred (treatment vs. control group) (Table 22). Sample size was thus included as a covariate of no interest in subsequent analyses of PRECIS-scores and attrition.

Table 22: Correlation analysis between total sample size at randomisation and domain-specific PRECIS-2 ratings. The table also shows the correlation analysis between sample size and differential attrition (last column) as measured by the difference in drop-outs between groups, irrespective of the 'direction' of attrition, i.e., in which group more patients were lost to follow-up (Spearman's Rank Order Correlation used as data not conforming with normality assumption; * p < 0,05, ** p < 0,01; 2-tailed). BL, baseline.

	PRECIS-2 domain	Eligibility	Recruitment	Setting	Organization	Flexibility (delivery)	Flexibility (adherence)	Follow-up	Outcome	Analysis	Total PRECIS-2 score	Different. attrition
Total sample size at BL	Corr. coeff.	.250	.029	**.451****	**.281***	.031	.193	.190	.161	**.273***	**.408****	**-.360****
	p	.061	.846	**.000**	**.046**	.822	.155	.157	.232	**.040**	**.002**	**.006**
	n	57	47	**56**	**51**	56	56	57	57	**57**	**57**	**57**

Correlation analyses

PRECIS-2 ratings and trial features

A range of planned correlation analyses were performed to identify potential associations amongst different trial features and with ratings of pragmatism, randomization method, and blinding status of participants (Table 23).

Specifically, it was asked if overall PRECIS-2 scores were associated with the number of trial centres, the source of trial funding (public, industry, mixed), the primary therapy investigated, the pain condition of participants (index pain disorder), and the employed analysis method (distinguishing a true ITT, modified ITT, and no ITT).

Table 23: Correlation analyses amongst trial methods with ratings of trial pragmatism (PRECIS-2 scores), randomisation methods, and analysis method. Statistical tests were part correlation analyses where covariates were controlled for and Spearman's rho where this was not indicated. †Sample size was used as covariate of no interest. Sensitivity analyses assess the same correlation without controlling for pre-identified confounding variables. * significant at p < 0.05 (2-tailed). DV, dependent variable.

	DV	Results	Sensitivity analysis
PRECIS-2 average	Sample size		
	Number of trial centres †	.190, *p* = .191, (df = 47)	
	Funding source †	-.048, *p* = 727, (df = 54)	
	Index therapy †	-.005, *p* = .97, (df = 54)	.089, *p* = .512, n = 57
	Index pain disorder †	**-.285*, *p* = .033, (df = 54)**	-.213, *p* = .112, n = 57
	Analysis method †	.198, *p* = .16, (df = 50)	.284*, *p* = .039, n = 53
Randomisation method	PRECIS-2 average †	-.197, *p* = .145, (df = 54)	-.137, *p* = .31, n = 57
	Baseline heterogeneity	*Analysis not conducted as no trials with sign. between-group age differences at baseline*	
	Sample size	.21, *p* = .122, n = 57	
	Analysis method	.0, *p* = .99, n = 53	
	Funding source	-.072, *p* = .594, n = 57	
Blinding of participants	PRECIS-2 average †	.214, *p* = .124 (df = 51)	.135, *p* = .331, n = 54
	Sample size	-.081, *p* = .562, (n = 54)	
	Analysis method	.2, *p* = .16, (n = 51)	
	Funding source	.111, *p* = .424, (n = 54)	

Controlling for sample size, only the variable 'Index pain disorder' was associated with PRECIS-2 ratings (r(54) = -.285, *p* = .033). Post-hoc analyses to see if specific diagnoses drove this correlation were not possible due to small case numbers in some of the categories.

Whether or not study participants were reported to be blinded to group allocation did not correlate with average PRECIS-2 ratings, the funding source, the size of the trial, or the employed analysis method.

Subgroup analyses

The sample was divided into drug and non-drug trials as well as populations of acute and chronic pain patients. Hence, it was examined whether total PRECIS-2 ratings differed between these groups of trials. Only trials providing a clear indication of the patient sample's duration of pain and fitting into the categories of acute (< 4 weeks) or chronic (> 3 months) were included into this analysis (n = 38).

One-way analysis of variance revealed no difference in average PRECIS-2 scores between trials of pharmacological and nonpharmacological therapies ($F(1, 55) = 0.27$, $p = 0.6$), with ratings averaging 3.7 ($\pm$0.7) and 3.8 ($\pm$0.6), respectively. Domain-specific ratings only differed for the flexibility with which treatments were delivered, with drug studies allowing significantly less flexibility (2.8 ± 1.2 vs. 3.7 ± 1.6; $F(1,54) = 5.14$, $p = 0.027$).

A significant difference in overall PRECIS-2 scores existed between acute and chronic pain trials ($F(1, 36) = 5.14$, $p = 0.03$), with higher scores in acute trials (4.3 ± 0.7 vs. 3.7 ± 0.6). Exclusion of an outlier, a chronic pain trial with the lowest overall rating (Qi et al., 2018), did not alter the statistical significance of this result ($F(1,35) = 5.32$, $p = 0.027$). Upon evaluation of individual PRECIS-2 domains, only the domain 'recruitment' differed significantly between groups ($F(1,28) = 5.88$, $p = 0.02$), with chronic pain trials investing more into patient recruitment than acute trials.

Discussion

This systematic review of methods describes current practice in the field of pragmatic trials in clinical analgesia research. A 'typical' trial is illustrated in Table 24.

Table 24: A 'typical' pragmatic or comparative effectiveness RCT in pain research. Based on a sample of 57 trials published between 2018 and 2020, using the most common design features and numerical averages (Hohenschurz-Schmidt et al., 2022c).

Design feature	**'Typical' implementation**
Interventions	Studying complex nonpharmacological interventions
Comparator	Comparison between two or more active interventions, often also with 'usual care'
Randomisation method	Individually randomised (at patient level)
Number of participants	Several hundred (Mean = 297)
Number of trial centres	Multiple (Mean = 12)
Duration	About one year follow-up period (Mean = 58 weeks)
Funding	Public funders
PRECIS-2 design features	
Participants (eligibility)	Relatively similar to the target population; Usually people living with persistent, often musculoskeletal pain
Recruitment methods	With more effort than what is normal in clinical practice, such as targeted identification through records or advertising
Settings	Rather similar to usual practice
Organisation	Difficult to assess as often not reported, but likely employing staff that are more experienced and better qualified than those normally delivering target interventions
Treatment delivery (flexibility)	Treatments delivered with some standardisation
Treatment delivery (adherence)	No more encouragement of patient adherence than in normal practice
Follow-up (assessments)	Elaborate and time-consuming follow-up assessments
Primary outcome measures	Disability or quality of life measures

Primary analysis	All randomised patients, irrespective of treatment adherence or data completeness (true intention-to-treat approach, using various forms of imputation for missing data)

Between 2018 and 2020, pragmatic trials were missing from research areas such as diffuse pain (e.g., fibromyalgia) or injury-related pain. They were also rare in neuropathic pain and headache research. They were rarely conducted in community settings or at patients' homes. Most pragmatic RCTs were parallel-group comparative effectiveness studies, and variations such as adaptive designs or stepped-wedge cluster randomised designs were almost absent. Finally, the sparse reporting of information about trial centres, providers, study populations, and comparator interventions undermined a pragmatic trial's primary purpose to inform decision-making by not allowing readers to adequately judge generalisability (Zwarenstein et al., 2017).

Notably, the sampled trials were not always 'pragmatic' as defined by the PRECIS-2 tool. For example, researchers employed more standardisation of intervention delivery than what is normal in clinical practice, they invested more effort into patient recruitment, employed more sophisticated treatment centres, and they conducted more elaborate outcome assessments than what patients experience in routine care. Some trials even included a placebo control arm or were conducted in single centres only, despite questionable alignment with pragmatic trials philosophy (Dal-Ré et al., 2020; Zwarenstein et al., 2020). In the publication of the review (Hohenschurz-Schmidt et al., 2022c), the thesis author hypothesised that such systematic deviations from the principle of pragmatic trials may reflect two challenges: First, the practical constraints of clinical trial research, where a certain number of participants need to be recruited and treated in a limited timeframe for successful project completion; or where incentives exist to perform more extensive outcome assessments to address additional research questions beyond the most pertinent pragmatic question. Secondly, the scientific need to answer a research question with high certainty may mean that 'pragmatism' is compromised to control bias. This could apply to the observed large degree of intervention standardisation in the present sample. Fidelity monitoring and limits on the use of concomitant treatments are similarly methods to reduce bias from unknown intervention exposure.

Beyond that, these deviations from the pragmatic 'ideal' may illustrate another important consideration in pragmatic trials and a limitation of the PRECIS-2 tool. As Zuidgeest et al. (2017) argue, it may occasionally be necessary to sacrifice 'pragmatism' in one PRECIS-2 domain to achieve larger generalisability in another. The most obvious example is the specific recruitment of patients from minority populations who represent the target population in clinical practice but are unlikely to engage in research endeavours. Other authors have argued similarly since, including that such efforts promote social justice in pragmatic trials (Ali et al., 2022b). In theory, a PRECIS-2 toolkit facilitates reporting of trial intent and justifications of methods (available at: https://www.precis-2.org/Help/Documentation/Toolkit). This toolkit includes a table for such purpose, the use of which, however, is rarely reported. Indeed, this review found that methodological choices are rarely justified, nor their impact on external or internal validity discussed.

For this PhD project, the review provided insights specific to nonpharmacological pain interventions and allowed for comparisons with research practice in drug trials. Not only were PPS interventions studied disproportionately often in pragmatic or comparative effectiveness trials, accounting for about 50% of studied interventions (Table 19), but they also employed slightly different methods (allowing for more flexibility in treatment delivery). This points to pragmatic trials being a topic of particular interest in PPS research which might require specific consideration of the practical circumstances of PPS interventions.

Limitations

To contextualise the review findings, a comparison with trials that were not declared 'pragmatic' would have been useful but was beyond the scope of this project. Other published reviews focus on specific pain populations or interventions, making comparisons difficult.

In the selection of studies, the review also relied on author-declared pragmatism and the assumption that comparative effectiveness trials aim to answer pragmatic research questions. Arguably, this was unavoidable due to the lack of conceptual and terminological clarity around

pragmatic trials (Pawson, 2019; Taljaard et al., 2022), and the continuous nature of trials along the explanatory-pragmatic spectrum. Nonetheless, misapplication of the term 'pragmatic' by study authors occurs (Dal-Ré et al., 2018; Taljaard et al., 2022) and may explain some outliers with low PRECIS-2 ratings. Unawareness of pragmatic trial reporting guidelines may partly explain non-compliance in comparative effectiveness research (see Hohenschurz-Schmidt et al., 2022c, for further discussion). Further, the PRECIS-2 ratings were based on considerable extrapolation and subjectivity due to inadequate reporting and the nature of the tool (which requires knowledge of 'normal clinical practice' in the context of individual trials) (Rothwell, 2010; Loudon et al., 2015a). These concerns are mitigated by the fact that ratings paralleled those from comparable reviews (Glasgow et al., 2012; Sepehrvand et al., 2019). Similarly, some challenges and reasons for the way that pragmatic pain trials are currently conducted had to be inferred, as reasons for design choices were rarely explicitly reported. Discussions with experienced collaborators of the IMMPACT statement presented in the next chapter helped clarify pertinent considerations.

Conclusion and next steps

In sum, the review findings and their discussion suggest that more nuanced guidance for pragmatic trials is required, balancing the need for generalisability with practical and scientific constraints of clinical trials. Decisions are likely dependent on each trial's circumstances and objectives (Witt et al., 2012), so that such a framework should also encourage better reporting of decision intent and method justification (Nicholls et al., 2020), not only of CONSORT items for pragmatic trials (Zwarenstein et al., 2008). Methodological guidance may be particularly important for PPS interventions, which are disproportionally represented in pragmatic trials research and possibly require special consideration regarding methodological challenges. Finally, the absence of pragmatic trials from certain areas of pain research may require further consideration.

This review ascertained the prevalence and nature of pragmatic and comparative effectiveness trials for the treatment and management of people with pain. It formed the basis of guidance development on trial methods presented in the next chapter. Ideally, this endeavour contributes to more researchers harnessing the potential of pragmatic trials in providing clinically applicable insights in a scientifically rigorous manner.

Chapter 5: Development of Best-Practice Considerations for Pragmatic Trials

Introduction

Most methodological publications about pragmatic trials focus on promoting such trials or on elaborating the concept of pragmatic trials. As described in the section introduction, pragmatic trials are commonly promoted because traditional placebo-controlled trials are thought unable of answering the full range of clinical questions (Ford and Norrie, 2016; Zwarenstein et al., 2017; Huebschmann et al., 2019; Cashin et al., 2020). For example, a pragmatic trial is valuable to assess whether a given therapy works as well as, or better than, established care when studied in a broad population and in non-academic settings (Cherkin, 2021). Publications clarifying or elaborating on the concept also abound, sometimes introducing the concept to a new field (Chalkidou et al., 2012; Rowbotham et al., 2013; Gordon et al., 2020a; Levy et al., 2022) or debating whether specific design features align with a pragmatic trial philosophy (Dal-Ré et al., 2020; Zwarenstein et al., 2020). Rarely, however, are methodological challenges discussed critically and comprehensively in the context of the need of trials to produce accurate and valid results (despite a clear emphasis on such issues in the seminal publication by Schwartz and Lellouch [1967]). Only recently have publications begun to emerge that seek to reconcile conducting clinical trials under real-world conditions with considerations of bias control. For example, authors are calling for balancing internal and external validity (Wasan, 2014; Keefe et al., 2022) or discuss the potential merits of methods traditionally perceived as 'explanatory', such as blinding (Christian et al., 2020), fidelity monitoring (Kerns et al., 2022), or targeted recruitment (Ali et al., 2022b). Unifying guidance is, however, not available. The systematic review of this PhD project mapped the current practice in the field of pragmatic trials for pain, also postulating a pervading influence of tensions between the conceptual ideal of pragmatic trials, practical constraints of trial research, and scientific requirements (Hohenschurz-Schmidt et al., 2022c).

This tension also emerged in the 2020 IMMPACT process to make methodological recommendations for pragmatic RCTs in pain research. The process brought together senior experts from both ends of the pragmatic-explanatory spectrum, and the thesis author negotiated and reconciled the often conflicting positions of these stakeholders.

Preliminary considerations: context of guidance development

The Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) publishes "*consensus reviews and recommendations for improving the design, execution, and interpretation of clinical trials of treatments for pain*" (IMMPACT.org, n.d.). Led by Professors Dennis C Turk and Robert H Dworkin, IMMPACT works under the auspices of the Analgesic, Anesthetic, and Addiction Clinical Trials, Translations, Innovations, Opportunities, and Networks (ACTTION), a "*public-private partnership with the United States Food and Drug Administration*". For its annual meetings, the IMMPACT directors invite stakeholders from academia, industry, advocacy groups, and US regulators and government agencies, and they discuss a designated topic, eventually producing and publishing guidance. Such guidance has been highly influential in the pain field, with the top 5 IMMPACT statements cited a total of 9,300 times by March 2023 (range 600 – 3,600 as per Google Scholar, e.g., Dworkin et al., 2005, 2010, 2012).

Nonetheless, IMMPACT has only recently begun to base its meetings on systematic evaluations of the field (Gewandter et al., 2014, 2015; Dworkin et al., 2016; Ferguson et al., 2022) and does not employ structured consensus-finding methodology (such as e.g., Moher et al., 2010b).

For the thesis author, the invitation to lead the IMMPACT XXIV endeavour on pragmatic trials presented an opportunity to engage with some of the leading pragmatic trial researchers and trial methodologists, and to develop recommendations with them. However, unlike when conducting the preparatory systematic review on pragmatic trials or the guideline development on (sham) control interventions, the author had little authority to influence the processes of the meeting and guideline development. Below, own contributions will be clearly identified and it

will be critically discussed where the lack of structured consensus methodology represented a challenge or limitation.

Traditionally, IMMPACT recommendations had focused on efficacy trials and internal validity (such as Dworkin et al., 2010, 2012; Gewandter et al., 2019). Also in 2020, only a small proportion of the meeting attendees were invited specifically for their expertise regarding pragmatic trials, (http://www.immpact.org/meetings/Immpact24/participants24.html). Similarly, several keynote presentations focused on methods to limit or measure heterogeneity. For example, Prof John T Farrar started his lecture with a definition of 'assay sensitivity', a concept foreign to the pragmatic trials literature. Only Professors Lynn DeBar, Robert D Kerns, and Dr Matthew J Bair had a clear focus on and expertise in pragmatic trials. In the two panel discussions (2.5 hrs in total) and single consensus discussion (1.5 hrs), many methodologic details were discussed, but this apparent polarisation between researchers with explanatory and pragmatic preferences was not resolved.[10] Individual participants' comments indicated that concerns had not been satisfactorily addressed. The absence of voting mechanisms made it impossible to gauge how opinions were distributed across all attendees. As a result, in writing the meeting's recommendation statement, the thesis author carried forward much of the bias from individual lectures; and had to infer the statement's contextualisation, main objectives, and conceptual direction from transcripts of thematically limited discussions. Demonstrating that indeed no consensus existed, the internal review of the author's draft manuscripts assumed a pivotal role in the process. Fundamental recommendations were discussed controversially, mainly in how they aligned with a pragmatic trials philosophy of replicating clinical practice and

[10] Only addressed briefly on pp. 199 - 200, transcript day 1 (22 Oct), A Wasan in response to N Katz: "*[assay sensitivity] is not necessarily the point of a pragmatic trial. I think a pragmatic starts with treatments that have already demonstrated efficacy and already have demonstrated assay sensitivity, and then says how do you compare them to each other in a more real-world clinical context? So, I think it's a bit much to expect pragmatic trials to meet their objectives and also demonstrate all of these explanatory components. We have to assume some stuff is proven for that treatment. We have to assume something and then go forward and plan.*"; continued on pp. 209 – 213 without adding much content; also see pp. 137 – 138, pp. 153 - 158, pp. 160 – 170, pp. 179 - 193, day 2 (23 Oct), for further discussions addressing the potential tension between internal validity and pragmatic trial design. For panel discussions, see transcripts pages 153 – 215, day 1 (22 Oct), and pp. 105 – 149, day 2 (23 Oct), and for the final consensus discussion pp. 149 – 214, day 2 (23 Oct), all available at http://www.immpact.org/meetings/Immpact24/background24.html.

assessing interventions under real-world scenarios. In the absence of formal agreement mechanisms, the discontent of co-authors with a main interest in pragmatic trials carried through three draft manuscripts and two rounds of peer-review, gradually softening recommendations that emphasised limiting and measuring heterogeneity to enhance internal validity and interpretability, or that proposed methods not currently practiced in the field. The role of the thesis author in this process was mainly to negotiate between the poles of co-authors, trying to protect the interests of each group while finding positions that were acceptable to all. This commenced in the role of moderator of live panel and consensus discussions, and was finally accomplished via in-document comments, by small-group discussion via email with the main pragmatic trial authors, and through negotiation with Prof Dworkin who conveyed concerns from members of the core IMMPACT group.

Reflection on own role in guideline development

Although I had a central role in this process, I had no personal preference or investment other than to complete the process and contribute a piece of research that addressed the challenges identified in our systematic review. I recognise that I may have been biased by my initial exposure to explanatory ideas through Prof Dworkin and limited by an initial difficulty to orientate myself in the conceptually unclear field of pragmatic trials. By conducting and presenting a systematic review of methods, I was however able to anchor all discussions to an empirical description of the status quo in the field. Through my deliberately respectful and patient negotiation with all involved stakeholders, I believe that the process resulted in an IMMPACT statement which raises critical discussion points for pragmatic trials in pain. In the process, I learnt a lot about the value of rigorous consensus methodology. This lesson crucially informed the design of my consensus process on control interventions in efficacy trials presented in the previous thesis chapters.

Methods

The guidance development process is illustrated in Figure 10.

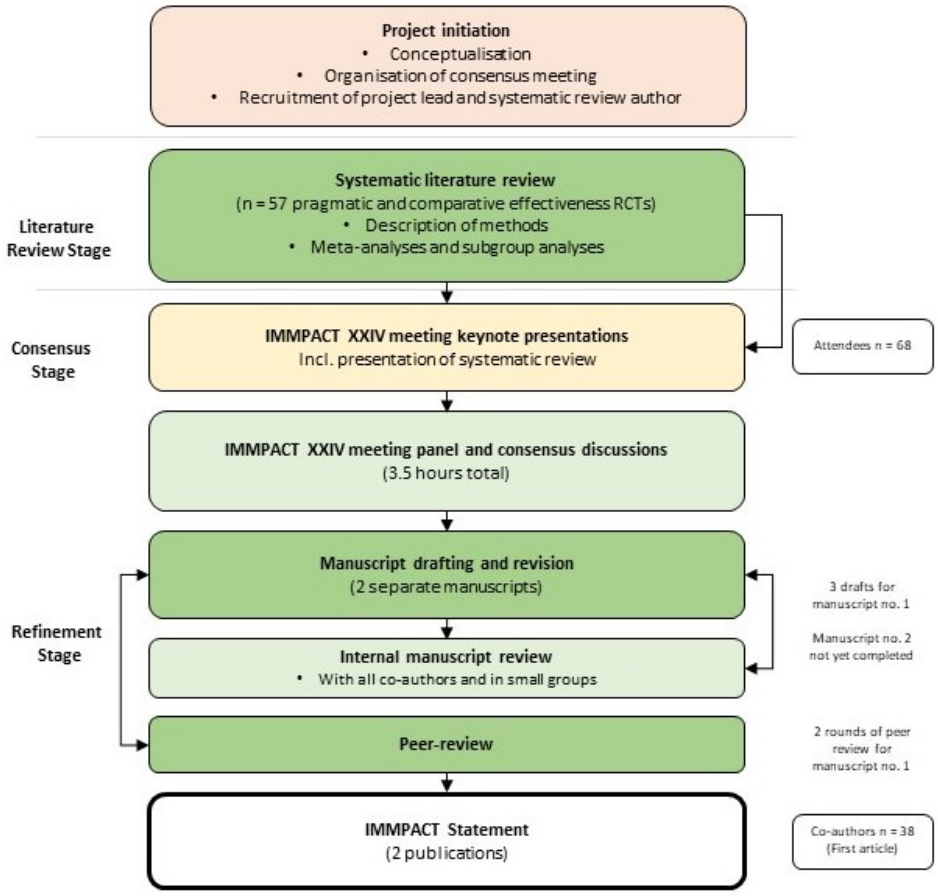


Figure 10: Flowchart of IMMPACT XXIV guidance development. The thesis author's level of involvement and responsibility is colour-coded: Red, no involvement; Yellow, partial involvement through a specific contribution; Light green, considerable involvement and responsibility; Dark green, leading role and sole responsibility.

On 22 and 23 October 2020, a videoconference consensus meeting was held by IMMPACT. Meeting participants were invited by the IMMPACT steering committee based on their expertise or experience involving pragmatic or pain-related clinical trials and to represent stakeholders from patient organisations, public institutions (such as the FDA and the National Institute of Health), and industry. In addition, all members of the ACTTION management, steering, executive, and oversight committees were invited. The meeting's objectives were to

discuss important considerations and provide best-practice suggestions regarding the design, implementation, interpretation, and evaluation of pragmatic clinical trials of pain treatments to inform the planning, conduct, and reporting of such studies.

Three consensus discussions were informed by nine 25-minute presentations by content experts and other academics. Presentations included the following topics:

- Definitions and general considerations (Prof Lynn DeBar)
- Statistical approaches (Prof Scott Evans)
- Lessons learnt from pragmatic trials in various settings (Prof Ajay J Wasan & Prof Robert D Kerns)
- Study population definition and patient recruitment (Prof John Markman)
- Study sites (Prof John Farrar)
- Concomitant and rescue treatments (Prof Michael Rowbotham)
- Outcome domains and measures (Prof Matthew J Bair).

Further, the results of the preparatory systematic review were presented by the thesis author. All participant details, lecture slides and meeting transcripts are available on the IMMPACT web site (http://www.immpact.org/meetings/Immpact24/participants24.html).

Following the meeting, the thesis author categorised the presented materials and meeting discussions into general and specific considerations, resulting in the plan to publish two manuscripts due to the number of topics covered. This decision was made in cooperation with the IMMPACT steering group. Starting with general considerations, the author drafted a consensus manuscript that was then reviewed iteratively by the designated co-authors (i.e., all meeting attendees who expressed interest). Once all designated co-authors agreed, the manuscript underwent peer review at PAIN, followed by revisions and further co-author review. The second manuscript is to contain pain-specific considerations. While this has been drafted, the second manuscript is currently undergoing internal review (as of 24 March 2023). Results of both manuscripts are presented below in combined form, with a brief intermittent paragraph prior to the second statement outlining methodological differences in their generation.

Results

The first of the two planned articles was published in the journal PAIN under the title "Research objectives and general considerations for pragmatic clinical trials of pain treatments: IMMPACT statement" (Hohenschurz-Schmidt et al., 2023a). The final IMMPACT statements were not termed 'guidelines' or 'recommendations' due to the lack of formal methodology (Moher et al., 2010b), but rather 'consensus statements of best-practice considerations'.

General considerations

Research questions and objectives for pragmatic trials

The IMMPACT statement highlights the need to clearly define the study intention, as the intention determines which design choices are most appropriate (Nicholls et al., 2020).

The IMMPACT co-authors agreed on the following scenarios under which a pragmatic attitude to trial design is considered appropriate. First, to assess whether a given therapy works as well as, or better than, established care when studied in a broad population and in non-academic settings (Cherkin, 2021). Such research questions were thought most pertinent for therapies with an established efficacy and safety profile. In fact, "*substantial evidence of the efficacy*" (p. 2342, Kerns et al., 2019) is considered a prerequisite for funding pragmatic trials by a large public Collaboratory in the US (https://painmanagementcollaboratory.org/pragmatic-studies/). However, the IMMPACT statement acknowledges that, in reality, pragmatic research questions are also formulated for therapies that have limited evidence of efficacy but are widely used in clinical practice. This applies to many complementary and integrative therapies (Gordon et al., 2020b; Hohenschurz-Schmidt et al., 2022c; Williams et al., 2021). Reflecting this, the statement holds that "*[w]hether [pragmatic trials in the absence of high-quality efficacy data] are appropriate depends on the research question and trial context*" (p. 4) (Hohenschurz-Schmidt et al., 2023a). Challenges of designing 'sham' control interventions and the study of widely established safe treatments are scenarios where comparative effectiveness designs "*can be*

considered" (p. 4). Nonetheless, IMMPACT points to the CoPPS Statement to indicate that specific guidance is becoming available to overcome this challenge for nonpharmacological trials.

Apart from testing comparative effectiveness, the IMMPACT statement lists the following possible objectives of pragmatic trials:

- Assessing different treatment-delivery strategies (e.g., stepped, stratified or matched care) (Gilron et al., 2020; Keefe et al., 2022).
- Testing effectiveness in different care settings and populations (from those in which interventions had already been studied).
- Evaluating differential effects for patient subgroups, phenotypes, and other questions aimed at personalised care.
- Questions of risk-benefit, cost-effectiveness, and other clinically relevant composite outcomes.

'High-quality' pragmatic trials

Under the heading 'Meeting trial objectives with high-quality designs', the IMMPACT statement postulates that "*study objectives should be achievable using the proposed trial methods, whether that means that the trial is closely aligned to typical clinical practice or not*" (p. 5, Hohenschurz-Schmidt et al., 2023a), thus pointing to the potential tension between external and internal validity and the implications for trial design. Indeed, the statement calls for a balancing of pragmatic and explanatory qualities, with the aim of producing clinically meaningful, relevant, and applicable results that are nonetheless reliable; the guiding principle being that "*[r]esearchers may emphasize generalisability when required by the research question but should preserve internal validity as much as possible*" (p. 5). PRECIS-2 was proposed as a useful tool to consider internal and external validity for each of the tool's domain, but researchers should additionally draw on other available tools to specifically evaluate internal validity (such as the Cochrane Risk of Bias tool; Sterne et al., 2019).

Optimising external validity

PRECIS-2 was widely acknowledged amongst meeting participants to be a useful tool in the design process of pragmatic trials. It was nonetheless felt that several methods would warrant consideration to align trials with normal clinical practice or enhance generalisability. These were:

- Use of real-world data (RWD) for eligibility criteria definition and recruitment
- Considering combined (e.g., risk-benefit) as well as responder and other subgroup analyses in addition to primary analyses
- Simplifying outcome choice, such as using measures with few scales as opposed to multiple-question disability questionnaires
- Employing real-world data collection tools, including consideration of wearables and mobile data sampling

In the main bodies of the IMMPACT manuscripts, each of those methods was discussed in relation to pain trials, alongside methods for each PRECIS-2 domain.

Optimising internal validity

Implementing the principle of considering both internal and external validity, the first IMMPACT manuscript then focuses on methods rarely discussed for their potential in pragmatic trials: Randomisation, blinding, and a list of other relevant bias minimisation methods.

Randomisation

Although most trials of the preparatory systematic review employed patient-level randomisation, stratification by site and randomisation of treatment clusters were also common (Hohenschurz-Schmidt et al., 2022c) (Table 20 in previous chapter). Cluster randomisation was discussed in more detail for its benefits and potential challenges. Less common but thought promising to answer research questions directly relevant for clinical decision-making were:

Including patient preference during randomisation processes, enrichment methods, and employing stepped-wedge designs. In the IMMPACT statement, these were presented as suggestions for consideration rather than recommendations, and literature with detailed methodological advice was referenced (Hohenschurz-Schmidt et al., 2023a).

Blinding

Blinding is at the heart of an ongoing controversy amongst (pragmatic) trials methodologists (Anand, et al., 2020; Christian et al., 2020; Dal-Ré et al., 2020; Wartolowska et al., 2022) and was controversial at the IMMPACT meeting. The agreed preamble is thus worth quoting in this thesis as an important contribution to contemporary debate:

"*For many pragmatic research questions, it is accepted that the non-specific (e.g., contextual or placebo) effects form part of treatments' real-world effectiveness. From this perspective, blinding may not be appropriate. Furthermore, a clinical decision may be between multiple interventions with doubtful evidence of efficacy but with different risks for harm or healthcare cost. These interventions are often well-established and commonly used in clinical practice. In this situation, non-blinded comparative effectiveness trials can answer important questions, while not negating the usefulness of improved efficacy research. Overall, we recommend blinding participants to group allocation where efficacy data are inconsistent and where compatible with trial objectives*." (p. 6, Hohenschurz-Schmidt et al., 2023a)

The statement then goes on to propose alternative methods to blinding participants to group allocation for when this is not possible (such as blinding to study hypotheses or measuring treatment expectations), and related techniques (e.g., having patients enter outcome data electronically to prevent assessor bias).

Other techniques for *minimising bias in pragmatic trials*

In this section, the thesis author again had to negotiate between members of the co-author group with conflicting views. The resulting paragraphs on techniques for bias

minimisation walk the thin line between advocating for more rigour in pragmatic trials while not disregarding the need for trials that are aligned with clinical practice, facilitating translation and implementation of evidence-based treatments. Co-authors felt that first it had to be acknowledged that "*[a]lthough randomization is commonly used to reduce bias in pragmatic trials, other bias-reduction methods such as treatment standardization and blinding study participants to treatment may conflict with pragmatic trial objectives.*" (p. 7, Hohenschurz-Schmidt et al., 2023a). The thesis author developed a table of a range of possible biases in clinical research and proposing methods to minimise bias in pragmatic trials of pain treatments (adapted here as Table 25). These methods were selected to be largely compatible with pragmatic trials objectives (i.e., not interfering with clinical practice in a manner that reduces generalisability of findings). Still, some co-authors requested the cautioning statement that "*the possible solutions for bias control in table 2 need to be examined for potential conflicts with a trial's pragmatic objectives. In this case they may not be suitable or their implications for the generalizability of findings should be declared.*" (p. 7). Notably, bias from differential expectations and exposure to contextual effects (i.e., from the placebo effect) was removed from the table during revisions due to concerns from co-authors regarding the compatibility with pragmatic trials objectives.

Table 25: Possible bias in clinical research and proposed considerations for methods to minimise bias in pragmatic trials of pain treatments. Potential biases are listed and defined in the left-hand column and potential approaches to minimise each bias in the context of pragmatic trials on the right (adapted from (Hohenschurz-Schmidt et al., 2023a).

Bias	**Possible solutions (and explanation)**
Recruitment bias *Recruiting predominantly or failing to recruit certain subgroups of eligible participants*.	• Enhanced recruitment / targeted recruitment strategies In pragmatic trials, recruitment bias is a problem when the trial fails to represent the clinical target population. If the research question requires a relatively representative sample, more effort may be required to recruit diverse participants (Ali et al., 2022b) even if recruitment methods no longer reflect standard clinical practice.
Selection bias	• Cluster randomised trials: site randomisation after participant recruitment

Selection of study participants skewed by factors such as participant characteristics (similar to recruitment bias but mainly driven by study staff).	• (Partial) Blinding of recruiting study staff (Christian et al., 2020) • Monitor baseline differences / Control for important covariates in the analyses
Allocation bias *Biased allocation of participants to study arms* (Higgins et al., 2011).	• Effective allocation concealment (Sterne et al., 2019)
Assessor bias *Knowledge of treatment allocation that influences outcome measurements*.	• Blinded outcome assessment (Christian et al., 2020) • Use of objective outcome measures (Christian et al., 2020), for example, actigraphy (Yu et al., 2021) • Use of disability and quality of life outcomes • Employ multiple follow-up assessments While outcomes in clinical practice are typically evaluated by providers, this is rarely necessary for pragmatic research questions. Therefore, bias control should be considered.
Attrition bias *Asymmetrical participant loss between study arms for non-random reasons* (Page and Persch, 2013).	• Assess risk during pilot phase • Monitor reasons for attrition • Include patient preference Although low adherence to treatments is common in clinical practice, it may undermine the interpretability of pragmatic trial results. This risk needs to be weighed against the relevance of employing low-touch ('pragmatic') strategies to increase adherence. Methods to evaluate reasons for participant loss do not interfere with pragmatic research questions and should thus be implemented, especially if adherence is not promoted.
Biased interpretation and reporting of results *Reporting bias typically refers to the selective reporting of positive results. Apart from*	• Evaluate overall internal validity • Pre-register trial and follow protocol • Accurately report non-significant results in superiority trials (not claiming comparable effectiveness) • Discuss limits of generalisability and avoid over-generalisation of findings

withholding negative results, alternative analyses can be performed (Gluud, 2006)*, and results can be misinterpreted or misrepresented* (Boutron et al., 2019; Draper-Rodi et al., 2022b).	• Adhere to reporting guidelines (e.g., Zwarenstein et al., 2008) Due to potentially greater heterogeneity, pragmatic trials may require more extensive reporting and more nuanced discussion than explanatory RCTs. This includes providing relevant contextual information to judge generalisability (Glasgow et al., 2001; Zwarenstein et al., 2008; Huebschmann et al., 2019), ensuring that claims are supported by data, and discussing study limitations.

Expanding the methodological arsenal of pragmatic trials

At the IMMPACT meeting, several trial designs were presented for their potential to answer clinically directly relevant research questions. For example, such questions referred to stepped care, personalised care, or the management of patients upon treatment failure. The proposed designs often required methods beyond simple randomisation of participants or clusters and deviated from the traditional two-armed parallel-group RCT design of most pragmatic trials (Hohenschurz-Schmidt et al., 2022c). Hence, there was considerable pushback from IMMPACT co-authors and manuscript reviewers, regarding them as "*broadly inconsistent with a pragmatic philosophy*" or "*ill-suited to answer pragmatic questions*" (reviewer comments). The focus on the 'pragmatic' research question as opposed to alignment with usual care was discussed controversially amongst co-authors, stating that such focus "*fundamentally disregards the impetus for conducting [pragmatic trials] in the first place. The emphasis on pragmatic trials […] emerged because the pipeline of efficacy trials to use of [evidence-based treatments] in everyday clinical care was getting stalled and approaches with established efficacy […] were not getting used. Thus, stepping around clinical practice elements more often than not won't help in translating [evidence-based treatments] into practice and, as such, misses the fundamental need for these types of trials in the first place*." (co-author comment during revisions). Cross-over, adaptive, and enrichment designs were most controversial in these regards.

In the end, such alternative designs were presented briefly and as 'suggestions' only, focussing on design elements that can or are already implemented in clinical practice (such as 'enrichment' of study samples through easy-to-use prognostic questionnaires such as STarT-Back). The possibility for such methods to reduce the applicability of results to a broad patient group was clearly highlighted, and general limitations identified. Nonetheless, co-authors agreed to defend these methods' usefulness against reviewer comments as these methods were thought able to produce results directly applicable to clinical and health policy decision-making or increase trials' efficiency while often not interfering with generalisability (as in cross-over designs). Under these premises, the following designs were presented: Large, simple trials (LSTs), cross-over designs, the stepped wedge cluster randomised design, enrichment designs and methods, and adaptive and other designs responsive to accumulating trial information (summarised in Table 40 in Appendix 3: IMMPACT Statement, additional content).

Example trials

Finally, the first IMMPACT article presented two pain trials as exemplary for their methods balancing internal and external validity, and to illustrate practical challenges and research constraints that can lead to methodological compromise. The thesis author informed this selection by screening the pool of studies included in the preparatory systematic review for these criteria, and co-authors also proposed one trial. Agreeing on exemplary trials required considerable negotiation from the thesis author. Beard et al.'s (2018) sham-controlled, three-armed, self-declared 'pragmatic' trial was selected for enhancing internal validity through the use of a sham control and other methods while embedding the trial within normal UK shoulder surgery practice. A pragmatic comparative effectiveness design was appropriate because shoulder arthroscopies were routinely performed in clinical practice. As there was insufficient evidence of efficacy, also including a sham control arm safeguarded against placebo effects. This trial also illustrated the benefits of pre-planned 'unpragmatic' secondary analyses (i.e., a per-protocol analysis), in this case assessing the effects of intervention non-adherence, which was high.

The second exemplary study compared mindfulness-based stress reduction (MBSR) to cognitive behavioural therapy (CBT) and to usual care for patients with chronic low back pain (Cherkin et al., 2016). Here, the thesis author had access to the direct experience of the trial's principal investigator who was an IMMPACT co-author. Rating the study with the PRECIS-2 tool illustrated the 'balance' achieved by the design: The researchers prioritised clinically relevant outcome measures, pragmatic data analysis (intention-to-treat), and low study questionnaire burden, while employing more explanatory methods in other domains, for example limiting the flexibility of intervention delivery, as well as using targeted participant recruitment and dedicated trial centres. Confirming the relevance of the IMMPACT statement's direction of thrust, the RCT's lead investigator then clarified that these methodological choices were partly made to enhance interpretability (e.g., by limiting and monitoring how interventions were delivered) and partly due to practical constraints (e.g., having to meet recruitment targets or requirements from funders).

Pain-specific considerations for pragmatic trials (second IMMPACT statement)

Preliminary methodological and contextual considerations for the second IMMPACT statement on pragmatic trials

The following section is based on meeting presentations and discussions which the thesis author compiled, extended, and contextualised with current literature. At the time of writing, the material is undergoing review with designated co-authors, aiming to publish the manuscript as a second IMMPACT statement on pragmatic trials. The thesis author is also discussing further consensus methods with the IMMPACT steering team, proposing to conduct a focus group with subject experts and patient partners, as well as anonymous voting on controversial formulations. Specifically, the thesis author has proposed to:

- Pre-register the below methods as a protocol at www.osf.io
- Gather comments on draft manuscript as outlined below

- Compile an agenda for focus group discussions consisting of controversial manuscript items and open-ended questions
- Invite all 2020 IMMPACT meeting attendees to attend an online focus group (expecting 10-15 people to attend, who contributed substantially to discussions during the first manuscript review)
- Conduct the focus group using accepted methodology (Ritchie et al., 2013; Holloway and Galvin, 2016) and modified with RAND/UCLA-type rating procedures of item appropriateness (Fitch et al., 2001)
- Report outcomes transparently, using the COnsolidated criteria for REporting Qualitative research (COREQ) (Tong et al., 2007)

Notably, the thesis author does not consider the use of a formal consensus method such as a Delphi survey (Jünger et al., 2017) useful or feasible in this instance (although this might be helpful for IMMPACT endeavours on less controversial topics) (Shrier, 2021). As the development of the first manuscript on pragmatic trials demonstrated, the field is too divided to aspire to consensus.[11] Instead, a modified focus group method appears appropriate to capture the different viewpoints, while producing guidance that can differentiate between items that are always appropriate to implement and others that should be considered under certain circumstances (as will be achieved by rating without implementing consensus thresholds). The author recognises their own dominant role and potential for bias from writing the draft manuscript and thus informing focus groups; For example, aligned with their understanding of their role as lead author, the thesis author has added paragraphs on stakeholder involvement and ethical considerations for discussion during focus groups. These topics had been absent from

[11] In a recent publication, Ray Pawson (2019) demonstrates convincingly that pragmatic trials could even be considered an 'essentially contested concept' due to i) its inherent complexity and existence of multiple sub-themes, ii) considerable contestation of subthemes, bringing into question the concept itself, iii) various stakeholders attributing different orders of relative importance to individual subthemes, and iv) the existence of incompatible views amongst various interest groups or groups of experts. Pawson only "*stop[s] short - just short - of terming the pragmatic trial an* essentially *contested concept*" (author's emphasis, p. 952) because he attests some movement towards an improved common understanding and application of the concept in recent years. During this process, "*the concept of the pragmatic trial has itself shifted markedly […] —from high aspiration to close operationalisation and from big idea to more circumspect claims*." (p. 952). Our first IMMPACT statement can likely be viewed as another realist approach to the topic.

the original meeting but are prominently discussed in contemporary pragmatic trials literature (Kalkman et al., 2015; Taljaard et al., 2018; Hung et al., 2017; Nicholls et al., 2019; Bastian et al., 2020; Keefe et al., 2022).

In the following sections, the author will outline the arguments they synthesised from the IMMPACT meeting and current literature and applied to the field of pain research. For the purpose of this thesis, the author added specific considerations for pragmatic trials of physical, psychological, and self-management (PPS) interventions. By drawing on findings of the systematic review (Hohenschurz-Schmidt et al., 2022, and chapter 4), it was ensured that the proposed trial designs address existing challenges in the field. Unless identified otherwise through reference to the IMMPACT meeting or to existing publications, the following are to be considered the author's own scholarly suggestions for the field rather than consensus recommendations. For publication, these considerations will likely change through discussions with IMMPACT co-authors.

Ethical considerations for pragmatic trials in pain research

This topic has not been discussed at the 2020 IMMPACT meeting, and this paragraph represents the author's own account of potentially relevant considerations.

Ethical and regulatory frameworks for clinical trials were primarily designed for tightly controlled efficacy studies. However, specific ethical considerations for pragmatic trials exist (Sugarman and Califf, 2014; Goldstein et al., 2018). In Table 26, these considerations are summarized and applied to pragmatic RCTs involving individuals in pain.

Additionally, one should consider whether conducting a pragmatic trial is ethical as opposed to a more explanatory study, which depends on the availability of high-quality efficacy research and the nature of the research question (Keefe et al., 2022; Hohenschurz-Schmidt et al., 2023a). Further, the aim of the IMMPACT statements on pragmatic trials is to promote trials that

are optimally designed to achieve their objectives, which is a prerequisite for ethical research in light of preventing research waste (Pirosca et al., 2022).

Multicentre pragmatic trials present particular regulatory challenges for review boards (O'Rourke et al., 2015). Conversely, a lack of standardised ethics procedures across institutions can be a barrier to conducting such trials (Eapen et al., 2014). In pain research, establishing large research networks can help standardise and centralise procedures, but challenges still remain (Ali et al., 2022a).

Table 26: Ethical considerations in pragmatic trials and in pragmatic trials for pain. The first column is largely based on Goldstein et al. (2018).

Ethical consideration in pragmatic trials	Application to pain trials and potential solutions
Do trials require the same ethical and regulatory oversight as efficacy trials if they test 'usual practice' and established interventions? This dilemma is known as the 'research/treatment distinction'. Some authors purport that this distinction blurs in pragmatic trials, reducing the need for ethical oversight (Kass et al., 2012; Faden et al., 2013). Conversely, if routine clinical practice or aspects thereof are changed to implement the trial, ethical assessment is required.	- Most pragmatic trials in pain will impose additional requirements on patients and medical personnel (such as additional outcome collection) (Hohenschurz-Schmidt et al., 2022c), for which ethical approval should be obtained. - Targeted recruitment of potentially vulnerable people is also common in pain research (Hohenschurz-Schmidt et al., 2022c; Ali et al., 2022b), requiring careful ethical assessment. - Ethical approval can be streamlined for all other aspects of a trial. For example, interventions may not require detailed review if they are not modified for the trial or are not directed at patients (Califf and Sugarman, 2015).
Is consent required in trials that test 'usual practice' and established interventions that are known to be safe? According to the principle of patient autonomy, voluntary informed consent is required for human involvement in any experimentation. If only regarding consent as necessary to inform about added risks of experimentation, consent may be waived or simplified in pragmatic trials between low-risk interventions for which there is equipoise. This is, however, controversial and defining 'low risk' is challenging (Lantos et al., 2015; Nicholls	- Consent can be waived or simplified under certain circumstances, many of which may apply to pragmatic trials of pain research (Council for International Organizations of Medical Sciences, 2016; Goldstein et al., 2018; Nicholls et al., 2019). - Especially nonpharmacological interventions are believed to hold little medical risk, and streamlining of consent has been called for (Kass et al., 2012; Faden et al., 2013). These interventions are often studied in pragmatic pain trials (Hohenschurz-Schmidt et al., 2022c). - Risk assessments in pragmatic pain trials should consider patient subgroups with different risk

et al., 2019); Also controversial is the need for disclosure of study methodology, including randomization (Goldstein et al., 2018).	profiles, for example depending on comorbidities or levels of health literacy (Nicholls et al., 2019). - Risk/benefit assessments may have to include consideration of other undesirable effects from testing and implementing non-efficacious interventions and low-value care, such as nocebo, behavioral, and socioeconomic effects (Hohenschurz-Schmidt et al., 2022d).
How can pragmatic trials promote justice and equity in trial research? Pragmatic trials are often considered opportunities to involve people in research that are not usually participants and to conduct research relevant to minority groups. However, involvement of such groups may also require additional methods to protect potentially vulnerable participants (Welch et al., 2015; Nicholls et al., 2019). Pragmatic trials, often embedded within health systems, also risk perpetuating their existing structural injustices (Ali et al., 2022b).	Ali et al. (2022) proposed strategies for addressing injustices and inequities in pragmatic pain research. Strategies involve: - consideration of socioeconomic trial context - effective and equitable stakeholder engagement - broad criteria for participating centers - recruitment of underserved and vulnerable populations even if additional monitoring may be required - flexibility and tailoring of interventions to subgroups - accessible data collection methods - promotion of diversity within the research team itself - ensured accessibility for populations with limited technology access or literacy (Ali et al., 2022b).

Patient and other stakeholder involvement in pragmatic trials

This topic has also not been discussed at the 2020 IMMPACT meeting, with this paragraph representing the author's own account of potentially relevant considerations.

Research questions for pragmatic trials are often considered 'stakeholder-driven' (Selby et al., 2015) or 'important for real-world decision-making' (Loudon et al., 2015a). As a result, early stakeholder engagement is crucial (Keefe et al., 2022). Identifying and involving relevant stakeholders holds the promise that the 'right' problems are addressed in a manner that is efficient and translatable to clinical practice or healthcare policy (Holzer et al., 2022). Helpful frameworks are available to guide stakeholder engagement throughout the research cycle (Hung et al., 2017), notably the Patient-Centered Outcomes Research Institute (PCORI) Engagement Rubric (Sheridan et al., 2017).

In the field of pain research, large institutional networks for the conduct of multicentre pragmatic trials have been established through extensive stakeholder engagement with a focus on developing personal relationships. For example, the Pain Management Collaboratory (PMC3), a joint project between three US government agencies, employs inwards- and outwards-facing mechanisms to optimize communication with key stakeholders (e.g., archiving of relevant information internally, and an outwards-facing website for stakeholders and the public). The PMC3 has a board of stakeholders affiliated with key government and nongovernment entities, and a 'patient resource group' able to support the Collaboratory as a whole as well as individual trials (Ali et al., 2022a). Further insights van be found in Bastian et al., (2020) and IMMPACT guidance for stakeholder engagement in pain trials is being developed (Holzer et al., 2022). (Also see Holzer et al. (2022) for a review of current methods of patient engagement in pain trials.)

Study treatments

In this section, based on several IMMPACT presentations and pertinent literature, considerations relating to the choice and implementation of study interventions are presented. Considerations include treatment standardisation, the duration of treatment phases, fidelity monitoring, and requirements for patient adherence.

Pain management is rarely standardised in clinical practice (Somerville et al., 2008), and a biopsychosocial approach to pain calls for treatment individualisation and interdisciplinarity, often involving one or more PPS intervention (Pizzo et al., 2011; Nicholas, 2022). Pragmatic trials represent both an opportunity and a challenge in this context: An opportunity as they allow for the study of interventions as delivered in clinical practice and irrespective of our understanding of treatment mechanisms (MacPherson, 2004); and a challenge because some oversight over the content of study treatment(s) facilitates the scientific interpretation of trials and assessment of generalisability (Huebschmann et al., 2019). Further, pragmatic trial designs lend themselves to the study not only of comparative effectiveness, but also to the investigation of treatment combinations, interactions, and tailoring strategies (for example through predictive enrolment,

augmentation or switching; although modifications of the common two-arm parallel design may be required (Hohenschurz-Schmidt et al., 2023a)).

Considerations underpinning the study of a particular therapy in a pragmatic trial need to be made explicit and the supporting evidence base reported clearly (Zwarenstein et al., 2008). Such evidence should include supportive efficacy and / or mechanistic research for the treatment or core components of a treatment package. Finally, intervention design for pragmatic trials needs to be based on a good understanding of how the intervention is applied in routine clinical practice. This can be obtained, for example, through practice surveys, and development ideally involves practitioner and other stakeholder engagement.

Although uncommon in real-world clinical pain management, authors have repeatedly called for some standardisation (or protocolisation) of study treatments in pragmatic trials, monitoring of fidelity, and detailed reporting of intervention content (Zwarenstein et al., 2008; Keefe et al., 2022; Kerns et al., 2022). Especially in trials of nonpharmacological interventions, provider expertise and 'style' of practice require consideration, balancing the need for a representative range of skillsets with acceptable levels of heterogeneity in treatment delivery (Keefe et al., 2022). The extent to which treatment flexibility needs to be controlled depends on the research question and the specifics of the intervention. Notably, dosage safety and therapeutic ranges are important to consider in drug trials, while the potential to introduce undesired heterogeneity through practitioners and intervention complexity is more pertinent in nonpharmacological therapies. However, any standardisation of treatment delivery, for example through trial-specific training and protocolisation, ought to be conservative and not exceed what can be expected in routine practice of a given intervention. Defining core intervention components and non-permissible modalities as a minimal step is advisable in order to promote fidelity and clearly differentiate trial interventions from one another (Kerns et al., 2022) (also see discussion of concomitant treatments below).

Where possible, fidelity monitoring may be preferable to protocolisation, seeking to document how the treatment is delivered during the trial. This can be done through a variety of, ideally low-burden, methods such as observing treatment delivery in a subset of treatment sessions, reviewing treatment documentation, or collecting patient-reported data on treatment adherence. High-quality fidelity (and adherence, see below) data may allow for sensitivity analyses to develop hypotheses about subgroups, mechanisms, intervention tailoring, or implementation challenges. Further considerations regarding treatment fidelity in pragmatic pain trials of nonpharmacological interventions are published (Kerns et al., 2022).

The duration of the treatment phase should be decided based on the characteristics of the pain, the treatment being evaluated, and the research question. Considerations are also intimately linked to how interventions are usually used in clinical practice, which outcomes measures are employed, and what patients expect from the intervention (considerations that likely benefit from patient involvement). In trials with people living with persistent or recurring pain, realistic interventions may range from intensive one-off treatment courses to continuous ('maintenance') care. Different pragmatic spinal manipulation trials illustrate these considerations well: For example, Goertz et al. (2013) enrolled young, otherwise healthy military personnel with acute low back pain, providing up to eight treatment sessions over four weeks. Conversely, Eklund et al. (2018) enrolled people with persistent or recurrent low back pain who had responded to an initial course of chiropractic, providing regular or symptom-driven maintenance care.

Patient adherence

At the IMMPACT meeting, attendees advocated for promoting adherence in line with best practice (not *normal* practice) in the respective therapeutic field to minimise missing data. This contrasts with PRECIS-2, which disincentivises encouragement or formal adherence requirements. In line with the emphasis of the first IMMPACT statement, however, encouraging patient adherence can be considered an opportunity to promote trials' internal validity with little

interference with most pragmatic research questions. If, however, acceptability and adherence levels in normal practice are part of the research question, adherence should not be promoted beyond normal practice.

As with intervention fidelity, treatment adherence should be monitored in all cases to facilitate interpretation and allow for corresponding sensitivity analyses. While often implemented in pain trials (Kerns et al., 2022), relying on retrospective assessment or patient self-report needs to be weighed against the need for high-quality adherence data. In some circumstances, technological tools could be developed help to corroborate patient self-report, for example user data from apps or activity tracking through wearables.

If participant adherence is disregarded, important information about the treatment's real-world acceptability is lost. Collecting qualitative information on reasons for non-adherence may supplement such assessments. Further, patient adherence could figure as a secondary study endpoint (discussed below under 'outcome domains').

Comparator and control conditions

Control conditions in comparative effectiveness designs

The efficacy-focused attendees of the IMMPACT meeting pushed for the promotion of control conditions (such as no-treatment, placebo, or treatment-as-usual). In comparative effectiveness trials, a basic level of efficacy and safety of the treatments under investigation is usually assumed. In trials without a control condition, however, this claim is only inferential, and attendees argued for this to be acknowledged as a limiting factor. Evidence to suggest that the compared treatments are efficacious and safe should be presented along with the level of recommendation (e.g., as per GRADE (Guyatt et al., 2011)). Uncontrolled trials should be avoided if the underlying efficacy research for the comparator is poor or, generally, in non-inferiority and equivalence designs (Food and Drug Administration, 2016), although this does not reflect current practice (Hohenschurz-Schmidt et al., 2022c).

Even in pragmatic superiority trials where the comparator is an established analgesic or complex therapy, a further control group may be beneficial: In community practice, many medications are not titrated to effective dosages, misdiagnoses can occur (Di Franco et al., 2011), and treatment adherence varies (National Institute for Health and Care Excellence, 2009; Pottegård et al., 2014); Many nonpharmacological interventions for pain yield inconsistent results (Rubinstein et al., 2019) and are applied variably and by therapists from a multitude of backgrounds (Thomson et al., 2014; Froment et al., 2019). Without a no-treatment, placebo, or treatment-as-usual control group, one cannot conclude confidently that the study treatments were comparable. If trial results showed no difference between groups, such a control group would demonstrate that the trial was able to distinguish an effective treatment from a less effective one, thus strengthening the conclusions drawn. It is acknowledged that the addition of a further trial arm holds practical and economic challenges for the overall trial but would urge investigators to carefully consider whether this is nonetheless required so that the trial does not result in uninterpretable findings. If no control group is implemented, this decision should be justified in the trial protocol or report.

Treatment as usual / Usual care

Twenty-five per cent of contemporary self-declared pragmatic or comparative effectiveness trials of pain therapies employ treatment or care as usual (TAU) as comparator, occasionally adding extra components such as advice or educational sessions. About 70% of those trials provide information on the potential content of TAU and most of those also collect data on the interventions actually received by patients in this group (Hohenschurz-Schmidt et al., 2022c).

'Usual care' is considered the 'comparator of choice' by many pragmatic trial methodologists, although operationalising this in a trial is fraught with challenges (Zuidgeest et al., 2017). Sometimes called 'standard care' (there is no consensus on terminology and interpretations differ (Nicholls et al., 2019)), what constitutes TAU differs widely, even within the same country, and may change (Smelt et al., 2010; Zuidgeest et al., 2017) – sometimes rapidly as last seen with the widespread adoption of telehealth during the Covid-19 pandemic (Gilbert et

al., 2020; Malliaras et al., 2021). Further, clinical recommendations and actual usual practice may differ widely so that TAU cannot be inferred from the consultation of clinical guidelines (González-Urzelai et al., 2003; Williams et al., 2010; Zuccaro et al., 2012; Sekhon et al., 2013; Tournebize et al., 2016; Ladeira et al., 2017). It is therefore paramount to describe in detail what the TAU condition entailed in the context of each trial (Smelt et al., 2010), which may warrant monitoring of the interventions actually received. Pascoe et al. (2022) proposed six aspects of TAU which should be reported in back pain trials, which were modified and endorsed by IMMPACT for pain trials more broadly:

- Who provided care (incl. profession, training, and experience)
- If and what type of self-management education or advice were provided (and in which format)
- Whether physical activity and movement were addressed and how
- If and what medications were used
- The dosage of any interventions received
- Whether care was consistent with current guidelines

Because of their potential to introduce undue heterogeneity, concomitant treatments will be further discussed below and similar considerations apply to variability detected in the TAU arm.

Waiting list controls

In waiting list conditions, patients expect treatment at a later timepoint (as opposed to 'true' no-treatment controls, where they are informed that they will not receive treatment as part of the study). Waitlist conditions can control for the natural course of disease and regression to the mean, although not for placebo effects. Waiting for treatment may, however, contribute to symptom persistence in anticipation of future treatment (Lynch et al., 2008; Furukawa et al., 2014). The feasibility and credibility of waitlist conditions (i.e., resemblance or proximity to usual practice) depend on the healthcare system. In the UK, for example, long waiting periods are

currently normal in the National Health System (NHS), especially for 'non-urgent' referrals such as for persistent pain (The Health Foundation, 2019; Public Health Scotland, 2021), and may be used as realistic control condition. Nonetheless, waiting list controls are uncommon in self-declared pragmatic RCTs of pain treatments (Hohenschurz-Schmidt et al., 2022c).

No-treatment controls

In no-treatment control arms, patients are informed that they will not receive treatment as part of the trial, also not at a later point. This may frustrate patients and negatively impact the self-reporting of symptoms, trial retention, or the patients' condition by way of the 'frustrebo' effect (Power and Hopayian, 2011). Nonetheless, no-treatment groups do exist in current pragmatic trials of pain treatments (Hohenschurz-Schmidt et al., 2022c); Frustration effects may be attenuated by means of deliberate communication as part of the consent process, but further research is required here.

Concomitant and rescue pain treatments

In trials of pain interventions, concomitant treatments may play an important role, but their permissibility and use is poorly reported (Hohenschurz-Schmidt et al., 2022c). On one hand, allowing patients to continue taking their usual medications or seeing other providers of nonpharmacological care is reflective of normal practice and thus important for the generalisability of results. Prohibiting concomitant therapies may also disincentivise recruitment and hold ethical challenges. On the other hand, concomitant therapy use may undermine trial interpretability if unknown to the research team. 'Rescue' analgesia may similarly confound trial results, meaning intermittent drug intake for insufficient pain relief during a trial. Confounding can occur when study groups use rescue medication differentially (for example when one treatment is not effective) (Dworkin et al., 2010), through interactions with the study drug, or by producing rescue-related adverse events. The choice of rescue medication also matters, with highly effective drugs reducing effect sizes and ineffective drugs considered unethical. Notably,

concomitant and rescue treatments may also consist of nonpharmacological approaches or, even more subtle and difficult to quantify, behavioural changes.

All these considerations should inform triallists in the design of pragmatic trials and monitoring of concomitant and rescue analgesia. Given the large risk for confounding effects, parameters of permissibility and monitoring processes of actual treatment use should be established before conducting a pragmatic trial and implemented across participating centres. If the trial objective is to replicate clinical practice, one should attempt not to limit the use of concomitant therapies unless based on a strong rationale and considering the potential effect on generalisability. The use of concomitant and rescue treatments should, however, always be carefully documented and possible implications for treatment effects and patient subgroups analysed statistically (Dworkin et al., 2010). To date, this is rarely done, and "*poorly described procedures and incomplete reporting are likely to hinder the interpretation, critical appraisal, and replication of trial results*" (p. 3, Grøvle et al., 2020), also in pragmatic trials of pain (Park et al., 2020; Hohenschurz-Schmidt et al., 2022c).

Obtaining reliable data on rescue and concomitant therapy use is challenging. Possible approaches to data collection include monitoring of electronic health records, billing and insurance data, and patient or physician report. Sampling of these data also needs to be weighed off against excessive research burden, and has to be tested and possibly simplified for the purpose of individual trials. When reporting a pragmatic trial and having access to feasible data sampling methods, the following items are considered important (Table 27).

Table 27: Concomitant treatments and rescue medication in pragmatic analgesia trials - Design and reporting considerations. List adapted from Grøvle et al. (2020).

Was rescue medication permitted (yes/no)? For what reason?
Brand and generic name(s), formulation, and administration
Allowed doses and frequency
Consequences of exceeding allowed dosage (withdrawal, treatment failure)
Specific procedures (e.g., dosage or timing in relation to pain symptoms)
Delivery (prescription, over the counter)
Payer (patient, sponsor)
How quantified (dosage, use/no-use, days taking rescue medication)
Assessor of consumption (patient: self-report, investigator: pill count)
Used as outcome? (if so, primary, secondary or explorative)
Prespecified statistical analysis plan (if used as outcome) and pre-registration on public trial registry
Report of rescue consumption in each treatment arm
Statistical analysis (if used as outcome)
Discussion whether rescue medication influenced trial results (if so, how?)
The same considerations apply and need to be adapted to non-drug therapies and pain self-management strategies used as concomitant or rescue treatment. In addition, also document:
Profession and qualification of treatment provider
Therapy-specific detail (e.g., techniques and approaches)

Similar to medications, admissible and prohibited nonpharmacological concomitant therapies ought to be pre-specified and their use documented, including the type of therapy, its providers, intensity, duration, and frequency. Behavioural change may have to be assessed at follow-up, depending on the study question, focusing on physical activity levels, various pain coping behaviours, and psychological health; ideally using accepted questionnaires. Where changes in these parameters are not expected to form part of the treatment mechanism, mediation analyses may be required.

Patient population and study sites

Eligibility criteria

Eligibility criteria in pragmatic trials are typically broad, reflecting routine diagnostic procedures for a specific medical issue (Loudon et al., 2015a). In the field of pain, this will involve triaging for serious pathology and application of standard diagnostic procedures. Frequently in pragmatic analgesia trials, eligibility criteria refer to pain duration and localisation only, but common comorbidities or drugs impacted participant eligibility in a large proportion of recently reviewed trials (25 and 12%, respectively) (Hohenschurz-Schmidt et al., 2022c).

Depending on trial objectives, the diagnosis of the underlying disease may be warranted to ascertain eligibility. If the diagnostic standard involves more sophisticated technology than what is commonly available, this is acceptable for pragmatic trials but may limit generalisability to other settings or healthcare systems. The same consideration applies to 'prognostic recruitment', where patients are eligible based on prognostic profiles (such as chronification risk (Delitto et al., 2021; Hohenschurz-Schmidt et al., 2023a)). If applicable, authors should discuss if the employed diagnostic technology is expected to become more widely available and otherwise justify the decision to conduct a pragmatic trial for their research question.

Trial designers may be tempted to reduce heterogeneity by applying eligibility criteria beyond ascertainment of diagnoses or simple stratification methods, but these should be kept similarly broad, based on clinical reality, and justified in light of the pragmatic question under investigation. For example, a trial could focus on children, adolescents and adults, older adults and elderly, or a single sex. Further specification is only warranted if reflective of a well-delineated clinical population for which a clinically relevant and population-specific question needs to be answered. In these cases, pragmatic trials can be highly selective, for example only including persons of a specific underserved population within clearly defined geographic, age-related, socioeconomic, and cultural boundaries; and authors must discuss the clinical or policy relevance of their selection and implications for generalisability (Noting, however, that sacrificing

'pragmatism' in the recruitment domain may be warranted to enhance generalisability by inclusion of representative patient sample (Zuidgeest et al., 2017)).

If the research question demands a broad clinical sample of patients, the thesis author proposed the following pain-related domains for consideration. As major eligibility criteria, these would be more inclusive than what is usually found in analgesia trials. Rather than following generic advice, however, researchers are encouraged to think about what is warranted depending on their individual research objective. Apart from necessary diagnostic criteria, we recommend to:

- Consider including pain of any intensity, not just ≥ 4/10, balancing clinical relevance with risk for flooring effects.
- Consider including any pain duration that fulfils diagnostic criteria.
- Consider not excluding patients for previous treatment non-response unless aligned with the pragmatic research question.
- Consider not excluding for any medical or psychiatric comorbidities unless for safety reasons.

Whilst broader eligibility criteria increase trial heterogeneity, they also enable subgroup analyses which may be able to answer clinically relevant questions. In line with applicable CONSORT reporting recommendations (item 3), authors should discuss "…*the degree to which* [the eligibility criteria] *include typical participants and/or, where applicable, typical providers (eg, nurses), institutions (eg, hospitals), communities (or localities eg, towns) and settings of care (eg, different healthcare financing systems)*" (Zwarenstein et al., 2008). The IMMPACT statement extends this recommendation to encourage statistical subgroup assessment where possible.

Finally, pragmatic trials are almost absent for several common pain conditions, including headaches, neuropathic pain, fibromyalgia, and injury-related pain (Hohenschurz-Schmidt et al., 2022c). IMMPACT should encourage trials with pragmatic attitudes for pertinent research questions in these groups of patients.

Patient identification through electronic health records

In line with pragmatic research questions, patients are usually recruited from existing patient populations of participating trial centres. Researchers, however, need to identify the population ultimately targeted in clinical practice and to complete trials on time. To facilitate this, they may have to employ more active recruitment methods. Identifying eligible patients through records may serve these aims, especially if an electronic health record (EHR) system is available. EHRs come with various challenges, some of which specific to pain research. People in chronic pain, for example, may be coded with various diagnoses, lack a precipitating and identifiable event, and are thus captured only through elaborate algorithms or natural language processing tools which require careful testing (Tian et al., 2013; Koleck et al., 2019; Dave et al., 2022). In less severe conditions for which patients may rely on self-management or over-the-counter medication, these filters may not be feasible at all. Although a rapidly developing technology, establishing EHR systems is costly and laborious (Kruse et al., 2016), data quality can be poor, interoperability challenging, and data protection needs careful consideration (Cowie et al., 2017). If feasible, we encourage the use of EHRs for the definition of eligibility criteria and for targeted recruitment. Currently, researchers might have to be cautious not to over-rely on real-world data for purposes integral to trial integrity, specifically outcome assessment, unless the EHR system in question has been shown to be reliable, data quality high, and its use ethically justifiable.

Recruitment and retention

According to PRECIS-2, recruitment through patients' usual appointments and from multiple diverse clinics is most 'pragmatic' (Loudon et al., 2015a). However, insufficient recruitment strategies pose a risk to the reliability of trial results, as they may lead to underpowered trials and the omission of relevant patient subgroups. The latter is of particular concern in chronic pain, which is more prevalent in populations that are arguably more difficult to recruit into research trials or have limited access to healthcare in general, such as various

underserved populations, ethnic minorities, those with psychiatric comorbidities, and elderly people (Sjøgren et al., 2009; Campbell and Edwards, 2012; Queiroz, 2013; van Hecke et al., 2017; Harris et al., 2018; Karran et al., 2020). Provided these are potential or intended end-users of the study treatment, maximising generalisability may require targeted recruitment methods, and trial results could be used to advocate for improved access to effective care for people from the studied populations. Indeed, pragmatic trials of pain therapies obtained the lowest average rating for 'pragmatism' in the recruitment domain (Hohenschurz-Schmidt et al., 2022c), likely reflecting the above practical and scientific concerns (Hohenschurz-Schmidt et al., 2023a).

Incentives can enhance recruitment, including financial compensation or free access to otherwise self-paid care. Since self-payment may be a barrier to trial participation and may cause differential attrition in the 'undesirable' trial arm, compensating patients for treatment costs is appropriate (Walsh and Sheridan, 2016). In any other capacity, incentives may bias recruitment and attract 'professional' patients. Further, trial designers ought to consider the vulnerability of patients, both financial and personal, which may create ethical dilemmas. In the field of chronic pain in particular, the desire to find effective treatment options may be strong (Ciechanowski et al., 2003; Jonsdottir et al., 2015).

The risk of participant attrition may be increased in pragmatic trials due to their longer duration. Retention of participants is an important factor in the planning phase of pragmatic trials and must be considered proactively to minimise the impact of missing data. The use of imputation methods for missing data can only address the issue retrospectively (Cai et al., 2020), and recruiting more participants to compensate for anticipated attrition is not only costly but also does not reduce bias. Instead, stakeholder involvement and feasibility testing can help anticipate and mitigate potential retention problems, as various factors such as the treatment and research burden, associated costs, the prohibition of concomitant or rescue treatments, and the consent process may interfere with successful retention. It is crucial to emphasise the importance of staying in the trial once enrolled to improve the validity of the results (Duncan et al., 2020).

Sample size

Good pragmatic trials provide a power calculation, ensuring adequate statistical power for the detection of all important differences to be examined. CONSORT highlights that "*[if] calculated using the smallest difference considered important by the target decision maker audience (the minimally important difference) then [authors should] report where this difference was obtained."* (item 7, Zwarenstein et al., 2008). Effect sizes obtained in efficacy trials with more controlled circumstances may not be transferable to pragmatic trials. If there are no good indicators to estimate anticipated effect sizes, rather than inflating sample sizes unnecessarily, adaptive trial components may allow for the adjustment of enrolment depending on emerging data, as discussed in the first IMMPACT statement (Hohenschurz-Schmidt et al., 2023a).

Outcome domains and measures

Choice of outcome measures in pragmatic pain trials

Ideally starting with a minimal set of core outcomes, trial designers ought to consider how each additional measure may alter the clinical workflow (Welsing et al., 2017). In doing so, IMMPACT meeting attendees recommended a very sober attitude to what constitutes a feasible and realistic outcome tool.

Commonly, outcomes in pragmatic trials should be "*of obvious importance from the patient's perspective*" and "*relevant to the people who decide whether to implement the intervention on the basis of its [the trial's] results*." (p. 9, Loudon et al., 2015a). In analgesia trials, the four core domains of pain outcomes proposed by IMMPACT in 2003 are well-established, covering intensity, interference, function, and change (Turk et al., 2003; Dworkin et al., 2005). Several outcome measures address these domains and are relatively quick to administer in pragmatic trials: This includes numeric rating scales (NRS) for each domain or composite score generating measures such as the Brief Pain Inventory (BPI), its modification the ultra-short three-item PEG-scale, the Multidimensional Pain Inventory (MPI), and subscales of the Patient-

Reported Outcomes Measurement Information System (PROMIS®) or the Short-Form Health Survey 36 (SF-36) (Kroenke et al., 2019). How relevant these outcomes are to patients, and thus pragmatic trials, may, however, vary (Turk et al., 2008; Gardner et al., 2015), highlighting the need for trial-specific stakeholder engagement or the exploration of individualised outcome measures (Antunes et al., 2020). Apart from the PEG-scale, other ultra-brief options for outcome assessment include simple clinical questions that are meaningful and conform with the practical restraint of research in primary care settings. Examples include “Is your pain tolerable?”, allowing for simple Yes/No answers or adapted to Likert-type ratings where required (Markman et al., 2020). Other outcomes relevant to the studied patient populations can be added, such as mood, sleep, individualised outcomes, or disease-specific measures. Depending on the research objective, outcomes in pain trials are likely long-term and calibrated to be clinically meaningful.

Objectifiable endpoints that would facilitate a low-effort data collection and may mitigate against expectancy and reporting bias are sparse in pain research (Wartolowska et al., 2022). However, their development appears particularly relevant in the context of pragmatic trials. Depending on the research question, intervention type, and target audience, changes in medication or healthcare utilisation could be assessed. Other promising examples include the use of remission-criteria or the (non-)discontinuation of treatment as indicators of therapeutic success. In the latter case, however, mechanisms need to be in place to distinguish treatment failure from the occurrence of adverse events. Relevance to stakeholders should be evaluated in all cases.

Economic considerations have the potential to alter clinical decision-making and health policy. It is therefore surprising how infrequently they are reported in recent self-declared pragmatic trials of pain treatments (Hohenschurz-Schmidt et al., 2022c). Especially when comparing interventions for which cost-savings of one over the other therapy are not apparent, data sampling for a formal cost-effectiveness analysis should be encouraged.

The influence of different patient priorities on outcome choice is again illustrated well by the above examples of pragmatic spinal manipulation trials: Goertz et al. (2013), having enrolled soldiers with acute back pain, studied the effects on pain intensity and physical function at four weeks (essentially recognising that their patients desired immediate relief from back pain and return to full activity). Contrastingly, Eklund et al. (2018) counted the number of days with 'bothersome' pain that impacted on daily activities of people with persistent or recurrent back pain over the period of a year (prioritising living well with pain over complete relief). Other, potentially meaningful outcome measures in this context are patient satisfaction or global impression of change (as used by both trials as secondary outcomes). Further, using time-to-relief or time-to-treatment-discontinuation as outcomes could be representative of patient priorities in private manual therapy practice, where patients may weigh treatment costs against acceptable pain levels or a sense of regained self-efficacy (Unsgaard-Tøndel and Søderstrøm, 2021). These outcomes would require trials with variable treatment lengths.

Assessment intensity, frequency, and follow-up duration

The most 'pragmatic' PRECIS-2 score is given when frequency and extent of follow-up assessments are limited to what would be seen in routine practice (Loudon et al., 2015a). Whilst self-declared pragmatic trials in pain tend to employ more laborious follow-up assessments than what would be expected in routine practice (Hohenschurz-Schmidt et al., 2022c), IMMPACT co-authors argued that excessive data sampling may interfere with successful patient retention, resulting in missing data that is not necessarily reflective of normal clinical practice. Intense follow-up will also increase research burden and trial cost. These undesirable consequences need to be weighed carefully against researchers' potential need and desire for expansive outcome collection.

The duration of follow-up periods is thought to be longer in pragmatic trials than in explanatory trials (Loudon et al., 2015a) (over 1 year on average in pain trials (Hohenschurz-Schmidt et al., 2022c)). Ideally, follow-up periods reflect the time horizon of patient and clinical

decision-making, drawing on stakeholder engagement. The need for data completeness poses challenges over such long time periods. Either, 'pragmatism' may have to be sacrificed to protect internal validity, for example scheduling additional data collection calls or appointments. Ideally, however, electronic data collection methods and low-burden outcome measures are employed to mitigate this risk, and pre-planned strategies for handling missing data are in place (see below).

Trial designers should consider carefully which data are needed to answer the research question and how data collection and follow-up duration may interfere with trial feasibility and internal validity.

Study monitoring

Commonly, pragmatic trials employ a mix of centralised and on-site study monitoring to ensure participant welfare, compliance with regulatory standards, and scientific integrity (Beever and Swaby, 2019). Especially in long-term and multi-centre studies, adequate data monitoring enables researchers to detect and correct problems with the sampling and relay of information.

Risk-based or risk-proportionate approaches to monitoring are increasingly employed in many pragmatic trials. Either, monitoring processes are streamlined according to an initial risk assessment or in-depth evaluations are triggered based on pre-specified indicators of risk to participant safety, data integrity or trial conduct (Beever and Swaby, 2019; Cragg et al., 2021). Such approaches minimise interference of monitoring processes with clinical practice and promote the feasibility of large trials (Meredith et al., 2011; E. Irving et al., 2017), and IMMPACT meeting attendees advocated for such approaches to study monitoring.

In PPS trials, the risks to patient welfare are arguably limited (Carnes et al., 2010; Hebert et al., 2015; Niemeijer et al., 2020), and the main risks may stem from problems with data collection or adherence of providers to trial procedures (due to the individualised nature of many such therapies or due to therapist allegiance). Potentially, some of the considerations developed for the CoPPS Statement (Hohenschurz-Schmidt et al., 2023 and chapter 3) can be utilised here

without interfering with a trial's pragmatic attitude, such as training providers and site staff about the need to adhere to scientific standards and trial processes.

Analysis

Analysis methods are an important area where pragmatic and comparative effectiveness trials can become more efficient. Novel analytical approaches also allow for more differentiated distinctions than statistically significant differences between group means (Luce et al., 2009).

As for all clinical trials, prospective registration is essential. This should include a detailed statistical analysis plan to prevent reporting bias. Objectives should be prioritised and accompanied by statistical formulations (Dworkin et al., 2021).

Analysis population and pragmatic approaches to data analysis

The recommended primary analysis in pragmatic trials is intention-to-treat (ITT), where all available data from randomised patients are analysed (Loudon et al., 2015a). In practice, so-called 'modified' ITT analyses are also common, featuring some post-randomisation exclusion criterion (Montedori et al., 2011; Hohenschurz-Schmidt et al., 2022c). It has been argued that ITT analyses can be impacted by trial-specific adherence patterns not necessarily reflective of real-world adherence (Hernán and Robins, 2017). In effect, an ITT analysis asks "What is the average effect on anyone commencing treatment in this trial?". Clinicians, however, may ask: "What is the most effective treatment for the patient in front of me?". This question is not addressed adequately by indiscriminate analyses at the level of all randomised trial participants. Pain therapies and patients' responses to them are complex, frequently defying simple distinctions into effective and ineffective or safe and unsafe treatments based on a population-average (Moore et al., 2010). Instead, the field should move towards analysing clinically meaningful subgroups of patients, for example by dichotomised responder analyses, or prognostic or data-driven subgrouping. To guide clinical and policy decisions, it may also be necessary to evaluate associations between outcomes, for example studying benefits and risks in the same population, which can then allow for the ranking of the desirability of different outcome combinations (Evans et al., 2015, 2019). While detailed approaches are discussed below, IMMPACT authors concurred with Hernán and Robins (2017) that "*pragmatic trials require an explicit and careful definition of*

the effect of interest […] and the collection of high-quality longitudinal data" rather than the "*uncritical reliance on the [IIT] principle*" (p. 1391).

In pain research there is the potential for a confounding effect from underlying pain-generating conditions, such as pain improving not because of better pain management but because of remission of the underlying disease (including the intermittent nature of persistent musculoskeletal pain), making randomisation essential. Conversely, (un)successful pain management may affect the underlying disease outcome, as in the discontinuation of chemotherapy due to insufficient pain relief or the non-adherence to exercise regimens due to fear avoidance or pain. Again, interpretation of one outcome (e.g., pain relief) requires clinical context in the form of another outcome for the same patient (e.g., disease or behavioural parameters). All these factors are trial-specific and require consideration when determining the outcomes of interest, combinations of outcomes, and analytical approach. The second IMMPACT statement will likely propose specific statistical combinations of risk-benefit outcomes (e.g., Evans et al., 2015; Evans and Follmann, 2016), but these appear less relevant to PPS interventions due to their relatively benevolent safety profile. Instead, more pertinent considerations may be benefit versus added costs, or other undesirable effects such as nocebo and false health beliefs (Hohenschurz-Schmidt et al., 2022d).

Estimands & missing data

Despite recommendations to avoid patient attrition and to improve data sampling, clinical trials will always have to deal with missing data. The IMMPACT statement will likely propose the use of estimands as a way of dealing with missing data which also facilitate personalised approaches to trial analysis, where composite outcomes are estimated for specific patients and subgroups. However, the thesis author expects some debate regarding the resemblance to real-world practice of such data analysis approaches and the impact on 'pragmatism' of necessarily associated additional gathering of data about intercurrent events.

In brief, common methods to account for missing data are last observation carried forward (LOCF) or baseline observation carried forward (BOCF). These methods have been widely criticised for their potential to miss the variance in true effects and create false-positive and false-negative findings; as well as for not reflecting the nature of most painful conditions, where symptoms fluctuate and are susceptible to placebo effects (Little et al., 2010; Mallinckrodt et al., 2017; Dworkin et al., 2021). Alternative imputation methods have been proposed, including control-based (or reference-based) imputation, whereby trial-specific assumptions inform the distribution of post-dropout outcomes and which can be tailored to the endpoints of the trial (Carpenter et al., 2013; Cai et al., 2020). To do so, the trial's objectives and estimands of interest need to be clearly specified. An estimand is "*the population quantity to be estimated in terms of the target population, outcome variable, handling of post-randomization (intercurrent) events, and population-level summary of the outcome variable*" (p. 2310, Cai et al., 2020); in other terms: the statistical model underlying the final analysis of trial results. In the most commonly employed estimand in pragmatic trials, the ITT analysis, missing data are often assumed to be missing at random, which does often not reflect the reality of a trial, let alone clinical practice. In practice, discontinuation can be linked to treatment tolerability, lack of effectiveness, treatment success, cost, patient preferences, interpersonal factors, and other subject characteristics. Researchers require knowledge of all these factors to approximate the true effect assignable to the randomised treatment (i.e., the de-facto estimand not confounded by discontinuation or alterations in treatments, use of out-of-protocol interventions, rescue medication, and considering reasons for dropout). Such factors are termed 'intercurrent events' (Cai et al., 2020). By classifying these intercurrent events and meticulously sampling data on their occurrence, different models (estimands) of the outcome of interest can be built and compared, either as sensitivity analysis or pre-specified and as primary analysis (Little et al., 2010). Knowledge of the nature and frequency of intercurrent events then allows researchers to make informed assumptions about the missing data, using control-based (or reference-based) imputation to replace missing data with data from the control condition in various ways. In effect, such an approach improves the assumptions made regarding unobserved data. This appears particularly relevant for pragmatic research questions as it can take into account some of the confounding

factors of real-world practice in a trial-specific manner. Detailed information on the use of various such imputation methods is available (Carpenter et al., 2013; Cai et al., 2020).

Reporting

The unifying theme of the present set of IMMPACT considerations was the balancing between internal trial validity and the research aim of obtaining clinically relevant information. Researchers will inevitably weigh these aims differently and have to find individual methods for any given research question and trial context. For a comprehensive understanding of a trial, the trial report or its supplementary files need to convey the reasoning behind the choice of methods, clarifying what the specific pragmatic question was and why the researchers believe that their trial, with its specific design, can address it. The CONSORT extension for pragmatic trials contains some items to that effect (Zwarenstein et al., 2008) and author self-assessment and justification of trial methods following the PRECIS-2 domains is desirable (Loudon et al., 2015a). Note that use of the PRECIS-2 *table*, not wheel-diagram, is preferable as it allows for a presentation of design rationales (Available at https://www.precis-2.org/Help/Documentation/Toolkit).

In reporting a trial, the results of which are to be used by clinicians and other decision-makers, the authors need to write the report with these end-users in mind: Trial authors are in the best possible position to report in a way that enables others to judge how applicable the findings are to their particular setting and stakeholder populations. Indeed, especially in primary care and private practice, the published article in a scientific journal may be the dominant source of information on which interested clinicians base their decision-making. The information required to translate the results into these readers' own context is frequently not reported, or superficially at best (Hohenschurz-Schmidt et al., 2022c). Who was treated, how, by whom and in which context are crucial pieces of information for the end-user of pragmatic trials and deserve detailed reporting. Arguably, word limit regulations of scientific journals prevent comprehensive reporting, too, and editors should consider allowing for extensive online supplementary material where authors can provide in-depth descriptions of trial populations, settings, and procedures.

Finally, trial reports should be written with a view on facilitating evidence synthesis. When involving pragmatic trials, systematic reviews may be enhanced by information necessary to judge the applicability of findings; and meta-analyses by being able to incorporate information about within-study heterogeneity to evaluate the appropriateness of data pooling (Mbuagbaw and Aves, 2022) or to explain variations in effect sizes (Sajobi et al., 2018).

For the reporting of pragmatic trials in pain research, the thesis author proposes the specific considerations in Table 28.

Table 28: Recommendation for the reporting of pragmatic trials.

- Fully adhere to the current CONSORT reporting statement (Moher et al., 2010a)
- Fully comply with the CONSORT extension for pragmatic trials (Zwarenstein et al., 2008).
- Provide all information necessary for the retrospective assessment of the trial by means of the PRECIS-2 instrument (Loudon et al., 2015a) or present a self-assessment.
- Outline how the trial methods are thought to enable answering of the pragmatic research question.
- Provide information of the permissiveness and nature or concomitant and rescue medication for each group and their probable influence on trial results (see Table 27).
- Report considerations of the relationship between benefits and risks, using statistical methods if appropriate.
- Discuss cost-effectiveness or justify why this has not been considered.
- Justify the choice of the analysis population (more specifically, the estimand) in relation to the research question.
- Acknowledge the origin, reason, and potential implications of missing data.

Effectiveness Section: Discussion and Limitations

Strengths and contributions of the IMMPACT statements

The IMMPACT statements provide a critical account of general methods in pragmatic trials as well as an overview of key considerations for pragmatic trials of pain treatments. These publications advance the field by not simply repeating currently established concepts. Instead, they may stimulate debate about the relative merit of protecting the real-world resemblance of trials versus enhancing internal validity or using novel methods to inform clinical decisions. Also, although not employing formal consensus or guideline development methods, the IMMPACT statements are novel in bringing together pragmatic trials proponents and aficionados of explanatory trials. Informed by a systematic review of methods (Hohenschurz-Schmidt et al., 2022c) and extensive reading of current literature, the IMMPACT statements address contemporary challenges of pragmatic trials and make them applicable to pain research.

The IMMPACT statements do not resolve the tension between internal and external validity. This was not to be expected because often "*the very features of pragmatic trials that support the generalizability […] of their results to real-world practice may also reduce assay sensitivity and therefore limit the interpretation of results*" (p. 462, Ford and Norrie, 2016). Instead, IMMPACT spotlights this conundrum, urging researchers to make explicit their considerations which underlie design choices, and not to follow current dogma blindly. Publications such as the one by Keefe et al. (2022), demonstrate that the IMMPACT statements are aligned with what can be considered an unease with aspects of the current pragmatic trials literature, specifically the replication of clinical practice at all cost (Godwin et al., 2003; Pawson, 2019). The existence of a certain degree of efficacy evidence for the tested interventions is a fundamental requirement for pragmatic and comparative effectiveness trials – at least according to several IMMPACT meeting attendees and some publications (Kerns et al., 2019). During consensus discussions, pragmatic trial researchers stated that such preceding evidence allowed pragmatic trial designers not to be concerned with bias from expectancy effects or questions of

treatment mechanism, instead valuing nonspecific effects as part of a treatment's real-world effectiveness. In the name of maximising generalisability, this attitude was then often extended to valuing heterogeneity across most features of the 'ideal' pragmatic trial. Other trial researchers at IMMPACT held that it was generally desirable to control and measure heterogeneity and understand the reasons for an intervention's (lack of) effectiveness; especially if such controls could be implemented without impacting a trial's ability to answer its primary research question. The fact that any controls on heterogeneity were contentious (not just those relevant for expectancy effects, but also limiting treatment flexibility or implementing minimum adherence requirements), illustrates that the debate was largely a philosophical one. Mirroring the terminological ambivalence discussed at the beginning of the section and the disagreements over pragmatic trials outlined by Pawson (2019), any method that was not 'like normal clinical practice' was controversial.

Although allegedly pivotal to justify a more pragmatic attitude to trial design, citing convincing efficacy evidence is not common practice in pragmatic trial reports. Lack of such reporting was last demonstrated by a systematic review of manual therapy trials in which the thesis author was involved (Roura et al., 2023, under review). This thesis has mentioned the dearth of efficacy trials in the psychology field (Williams et al., 2021), but the consequences can be illustrated just as well in other fields: Referring to the study of exercise interventions for osteoarthritis, Prof Haxby Abbott of Otago (NZ) acknowledges on twitter that the field may have moved from efficacy to effectiveness research too early (Haxby Abbott [@Haxby_Abbott], 2023). Contrasting with over 130 trials of exercise against TAU (Smedslund et al., 2022), only thirteen efficacy trials existed in 2021; and their synthesised evidence is of low certainty (Dean et al., 2021).

In the end, the IMMPACT manuscripts made two important qualifying statements, aligned with the more cautious schools of thought in the pragmatic trials literature (Godwin et al., 2003). First, all co-authors accepted that "[…] *measurable variables that account for heterogeneity should be identified and controlled or included in statistical modelling where the research question permits it; but heterogeneity should be accepted and incorporated into the trial design where required by the objectives of a trial*" (p. 10, Hohenschurz-Schmidt et al., 2023a). Secondly,

both manuscripts highlighted the need to clearly identify existing evidence of efficacy. The thesis author will propose to urge triallists in the second IMMPACT statement to refer to the level/quality of evidence (e.g., employing the GRADE system). Especially for PPS trials, where pragmatic trials and limited efficacy data are common, a clear recommendation to this effect was warranted.

Finally, the second IMMPACT statement will make a field-specific contribution to pain research (barring further development of the author's draft manuscript with IMMPACT co-authors and external peer review). These recommendations address the gaps and challenges identified by this project's preliminary systematic review, such as best-practice considerations for both internal and external validity in each of the PRECIS-2 domains and beyond (Hohenschurz-Schmidt et al., 2022c).

Pragmatic trials are increasingly perceived as limited to the respective scenario in which they are conducted and to closely related target settings (Pawson, 2019). This is because judging generalisability requires detailed knowledge of target settings and populations (Rothwell, 2010), and generalisability cannot be assumed simply because a trial conforms with 'pragmatism' as per PRECIS-2 or other tools. In turn, designing practically relevant trials requires knowledge of clinical processes and involvement of key stakeholders. Here, the IMMPACT statements provide a comprehensive list of important considerations for trial designers in pain research. This achievement can be attributed to the combined expertise of the IMMPACT group in pain science and clinical practice, as well as the thesis author's in-depth engagement with the current literature and trials. Eventually, however, design choices depend on individual research questions and context. Ideally, the IMMPACT statements will promote more comprehensive preliminary considerations in the design phase of pain trials and more transparency in the reporting of trial rationales. Depending on their impact and successful dissemination, the IMMPACT statements may also make trial funders more cautious regarding methodological rigour of pragmatic trials and the justifications of research questions. With its formal link to the FDA, IMMPACT is well-positioned to affect such change.

Limitations of the IMMPACT statements and reflections on guideline development methods

The absence of formal consensus methods limits the statement's claims to objectivity and makes it liable to authority bias; a limitation aggravated by the relatively small and US-centric group of contributors. Whether consensus would have been achieved in a Delphi process and without the multiple manuscript iterations and negotiations is, however, unlikely. Indeed, it was clear that no consensus existed amongst meeting attendees about best-practice methodology in pragmatic trials. The thesis author's ability to negotiate between 'camps' while not belonging to either side may, in the end, have facilitated this compromise statement which hopefully advances the field. On the other hand, the attempt to find compromise may have 'watered down' recommendations to the lowest common denominator (Scott and Black, 1991; Blazey et al., 2022). Results may also have looked very differently had other 'experts' been invited or if individuals (including peer reviewers) had taken a harder or softer stance in the negotiations. However, the fact the IMMPACT considerations deviate from most current conceptual guidance suggest that some advances were made, reaping the benefit of a conversation between triallists with different preferences and experiences. Finally, in the presentation of main considerations, the author attempted to draw on quantitative evidence where possible. Contrary to this project's review of (sham) control interventions (Hohenschurz-Schmidt et al., 2023b, 2023c), the systematic review for IMMPACT did not perform meta-regression analyses to study links between methods and outcomes (Hohenschurz-Schmidt et al., 2022c). For example, whether the degree of 'pragmatism' is in fact associated with altered effect sizes, remains unclear (Mbuagbaw and Aves, 2022).

As a next step, the proposed approach must be tested in pragmatic trials. Whether the recommendations indeed make such trials more reliable, feasible, and better to interpret while not reducing generalisability remains to be confirmed. Conceptual disagreements with

established authors in the pragmatic trials field are also likely and IMMPACT authors may yet have to defend their newly found compromise.

Reflections on guideline development methods

'Consensus statements' are criticised in many fields. The main criticisms – and these certainly apply to many IMMPACT statements – are: "*[R]elevant stakeholders are absent from the consensus process; authors inadequately report the method of evidence synthesis; consensus group members may be 'coerced' to agree; and the reports often fail to capture the rich discussion that occurs during a panel meeting*" (p. 2, Blazey et al., 2022). Reflecting on the role of the ACTTION partnership and IMMPACT meetings in the pain field, it transpires that their methods have evolved despite ongoing limitations. Early, highly cited IMMPACT publications were exclusively based on untransparent expert consensus processes, always involving the same individuals. For example, the 2005 recommendations for core outcome measures in pain trials (Dworkin et al., 2005), the 2008 statement on clinically important differences (Dworkin et al., 2008), or the 2017 publication on pain biomarkers (Smith et al., 2017) involved only 'unsystematic' literature reviews and no patient involvement (also see: http://www.immpact.org/meetings.html). Only in 2008, IMMPACT published a survey of affected patients to identify important outcome domains (Turk et al., 2008). Systematic reviews became established at IMMPACT from 2014 onwards (e.g., Gewandter et al., 2014, 2015; Dworkin et al., 2016; Ferguson et al., 2022). Despite this advance, to which the author also contributed with a systematic evaluation of pragmatic trials (Hohenschurz-Schmidt et al., 2022c), the actual meeting and manuscript writing processes remain untransparent and largely inaccessible to outsiders. The thesis author attempted to introduce formal voting mechanisms into the IMMPACT XXIV meeting, which was disregarded by the steering group. Pointing out the existence of disagreements in the present IMMPACT statement (Hohenschurz-Schmidt et al., 2023a) and publishing detailed meeting transcripts on the website (http://immpact.org/meetings/Immpact24/participants24.html), similarly represent small but important steps towards more transparency. In the writing of the second IMMPACT manuscript

on pain-specific considerations for pragmatic trials, the author hopes that the directors will agree to an additional focus group with subject experts and patient partners, as well as anonymous voting on controversial formulations. In the future, IMMPACT statements could be made much more valuable by two simple steps: First, implementing best-practice consensus or guidance development methodology (Brouwers et al., 2010; Moher et al., 2010b), and second, broadening access to meetings, e.g., by open invitation or targeted recruitment of patient partners, other stakeholders, and non-US researchers. Choosing guideline topics based on priority setting methods may also be beneficial (Petit-Zeman et al., 2010).

Practical Section: Applying Recommendations to a Specific Example

Chapter 6: Trial Protocol and Control Intervention for a Novel Painful Diabetic Polyneuropathy Treatment

Chapter and section summary

Having made largely conceptual contributions to the field of PPS trial research so far, the final part of the PhD project applied the recommendations to a practical example in pain research. To do so in a clinically relevant manner and to potentially broaden the treatment options of a population of people with pain who currently lack nonpharmacological treatments, a novel intervention was developed. Aimed at people with painful diabetic polyneuropathy (pDPN), 'Augmented osteopathic care for people with painful diabetic neuropathy', or *NeuOst*, combines core osteopathic therapy components, exercise, psychologically-informed approaches, and elements from other interventions relevant for people with diabetes and pain. Intervention and trial protocol development involved a multidisciplinary expert group as well as several patient and clinician partners.

This chapter summarises the intervention development, and presents the protocol for a feasibility RCT and the development of a control intervention. Further, the usability of the IMMPACT and CoPPS Statements was tested on this example by means of a modified heuristic usability evaluation framework.

Funding to conduct the study was obtained from *The Osteopathic Foundation*, the intervention developed, and a trial protocol completed. The subsequent trial conduct is not part of this PhD project. The IMMPACT and CoPPS Statements guided the decision to study the efficacy of the manual therapy components of the new intervention. The CoPPS Statement was followed to develop the control intervention. While some proposed trial features are reflective

of the clinical reality of intervention delivery and the characteristics of the target population, following own IMMPACT recommendations, the protocol contains safeguards to monitor the ensuing heterogeneity. The control intervention replicates all but selected manual therapy aspects of the *NeuOst* intervention and aims at balancing expectation effects between groups and concealing group allocation from participants. Patient partners were involved in the development and pre-trial testing to promote the control intervention's acceptability and blinding success. The trial protocol plans to train trial providers extensively to deliver interventions and comply with trial processes, following CoPPS recommendations. Blinding effectiveness and multiple feasibility outcomes will be measured to decide on design adaptations, intervention refinement, and progression to a full-scale trial. Based on the present example, both the IMMPACT and CoPPS Statements were found able to inform almost all decisions with regards to trial design. While less user-friendly, the IMMPACT statement's core tenet was deemed more broadly applicable, i.e., the recommendation to allow for heterogeneity whenever possible but measuring or limiting heterogeneity where required by the research question. The CoPPS Statement was easier to use but identification of potential treatment mechanisms and their omission in the control intervention were challenging recommendations to apply in the field of manual therapy.

This final part of the project brings together the conceptual considerations about best-practice trial design in PPS interventions, also laying the ground for the conduct of a feasibility RCT in the immediate postdoctoral phase. The *NeuOst* project will add knowledge in two important ways: It is an evidence-driven attempt to broaden osteopathic practice beyond musculoskeletal pain; and it will be the first rigorous evaluation of the recommendation frameworks for trials of PPS interventions proposed during the PhD project. Future, independent evaluations of the guidance frameworks are required and will inform updates.

Introduction

Before presenting the development of a feasibility trial protocol and a control intervention, the thesis first focuses on an often disadvantaged clinical population: people with painful diabetic polyneuropathy (pDPN). In practice, there is a lack of nonpharmacological treatment options for these people and pharmacological analgesia is often unsatisfactory. Due to this need and anecdotal reports of patients having made good experiences with osteopathic care, a manual-therapy-based intervention was developed for people with pDPN. On that basis, a trial protocol was designed. To appraise the trial protocol and intervention rationale, it is important to first understand the clinical population and the evidence base to date.

Painful diabetic polyneuropathy (pDPN)

With an estimated prevalence of 9.3% of the adult population, diabetes mellitus (DM) is one of the leading health concerns of the modern world and projected to rise over the next decades, especially in younger people and in countries moving from lower to higher incomes (International Diabetes Federation, 2019). Estimates for associated diabetic peripheral neuropathy (DPN), which is often undiagnosed and may result in severe diabetic foot complications, range from 16% to 87% (Sobhani et al., 2014; International Diabetes Federation, 2019). In about 26% of patients, diabetes is associated with a *painful* diabetic neuropathy, contributing to disability and reduced quality of life (Davies et al., 2006) and associated with higher risk of anxiety and depression (Kioskli et al., 2019; Kec et al., 2021). Musculoskeletal pain is also up to twice as common in a diabetic population compared to healthy controls (Themistocleous et al., 2016). For example, in a Danish study, 60% of diabetic patients reported low back pain compared to 30% of matched healthy controls, also associated with typical chronic pain risk factors such as higher body mass index and sedentary life style (Molsted et al., 2012). Both diabetes and low back pain are also more prevalent in socioeconomically deprived communities, highlighting the importance of social determinants of health (Karran et al., 2020; Hill-Briggs et al., 2021).

Treatment options for people with pDPN

Nonpharmacological interventions are discussed frequently in the context of diabetes but not necessarily in relation to the management of neuropathy symptoms (Pandey et al., 2011; Raveendran et al., 2018). Treatments include physical exercise, dietary interventions (Fjeldsoe et al., 2011; Davies et al., 2015; Yamamoto et al., 2018), bariatric surgery (Maggard-Gibbons et al., 2013), sleep and stress management (Hamid, 2011; Touma and Pannain, 2011; Wagner et al., 2016), and cognitive behavioural interventions (Safren et al., 2014; Castelnuovo et al., 2017). Education, lifestyle interventions, and dietary control form part of clinical guidelines for diabetes alongside pharmacotherapy (McGuire et al., 2016).

In practice, the treatment and management of people *with pain* and DPN focuses on pharmacological interventions (Tölle et al., 2006; Meisinger et al., 2018), despite the limited effectiveness of many drugs and range of side effects (Dixit and Asiri, 2014; Finnerup et al., 2015; Javed et al., 2015). Contrastingly, clinical guidelines for DPN and expert consensus advocate for nonpharmacological options to be considered alongside pharmacotherapy, notably lifestyle modification, optimal diabetes treatment, and multifactorial cardiovascular risk interventions (McGuire et al., 2016; Ziegler et al., 2021). Additionally, these experts call for patient management beyond pain relief, also addressing quality of sleep, functionality, and general quality of life (Ziegler et al., 2021).

The evidence for nonpharmacological interventions for pDNP

Despite recommendations in guidelines, few clinical trials have investigated nonpharmacological treatments for people with pDPN (with the exceptions of trials for dietary supplements and neurostimulation, which are more abundant but not discussed here due to the scope of the project). Recent systematic reviews on neuropathic pain (Moisset et al., 2020) found high-quality studies only for mindfulness-based stress reduction in people with DPN (notably

Nathan et al., (2017)). Bai et al. (2022) reviewed eight cognitive behavioural or mindfulness intervention trials. Studies were small (total average sample size of 48 participants), mainly not participant blinded, and of mixed methodological quality; Psychological interventions were superior to usual care or educational controls at the short term for the reduction of pain, but no long-term effects were found for any outcomes. The small number of psychological intervention trials is remarkable in light of the high comorbidity of psychological problems and the potentially mutually reinforcing nature of neuropathic pain and psychological risk factors (Kioskli et al., 2019).

Regarding exercise interventions, the available systematic reviews differ in their assessment: Moisset et al. (2020) argue that high-quality research into massage- or exercise-based interventions is sparse. A narrative review of physiotherapy interventions identified eight exercise trials but the methodological quality of the review was low (Jahantigh Akbari et al., 2020). Hernández-Secorún et al. (2021) included 22 studies of exercise for DPN, reporting signs of effectiveness of various forms of exercise on functional outcomes and pain, mainly compared to no treatment. This included 4 trials of strength training, 12 trials of balance programmes, and 7 aerobic-type training programmes; about a third of studies were classed as low risk of bias according to the PEDro scale (Hernández-Secorún et al., 2021).

A systematic review of various types of manual therapy for DPN concluded that manual therapy was effective short-term, albeit based on only six low- or mid-quality studies comparing interventions against treatment as usual (Hernández-Secorún et al., 2021). Other reviews dedicated to nonpharmacological therapies for people with DPN (Çakici et al., 2016; Amato Nesbit et al., 2019) or painful peripheral neuropathies (Liampas et al., 2020) "*underscore[…] the lack of good quality studies […]*" (p. 4102, Liampas et al., 2020), concluding that *"[f]urther research should address long-term outcomes and other non-pharmacologic treatments.*" (p. 15, Amato Nesbit et al., 2019).

Osteopaths are familiar with treating people experiencing pain (Fawkes and Carnes, 2021). A recent estimate placed the proportion of patients with diabetes in UK osteopathic practice at about 4% (Plunkett et al., 2021). Although the proportion of patients seeking help for

or with painful DPN is unclear, authors have speculated about a potential role of osteopathic manual therapists in the treatment and management of people with diabetes and concomitant pain. Treatment rationales, as hypothesised in the literature, include the facilitation of exercise and active lifestyles, the reduction of pain by treating associated neuropathy or concomitant musculoskeletal conditions, and, to a lesser degree, indirect effects on diabetes disease markers via enhanced engagement in activities (Shafer et al., 2020). Existing studies to these effects are, however, sparse and of questionable methodological quality. They include a subgroup analysis of a larger osteopathic trial showing that low back pain in patients with DM improved compared to a sham intervention; as it did in the overall sample irrespectively of diabetes (Licciardone et al., 2013). Further, several small studies and student projects on disease parameters or disease complications were performed, including diabetic ulcers (King, 2016), intraocular pressure (Díaz Cerrato, 2008; Kuhmann, 2008), blood glucose (Bono Mira, 2008; Carpenter, 2016; Vermeersch and Quaghebeur, 2018), insulin levels (Pêgas de Oliveira, 2008), and HbA1c (Kiegerl, 2007), suggesting an interest in the field but not allowing for conclusions regarding treatment efficacy.

The sparsity and low quality of research into nonpharmacological treatments and management approaches for people with pDPN is a notable gap, given how common and debilitating this complication of DM is, and given the inadequacy of existing drug treatments and their associated side-effects (Dixit and Asiri, 2014; Finnerup et al., 2015). Osteopaths, with their expertise in pain management may be well-placed to look after people with pDPN, especially when provided with additional training regarding diabetes complications and in behavioural and other relevant interventions. Combinational or adjunct psychological approaches are increasingly explored for pain management, including psychologically-informed physiotherapy (Keefe et al., 2018). ‘Augmentations’ of osteopathy with cognitive-behavioural techniques have been successfully developed before (Carnes et al., 2017). Whilst the role of manual techniques in a pDPN population is under-researched, limited evidence suggestive of efficacy exists (Hernández-Secorún et al., 2021). It may be hypothesised that manual treatment and other therapeutic components may interact with one another and with patient characteristics to facilitate

behavioural change and promote health (Underwood et al., 2007; Bishop et al., 2015; Bialosky et al., 2017a; Keefe et al., 2018).

However, to work effectively and safely with people with pDPN, osteopaths may require further training regarding specific disease complications and patient management. Osteopaths in the UK are trained according to the regulating body's standards. These include, '*a sound understanding of […] non-neuromusculoskeletal case presentations*' and *'[a] patient who is under the care of another healthcare professional for a condition where osteopathy may not be a suitable approach to treating the underlying issue, but where osteopathic treatment may help to alleviate symptoms.*' (p. 15, General Osteopathic Council, 2022). Theoretical familiarity with diabetes and polyneuropathies can thus be assumed, but the profile of osteopathic practice (Plunkett et al., 2021) makes it unlikely that practitioners have much practical experience with patients with advanced disease complications or considerable disability.

Aims and objectives

To enhance the applicability of osteopathic care to a clinical population of people with neuropathic pain and diabetes, an intervention was developed, drawing on core osteopathic therapy components, psychologically informed approaches, and combining them with elements from other interventions relevant for people with diabetes and pain. This intervention was called 'Augmented osteopathic care for people with painful diabetic neuropathy', or *NeuOst*.

To prepare an efficacy trial of this intervention, a protocol for a feasibility trial and a control intervention had to be developed. This process was also used to test the applicability and usability of the CoPPS and IMMPACT guidance frameworks for clinical trials of PPS interventions (Hohenschurz-Schmidt et al., 2023d, 2023a, and chapters 3 and 5), developed as part of this PhD project.

Intervention development of *NeuOst* and the protocol development for a feasibility RCT formed part of this PhD project. In the immediate postdoctoral phase, *NeuOst* will be taught to fully qualified osteopaths who then deliver the intervention as part of a feasibility RCT.

Methods

Academic collaboration

For this project part, a collaborative network of leading research experts was assembled to refine the mechanistic rationale for manual therapy in pDPN, develop the combination of osteopathy with other relevant interventions, decide on adequate trial designs for a clinical study, and to prepare and conduct a feasibility trial. The thesis author led this project and sought input from collaborators at selected and specific points of the project, identified below. Input was sought according to collaborators' expertise at the ideation and protocol development stage on specific aspects of the intervention development (for example during research visits or virtually), and to deliver components of the provider training for *NeuOst*.

Funding and conflicts of interest

Research visits to academic collaborators were facilitated through a personal travel fellowship by the *Society for Back Pain Research*. A personal £20,000 project grant from *The Osteopathic Foundation* (oF) was awarded as a seed fund to develop the intervention and provider training course and to conduct the feasibility study. Both grants were awarded as part of open competitions. Quarterly project reports are submitted to the oF and, as part of the funding agreement, the research will be presented at funder-related conferences once completed. The funders have no role in the planning, conduct, or reporting of the study but their contributions will be acknowledged. All contributors' potential conflicts of interest will be reported as per International Committee of Medical Journal Editors (ICMJE) guidance (www.icmje.org) in final research reports.

Intervention development

Frameworks for intervention development and project organisation

Development of the *NeuOst* intervention drew on the current framework for complex intervention development and evaluation by the Medical Research Council (MRC) (Figure 11) (Skivington et al., 2021).

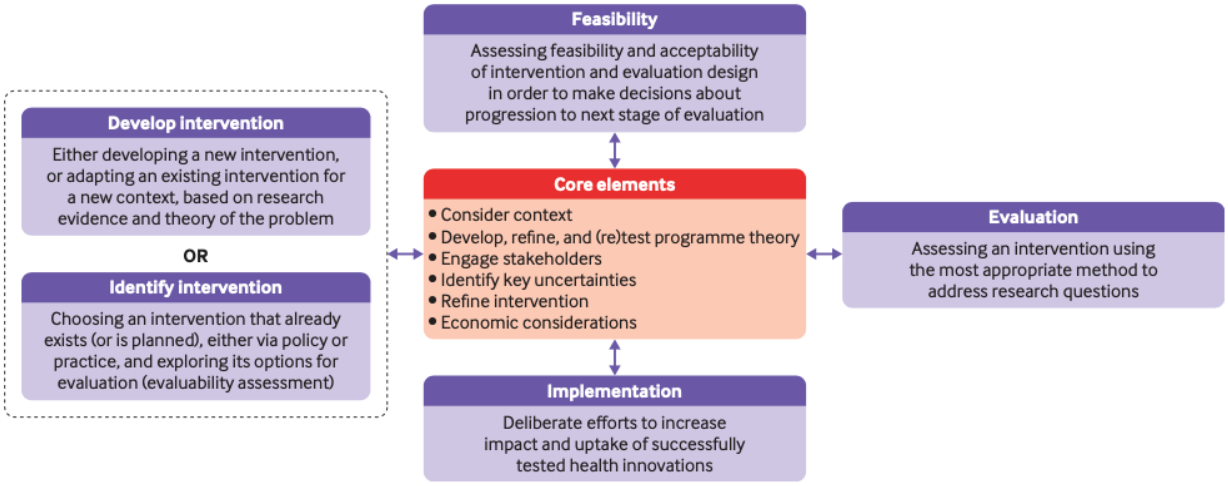


Figure 11: MRC Framework for developing and evaluating complex interventions. *Context*: any feature of the circumstances in which an intervention is conceived, developed, evaluated, and implemented; *Programme theory*: describes how an intervention is expected to lead to its effects and under what conditions—the programme theory should be tested and refined at all stages and used to guide the identification of uncertainties and research questions; *Stakeholders*: those who are targeted by the intervention or policy, involved in its development or delivery, or more broadly those whose personal or professional interests are affected (that is, who have a stake in the topic)—this includes patients and members of the public as well as those linked in a professional capacity; *Uncertainties*: identifying the key uncertainties that exist, given what is already known and what the programme theory, research team, and stakeholders identify as being most important to discover—these judgments inform the framing of research questions, which in turn govern the choice of research perspective; *Refinement*: the process of fine tuning or making changes to the intervention once a preliminary version (prototype) has been developed; *Economic considerations*: determining the comparative resource and outcome consequences of the interventions for those people and organisations affected (Figure from Skivington et al., 2021; used in line with the applicable Creative Commons Attribution CC BY 4.0 license).

NeuOst was expected to adapt and combine existing interventions for a new context. Intervention development thus focused on the MRC framework's core elements (e.g., developing and refining the programme theory, and engaging stakeholders). An overview of the process is provided in Figure 12 (adapted from Craig and Campbell, 2015). Methods for the feasibility evaluation will be described below.

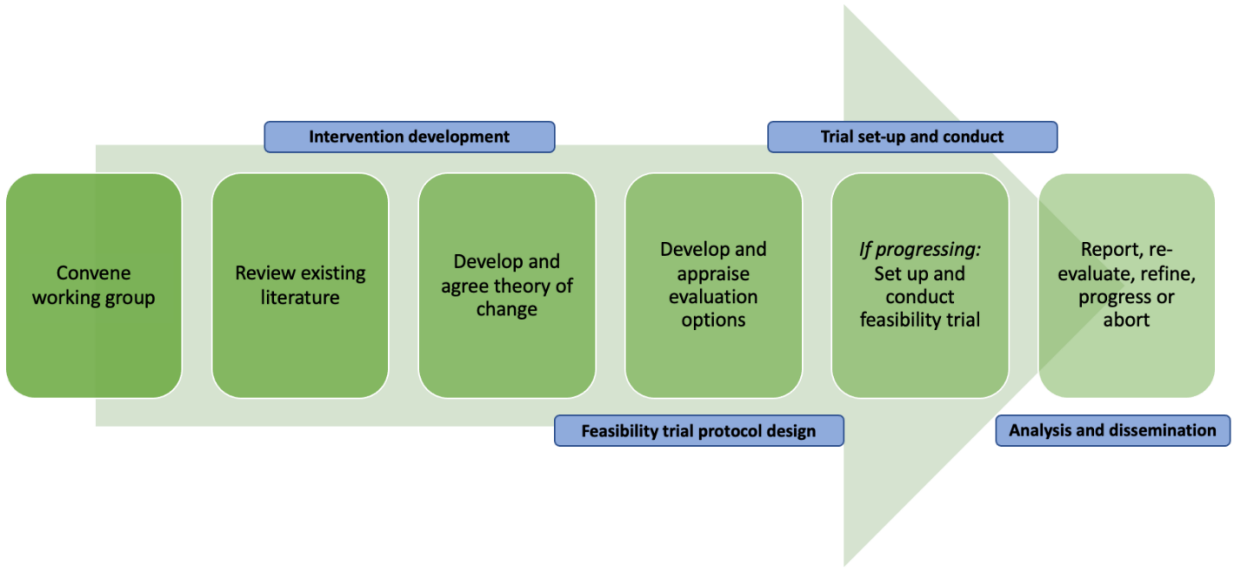


Figure 12: Process flowchart of intervention development and evaluation. The first four tasks form part of this PhD project. Adapted from: Craig P, Campbell M. Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies: a What Works Scotland Working paper. 2015.

Literature review

It was considered to perform an independent literature review of PPS treatment modalities that were potentially suitable for the NeuOst intervention. However, several recent systematic reviews in neuropathic pain and peripheral neuropathy were available (Amato Nesbit et al., 2019; Liampas et al., 2020; Moisset et al., 2020; Hernández-Secorún et al., 2021; Jahantigh Akbari et al., 2020; Bai et al., 2022; Jones et al., 2022). While each of those reviews had limitations of methodology or scope, in their combination they provided a useful and current overview of available literature, and no independent systematic review was performed. However,

osteopathic (https://ostemed-dr.contentdm.oclc.org/) and chiropractic (https://www.chiroindex.org/) databases were searched specifically for the terms 'diabetes', 'neuropathy', and 'neuropathic pain' to also capture articles not published in scientific journals. Pathophysiological mechanisms of pDPN were reviewed as part of a co-authored narrative review with Miss Sasha Smith (Imperial College London) (Smith et al., 2022a) and so were relevant prevention strategies for DPN (Smith et al., 2022b).

Drawing on the above reviews, the available evidence regarding individual treatment modalities was plotted against target mechanisms in people with pDPN. In this table (provided in Appendix 4: NeuOst intervention, additional content'), therapeutic modalities or techniques were annotated as to their potential role within a multimodal intervention and their expected interaction with other components.

Theory of change development

Drawing on the MRC framework (Skivington et al., 2021) and logic models for complex interventions (UK Office for Health Improvement and Disparities, n.d.), the context for the proposed intervention was defined, an intervention rationale developed, and intervention boundaries specified. Further, the required resources, implementation strategies, desired outcomes, and broader anticipated impacts were outlined (See logic model in Appendix 4: NeuOst intervention, additional content). The above 'evidence table' of potential intervention components served as a guide for conversations with the *NeuOst* working group: Collaborators commented on individual components according to their expertise and suggested how components could be operationalised in a treatment protocol.

Stakeholder engagement

Having identified people living with pDPN and osteopathic clinicians as the two main stakeholder groups, these were recruited for stakeholder engagement activities through social

media adverts, leaflets in South London diabetes community clinics, newsletters of relevant organisations (*Diabetes UK* and the *Institute of Osteopathy*), personal networks, and a dedicated project website (https://www.uco.ac.uk/NeuOst, updated since). Stakeholders were recruited for individual interviews or sustained accompaniment of the study, as per their preference.

Initial engagement prior and during intervention development involved semi-structured interviews with patients to understand the experience of people living with pDPN as well as their interest in nonpharmacological care and potential barriers to treatment participation. Interviews with osteopaths and physiotherapists who had experience working with people with painful polyneuropathies focused on understanding their experiences and clinical rationales. Once core intervention elements had been agreed amongst the working group, patient-centred (PPI) activities involved practical treatment sessions with patient partners. Here, the draft intervention protocol was tested regarding its feasibility (for example in terms of time requirements), and informal feedback collected from patients regarding preferences and acceptability, including negative experiences.

Further, it was aimed to hold regular meetings with stakeholders at all stages of the project. We also recruited one person with a lived experience of pDPN to an independent trial steering committee to monitor the feasibility study (see below).

Ethical approval for stakeholder engagement and intervention development was not required (NIHR Oxford Biomedical Research Centre, n.d.), but the UK National Institute for Health and Care Research (NIHR) Standards for Public Involvement were followed (National Institute for Health and Care Research, n.d.). For example, barriers to involvement were reduced by reimbursing stakeholders for their time and by facilitating virtual interviews. A participant information sheet was made available, detailing PPI content, objectives, and governance processes, and informed consent was obtained at the beginning of each patient's involvement.

Whilst interim results of the stakeholder involvement are reported and reflected on below, the final study report will include the presentation of complete results, reflections, and

critical appraisal of stakeholder engagement activities in this study, following the Guidance for Reporting Involvement of Patients and the Public (GRIPP2) (Staniszewska et al., 2017).

Trial protocol and control intervention development

In developing a trial protocol, the thesis author first considered key uncertainties with regards to the newly developed intervention to define a research question. Secondly, potential research perspectives were matched to the primary research question and other key uncertainties. Research perspectives and resulting trial design choices were fundamentally informed by the trial methodological frameworks developed during the earlier stages of this PhD project. Practically, the author drafted a trial protocol and sought feedback from collaborators before submitting it to a research ethics committee. In the autumn of 2021, the author also participated in the short course 'Contemporary Clinical Trials' by the Imperial College Clinical Trials Unit (https://www.imperial.ac.uk/clinical-trials-unit/contemporary-clinical-trials-course/), using this 10-session course to structure and inform the protocol development.

As a core design element of the *NeuOst* trial, a control intervention was developed, following the steps described in the CoPPS Statement (Hohenschurz-Schmidt et al., 2023d). Dedicated stakeholder involvement activities were also performed. The result is presented as descriptive text and a completed CoPPS checklist below.

Usability evaluation of IMMPACT and CoPPS frameworks

This thesis chapter also describes the advantages and challenges of applying the guidance offered by the CoPPS Statement and the IMMPACT statements on pragmatic trials, as applied to the design of the *NeuOst* feasibility trial and its control intervention. To do so, the author modified a usability testing method, specifically Nielsen and Molich's (1990) Heuristic Evaluation of User Interfaces. For this, Nielsen's (1994) ten heuristics were adapted to the evaluation of a research guideline (resulting in the eight criteria shown in Box 1). Adaptation also considered

AGREE II, a framework for the appraisal of clinical practice guidelines that, however, focuses on guideline development rigour rather than user experience (Brouwers et al., 2010). Next, the thesis author wrote a reflective account of the experience of developing the *NeuOst* trial protocol, the present thesis chapter, and the control intervention. This report was then scored against the items in Box 1, using a 5-point Likert-type scale (as per Nielsen, 1994) and brief descriptions of the author's experiences were presented in tabular form. This usability evaluation was performed for IMMPACT and CoPPS Statements separately.

Box 1: Heuristic evaluation of research guidelines. Adapted from Nielsen (1994). The rating of each heuristic followed a 5-point Likert-type scale, ranging from 0 = The guideline does not fulfil this criterion at all, to 5 = fulfils it completely.

1. **Visibility of research steps:** The guideline should provide clear information on the stage of the research process to which recommendations apply. The user should be able to easily understand what stage of the research they are at, what needs to be done next, and what the outcomes are.
2. **Match between guideline and the real world:** The guideline should use terminology and concepts that are familiar and relevant to the user, and the research process should align with real-world research practices (including feasibility in light of practical constraints of clinical trials). The guideline should not introduce unnecessary jargon or complex terminology that might confuse the user.
3. **Consistency and standards:** The guideline should use consistent terminology, formatting, and labeling throughout the document. There should be no conflicting messages. This consistency makes it easier for the user to understand the guideline and use it effectively.
4. **Recognition rather than recall:** The guideline should present information in a way that is easy for users to remember and apply.
5. **User control and freedom:** The guideline should provide the user with options to review and modify their research plan as needed. Users should be able to make choices about how to approach their research.
6. **Error prevention:** The guideline should be designed to anticipate and address common errors users might make, such as missing key steps in the research process. The guideline should provide clear instructions and guidance to help the user avoid mistakes.
7. **Help and documentation:** The guideline should provide clear and accessible documentation that explains the research process and how to use the guideline. This documentation should be easily accessible and provide answers to common questions and problems that users might encounter.
8. **Other features.** As the heuristics framework was newly adapted and is not usually used to evaluate research guidelines, a 'other' field was included to allow for evaluation of unanticipated challenges.

The thesis author acknowledges that independent evaluations of both frameworks will be required and seeks to identify their biases throughout the presentation of the evaluation's results.

Results

Working group

The members of the *NeuOst* working group and their respective areas of expertise are listed in Table 29. In 2022, the thesis author visited Prof Esther Pogatzki-Zahn in Münster, Germany, mainly discussing the local group's experiences with developing an intervention for postsurgical pain management, interdisciplinary care, and trial conduct. The author also visited A/Prof Annina Schmid in Oxford, UK, focussing on the potential role of physical interventions in the management of neuropathic pain and intervention protocol development. A further research visit to Dr Esther Williamson (Oxford, UK) will be conducted when the trial is running to discuss monitoring and interpretation of findings.

Table 29: *NeuOst* working group.

Name	**Role and contributions**	**Qualifications and expertise**
Mr David Hohenschurz-Schmidt	Project lead; Senior Responsible Officer (SRO) during feasibility trial Contributions: • Project planning and management, Budgeting and financial administration together with UCO finance department. • Completion of Grant Management Reports • Research protocol planning, research conduct, and write-up (intervention development and feasibility trial) • Working group coordination	GOsC-registered osteopath; Clinician-academic; PPS trial design
Mr Steven Vogel, DO	Project Sponsor (representing the sponsor institution) Contributions: • Project governance and oversight • Contractual arrangements with funder and sponsor institution	Deputy Vice Chancellor (Research) of University College of Osteopathy, London; Editor in Chief of the International Journal of Osteopathic Medicine; Clinician-academic; Manual therapies

	• Interpretation of results, contribution to write up	
Prof Andrew SC Rice	Collaborator Contributions: • Project strategy development • Academic guidance and project oversight • Inputs regarding the medical management of neuropathic pain into intervention development	Professor of Pain Research at Imperial College London and Hon. Consultant in Pain Medicine at Chelsea and Westminster Hospital; Neuropathic pain
Prof Lene Vase	Collaborator Contributions: • Academic guidance with a focus on sham control intervention design	Professor of Psychology at Aarhus University, Denmark; Placebo effects and mechanisms; Sham control interventions
Univ.-Prof. Dr. med. Esther Pogatzki-Zahn	Collaborator Contributions: • Academic, notably regarding trial planning • Inputs regarding the medical management of neuropathic pain • Trial monitoring • Trial analysis and interpretation	Prof in Anaesthesiology and Pain Medicine; Anaesthetist and pain physician; Group leader of the Experimental and clinical pain research group at University Hospital Münster, Germany.
A/Prof Annina Schmid	Collaborator Contributions: • Academic, notably regarding intervention development (manual therapy, exercise components for neuropathic pain, and peripheral neuropathies)	Associate Professor, Wellcome Clinical Career Development Fellow, NIHR Oxford BRC Senior Research Fellow, Specialist Musculoskeletal Physiotherapist MMACP
Dr Esther Williamson	Collaborator Contributions: • Trial design, monitoring, and analysis	Deputy Director, Centre for Rehabilitation Research in Oxford; Clinician-academic; Physiotherapy; Clinical trials
Dr Kirsty Winkley	Collaborator Contributions: • Intervention development (components regarding diabetes management) • Provider training course development and delivery	Diabetes specialist nurse and health psychologist; Reader in Diabetes and Primary Care at the Florence Nightingale Faculty of Nursing & Midwifery, King's College London; Clinician-academic; Diabetes and DPN complications management

	• Assistance with patient involvement of people with painful DPN and • Patient recruitment into feasibility trial	
Dr Matthew Evans	Collaborator Contributions: • Academic, notably regarding treatment mechanisms and epidemiology of people with pDPN • Inputs regarding the medical management of neuropathic pain into intervention development • Provider training course development and delivery	NIHR Clinical Lecturer in Neurology; Clinical Research Fellow at Pain Research Group, Faculty of Medicine, Dept. of Surgery & Cancer, Imperial College London. Clinician-academic; Neurology
Dr Whitney Scott	Collaborator Contributions: • Academic guidance • Intervention development regarding psychological pain management • Design of provider training programme and possibly delivery	Lecturer in Clinical Health Psychology & Clinical Psychologist and Research Lead at INPUT Pain Unit; King's College London, Department of Psychology, Mental Health & Psychological Sciences, Institute of Psychiatry, Psychology & Neuroscience; Clinician-academic; Psychology
Ms Sasha Smith	Collaborator Contributions: • Trial design and assistance with practical conduct	PhD student and Clinical Trial Manager at Imperial College London, based at the Section of Vascular Surgery, Department of Surgery and Cancer.
Dr Phil Bright	Collaborator Contributions: • Intervention development regarding integration of physical exercises	Head of Research at European School of Osteopathy; Clinician-academic; Manual therapies; Exercise interventions
Mrs Soran David	Collaborator Contributions: • Intervention development regarding integration of psychologically informed approaches	Unit Leader University College of Osteopathy; Osteopathy; Psychologically informed practice

	• Delivery of provider training module	
Dr Jan Vollert	Collaborator, Project Statistician Contributions: • Assisting with statistical analysis plan • Conduct of statistical analysis • Assistance with trial monitoring report compilation for independent steering committee review	Biostatistician and Postdoctoral Researcher at Imperial College London; Neuropathic pain research
Dr David Evans	Trial Steering Committee, Chair Contributions: • As per NeuOst Trial Steering Committee and Date Monitoring and Ethics Committee terms of reference	Research Fellow at the Centre of Precision Rehabilitation for Spinal Pain (CPR Spine) at the University of Birmingham; Honorary research fellowships at the Research Centre at the University College of Osteopathy in London and the Warwick Clinical Trials Unit; Clinician-academic; Osteopathy
Dr Siobhan Stynes	Trial Steering Committee, Member Contributions: • As per NeuOst Trial Steering Committee and Date Monitoring and Ethics Committee terms of reference	HEE/NIHR ICA Clinical Lecturer in Physiotherapy at Keele University; Clinician-academic; Back pain
Mrs Elizabeth Pigott	Trial Steering Committee, Patient representative Contributions: • As per NeuOst Trial Steering Committee and Date Monitoring and Ethics Committee terms of reference • Input during intervention development and trial planning	Retired nurse; Person living with pDPN

Stakeholder engagement

Here, interim results from stakeholder engagement are presented that focus on the development of the *NeuOst* intervention. Stakeholder engagement regarding the development of the trial protocol and control intervention is presented in the respective section describing the trial protocol.

Five people with pDPN were interviewed online, one of them twice, and one at regular intervals amounting to four interviews. This last person also attended an in-person meeting where the first *NeuOst* treatment session was tested practically. In total, eight interviews focused on patients' experiences of living with painful neuropathies and with various interventions; two later-stage interviews were about a draft trial protocol and the control intervention. All but one patient had Type 2 diabetes (T2DM) and they experienced varying degrees of neuropathy (in some cases undiagnosed and even questionable). One person had chemotherapy-induced neuropathy. There were no amputations, the mean age was 74 years, and there were three women and two men.

Further, four osteopaths and a physiotherapist were interviewed. Two of the osteopaths had worked as clinic tutors in a community clinic for people with HIV for many years. The third osteopath had been treating their partner for chemotherapy-induced neuropathy (interviewed as a couple) and had thus become interested in the treatment of people with neuropathies. The fourth osteopath was interviewed about the protocol of the planned trial. She had experience with providing control interventions due to her involvement in a recent efficacy trial of an osteopathic intervention. The physiotherapist led the physiotherapy component of a regional outpatient pain management programme in Germany and was interviewed during a research visit to Münster. Providers had been practising for an average of 15 years and there were two women and three men.

At the date of writing, eleven hours of stakeholder engagement were conducted for the *NeuOst* project, 8.5 with patients and 2.5 with providers.

Patient interviews highlighted the following topics relating to the development of the *NeuOst* intervention. (Results of stakeholder interviews regarding the trial protocol and control intervention are presented separately.)

1. Diagnostic uncertainty
    a. Although all patients believed they had painful neuropathy, not all patients had been formally diagnosed. In one uncertain case, a diagnostic screening questionnaire (the DN4) was used and the patient did not qualify for a diagnosis of neuropathic pain according to the most commonly used threshold value (Spallone et al., 2012).
2. Knowledge variability
    a. Whilst some were aware of neuropathy and its basic mechanism, others had hardly ever heard of the role of their nerves and were glad to be explained some of the principles. Similarly, the link between neuropathy and Charcot foot or ulcerations was not apparent to all interviewees. Some patients also complained that nobody had taken the time to educate them about the basics of diabetes complication management, including neuropathy and foot care.
3. Impact of symptoms
    a. Most patients reported considerably reduced quality of life due to neuropathy symptoms. Pain at night and interference with sleep seemed the most troubling. A single patient suffered from recurrent ulcerations but only had temporary neuropathic pain. This patient was most preoccupied with foot care and the avoidance of wounds. Finally, balance problems and the fear of falling figured strongly in those with also otherwise more debilitating experiences.
4. Diabetes management
    a. DPN cannot be considered independently from living with diabetes. This was most apparent from a somewhat inconclusive call after which the interviewee

contacted the study team a few hours later, apologising that they had become "rather low" (in blood glucose) and could not remember much of the conversation. They hoped to "have made any sense". This incident demonstrated that only repeated work with people with diabetes and / or careful training would sensitise providers to such clinical scenarios and prepare them to provide first-aid hypoglycaemic treatment.

5. Desperation
    a. Most patients were desperate for non-drug treatment options for their pain. Those who had tried available medication disliked it for its side effects and did not take it regularly. Others were trying various alternative creams and ointments.
6. Different attitudes towards exercise
    a. Patients felt overall positively towards physical activity but engaged in it to very variable degrees and felt differentially confident about performing activities. Again, the fear of falling and of wounds were strong for a person with respective balance problems and previous experiences. This person nonetheless engaged in daily dance exercises at home, broken up into several brief sessions a day, and advocated for safeguards, such as opportunities to hold on to something whilst performing the exercises and to adapt the degree of challenge for balancing exercises. Others considered themselves active but mainly mentioned performing household chores or regular walks as exercise, not engagement in structured programmes or leisure sports.
7. Positive attitude to manual therapy
    a. Although only one patient (the partner of an osteopath) had had manual therapy for their symptoms, all were very positive about the possibility and were willing to try it in theory. Nobody voiced concerns regarding safety or practicability. There was a sense of hope at the possibility (likely following from the feeling of desperation described above). The person with chemotherapy-induced neuropathy described improved sleep after foot mobilisations and massage in the

evenings, but only for the following night so that their partner was treating their feet most evenings.

The person treated by me as part of the intervention development provided written feedback afterwards, citing no adverse effects and a reduced feeling of 'tiredness' in her legs. She also wrote: "*The foot exercises* [meaning the manual therapy techniques] *are good. They help release a 'floppy' feeling / not having to unknowingly clench muscles in your feet and legs while you are waiting for the inevitable pain to come. So, there is already a chance that if neuropathy pain comes on it may not be as bad as you normally get as your limbs were more relaxed. It would be a great feeling for persons with pain as it was a relief you did not know you needed. Something different to take your mind off the one constant annoyance you endure. […] I enjoyed the session completely.*"

Provider interviews led to the following insights:

1. Treatment approach and rationale
    a. Practitioners did not have an elaborate reasoning behind the application of their techniques other than to "move the feet" and "provide sensory input". They mainly advocated for a prolonged mobilisation of the feet to "stimulate the Golgi tendon organs" as well as generic articulatory and soft tissue treatment of the lower extremities and potentially lumbar spine.
2. Dampened expectations
    a. Providers did not perceive the manual component as life-altering for patients, but believed it provided short-term relief. They related patient stories where even a few hours relief were deemed worthwhile by patients, but highlighted that this meant regular (e.g., weekly) treatment that was unlikely to be possible in a private practice setting.
3. Effects beyond the musculoskeletal domain
    a. Practitioners highlighted areas of treatment that were often as important to them as the physical side. For one person (the physiotherapist), this was behaviour

change and changing unhelpful beliefs around the pain experience. For the osteopaths that had worked in community clinics, this meant social support and the ability to manage disease complications (mainly of HIV), for example by integrating with medical care or by signposting accordingly.

Description of the *NeuOst* intervention

The above development process resulted in the *NeuOst* intervention, a manual therapy-based multimodal intervention also incorporating elements of diabetes management, education, exercise, and ACT-informed activities. Specifically, the *NeuOst* intervention includes the components listed in Table 30. *NeuOst* was developed for board-registered osteopaths as providers who are to be trained in the specific intervention components during a 2-day training course, where they will also be equipped with relevant population-specific knowledge and diagnostic skills.

Table 30: Components of the NeuOst intervention.

Component	Techniques	Integration into *NeuOst*
Manual therapy	- Soft tissue techniques (massage and stretching) - Articulatory / joint mobilisation components - 'Subtle' Manual Therapy components (e.g., Balanced Ligamentous Tension and Craniosacral techniques) - Thrust components ('Spinal manipulation') - Soft tissue component 2 (Neurodynamic techniques) - Active resisted muscle activation techniques	Provided as a core component in every treatment session.
Physical activity / Exercise	- Exercise-related advice and education - Sensorimotor / balance training	Provided as a home-based graded exercise plan, with new exercises practised in every in-person session and the ability to

	- Static and / or dynamic muscle stretching - Anaerobic / strengthening exercises - Aerobic / cardiovascular exercise	personalise the plan and any exercises.
Communication, Education, and ACT-informed components	- Pain self-management advice - Pain education - Sleep management education - Techniques from an ACT workbook on mindfulness, awareness, openness, and shifting of attentional focus.	Provided in in-person sessions, some components optional and as per patient needs. Patients encouraged to practice ACT-based skills at home.
Diabetes management components	- Diabetes complication screening and advice on disease management - Education on diabetes and its complications - Foot care education	Provided especially in first in-person session and revisited in each subsequent treatment session. Personalised to patient needs and may require engagement of other managing healthcare providers.
Integration with standard medical care	- Providers aware of standard medical care and local care pathways	Acknowledgement of patients' medication list and communication with managing / prescribing clinician as needed.

A detailed description of the intervention theory, a review of evidence for individual components, and an intervention protocol for the purpose of the *NeuOst* trial are provided in Appendix 4: NeuOst intervention, additional content.

Trial protocol

Key uncertainties, research questions, and major design decisions

Before describing the *NeuOst* trial protocol in detail, key uncertainties with regards to the *NeuOst* intervention are outlined, research questions specified, and the resulting design decisions described. It is also described how this decision-making process was informed by the IMMPACT and CoPPS frameworks as well as specific stakeholder engagement.

Pragmatic versus explanatory attitude to trial design

Little is known about the potential role of osteopaths in the care for people with pDPN, or even with any type of peripheral polyneuropathy. The need for improved person-centred care for people with pDPN and the good medical education of UK osteopaths, however, justified exploring their potential contribution in this clinical population.

There was a good scientific rationale for the incorporation of several nonpharmacological interventions into the *NeuOst* intervention, notably exercise (Hernández-Secorún et al., 2021), education (He et al., 2017), and diabetic foot care and complication screening (NICE, 2015). Although established for the management of people with pDPN, it needs to be noted that supporting evidence comes almost exclusively from effectiveness studies. Studies either showed benefits from these interventions on patient-reported outcome measures relating to pDPN, or on underlying disease parameters and complication risk (for the complete 'evidence matrix', see Appendix 4: NeuOst intervention, additional content). For CBT and mindfulness interventions, there was a weak recommendation based on moderate evidence from open-label trials (Moisset et al., 2020). For most manual therapy techniques, there was a lack of studies. Common techniques used by osteopaths had not been studied, including joint articulation, passive muscle stretching, or spinal manipulation. Limited evidence existed only for the efficacy of foot massage compared to educational controls (Chatchawan et al., 2015) and there was questionable evidence of effectiveness for foot massage compared to various other interventions (synthesised by Fu et al., 2020). Therefore, a full-scale trial of *NeuOst* would have to include a specifically designed control intervention to test the efficacy of the manual therapy components, controlling for expectancy and learning effects.

This design choice is aligned with the IMMPACT framework for pragmatic trials in pain research, which postulates that '*before testing the effectiveness of new treatments in routine practice, existing efficacy and safety evidence for a treatment or for core components of a multi-modal treatment ought to be considered to determine whether more efficacy research is needed*' (p. 4, Hohenschurz-Schmidt et al., 2023a). The second scenario under which comparative effectiveness trials are commonly conducted, even in the absence of high-quality efficacy data,

is when interventions are low-risk and already widely used in clinical practice (Hohenschurz-Schmidt et al., 2023a). This may be the case for diabetes education, foot care, exercise, and possibly psychological interventions, but is not the case for manual therapy which is not commonly used by people with pDPN or other polyneuropathies.

The final consideration regarding a generally more pragmatic or explanatory attitude to trial design concerns the target audience; in the words of the IMMPACT statement: '*The overall trial design is thus guided by how trial results will be used*.' (p. 5, Hohenschurz-Schmidt et al., 2023). If the objective was to directly inform clinical decisions or policy, for example funding decisions in the NHS, a more pragmatic attitude and thus a comparative effectiveness study between *NeuOst* and usual care could be considered appropriate. Instead, however, the objective here is to determine whether osteopaths may have any role to play in the care of people with pDPN and if further research regarding real-world effects may be warranted. This scientific rather than policy question is best answered with an efficacy-focused RCT studying the effects of osteopaths' unique contribution, the manual treatment.

***Decision 1**: To conduct a trial focused on evaluating the efficacy of a manual therapy component within a multimodal package of care for people with pDPN.*

The IMMPACT statement is clear that trials should be designed to achieve their objectives, not along a false pragmatic/explanatory dichotomy (Hohenschurz-Schmidt et al., 2023a). As such, the objectives of this trial were to obtain reliable data on the potential clinical benefit of the manual therapy component of the *NeuOst* intervention under best-case circumstances (i.e., with ideal adherence etc.) and controlling for placebo effects. In other words, the main objective was to test the superiority of *NeuOst* over a 'sham' control intervention.

Apart from discussions below, a completed PRECIS-2 table documents design choices on the pragmatic-explanatory continuum and describes the underlying rationale, available in Appendix 5: Supplementary information for *NeuOst* feasibility trial').

Blinding and controlling for placebo effects

In line with an efficacy-focused research question regarding the manual therapy component of *NeuOst*, controlling for placebo effects was an objective of the specifically designed control intervention. To achieve this, a control intervention was designed that aims to blind participants to their group allocation and replicates all components of the investigated treatment, apart from the components whose effect the trial aims to study. The design of this control intervention is described in detail below. This design decision was based on the 'similarity principle' of control design in efficacy trials, described in section 2.1 of the CoPPS Statement (Hohenschurz-Schmidt et al., 2023d).

***Decision 2**: To blind participants to their group allocation and ensure replication of all intervention components, apart from the manual therapy component of* NeuOst.

The first IMMPACT statement also comments extensively on 'blinding and accounting for participant expectancies': First, this project was not interested in real-world effectiveness of *NeuOst* at this stage (for which '*it is accepted that the non-specific (e.g., contextual or placebo) effects form part of treatments' real-world effectiveness"*, p. 6), so teasing out placebo and specific effects was appropriate. Second, the project does not study which of two well-established interventions is to be used; another scenario where blinding may not be indicated. "*Overall"*, IMMPACT states "*we recommend blinding participants to group allocation where efficacy data are inconsistent and where compatible with trial objectives*." (p. 6, Hohenschurz-Schmidt et al., 2023a). This was certainly the case here.

Three-armed designs

Both the CoPPS Statement and the IMMPACT statement highlight the potential merit of including a third study arm, such as waitlist or no-treatment groups. In CoPPS, the highlighted benefits are to understand the extent of the placebo effect and the potential real-world

effectiveness of the intervention (CoPPS section 2.2.) (Hohenschurz-Schmidt et al., 2023d). IMMPACT also discusses a trial by Beard et al. (2018) as example of three-armed RCTs with both a 'sham' control and a no-treatment arm. Applying both frameworks to *NeuOst* showed that a third arm was not indicated from either perspective: In CoPPS, it would be a way of answering an additional research question about placebo effects and thus not obligatory. From the IMMPACT point of view, the 3-arm design would have been justified if *NeuOst* was an already well-established treatment, which it was not, or to ensure assay sensitivity in a comparative effectiveness study, which, again, is not planned for the reasons outlined above. On this point, it needs to be noted that the control intervention as planned can be considered 'another active treatment' due to multiple active intervention components. Nonetheless, the research question was not one of comparative effectiveness or whether osteopaths can deliver education and exercises effectively, but rather one of superiority of the manual therapy component over placebo for reasons outlined above.

Decision 3: To conduct a two-armed trial.

Full-scale versus feasibility trial

Although at this early stage of the *NeuOst* intervention development and evaluation life cycle (Skivington et al., 2021), efficacy may have been the most pertinent academic question, other important uncertainties existed. These related to 1) the feasibility and acceptability of training osteopaths in this multimodal intervention, 2) the acceptability of *NeuOst* to patients, and 3) the feasibility of a participant-blinded RCT with a specifically designed control intervention. The latter had to be studied as part of the development of a high-quality control intervention, as per CoPPS statement. (Section 3.2) (Hohenschurz-Schmidt et al., 2023d).

Further, needs for intervention refinement may also have to be identified, acknowledging the complexity of the intervention and unknown interactions between components (Skivington et al., 2021).

These feasibility-related questions had to be addressed prior to embarking on a full-scale efficacy study.

Decision 2: *To investigate the feasibility of an efficacy RCT of NeuOst.*

To evaluate these questions, a feasibility trial with pre-defined and stakeholder-informed feasibility criteria will be conducted as per the protocol below. In line with the MRC framework's comments on research perspectives, other research questions such as those relating to implementation, effectiveness, cost-effectiveness, translation to other settings, and scalability, are to be addressed if acceptability is ascertained (Skivington et al., 2021) and once signals of efficacy are obtained from a full-scale RCT (Freedland et al., 2019; Hohenschurz-Schmidt et al., 2023a).

With these fundamental considerations in mind, the next paragraphs describe other design elements of the *NeuOst* RCT. While the design was developed as for a full-scale RCT, methods of a full-scale trial may change based on experiences from the feasibility trial.

A detailed protocol for the feasibility RCT can be found on the Open Science Framework once agreed by the working group (https://osf.io/sbqce/). That protocol follows the CONSORT reporting guideline extension for nonpharmacological treatments (Boutron et al., 2017) and pilot and feasibility trials (Eldridge et al., 2016a), and the SPIRIT reporting guidelines for trial protocols (Chan et al., 2013), and will be submitted for institutional ethics review after a final consultation with the *NeuOst* work group and monitoring and steering committee. Once approved, the protocol will be pre-registered on clinicaltrials.gov. Please consult the protocol for detail on how the below design choices are operationalised.

Trial objectives

The primary objective is to test the feasibility of a single-site parallel-arm participant-blinded controlled randomised trial of a 6-week course of *NeuOst* in addition to usual care, compared to a designated control intervention for adults with pDPN.

Secondary objectives are to obtain indications of potential efficacy and safety of *NeuOst*, including estimates of between-groups treatment effects and their variance on key pain/disability outcomes, exploring ITT, per-protocol, and a mixed-effects model adjusted for baseline differences.

Further, qualitative data on recruitment routes, reasons for exclusion during screening, reasons for attrition, and participant and provider experiences during the feasibility trial will be collected where possible.

Trial design

The *NeuOst* feasibility RCT will be a two-armed, parallel-group, participant-blinded, single-site, individually 1:1 randomised clinical trial, designed to study the feasibility of a full-scale efficacy trial testing the superiority of the *NeuOst* intervention over a control intervention on pain-related outcome measures in people with pDPN. The feasibility trial will consist of six treatment sessions and have a 4-week post-intervention follow-up.

Participant flow and timelines are illustrated in Figure 13.

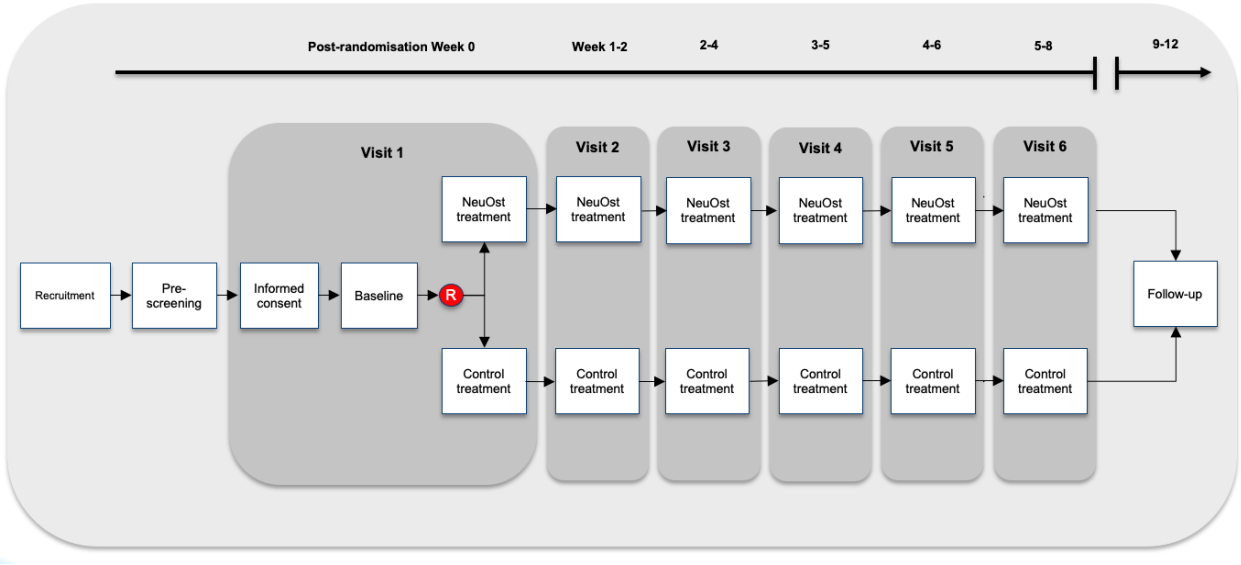


Figure 13: Flowchart of the NeuOst feasibility trial. A Multimodal Manual Therapy-Based Intervention for People with Painful Diabetic Neuropathy: Feasibility of a Randomised Controlled Efficacy Trial.

Study setting

The trial will be conducted in a single centre, the University College of Osteopathy teaching clinic in central London (https://www.clinic.uco.ac.uk).

Participants (eligibility criteria)

Eligible are adults with pDPN, irrespective of comorbidities and medication use (Table 31).

Table 31: Participant eligibility criteria for the *NeuOst* feasibility RCT.

Inclusion criteria
- Medical diagnosis of diabetes mellitus (type 1 or type 2) as per patient self-report - Distal symmetrical peripheral neuropathy (defined as a score of ≥4 on the Michigan Neuropathy Screening Instrument (questionnaire part) (Herman et al., 2012)) - Neuropathic pain as assessed by DN4 questionnaire (defined as a score of ≥3 of 7 self-reported items) (Spallone et al., 2012) - Participants capable of giving informed consent themselves - Age: 18 and older
Exclusion criteria
- Contraindications to manual therapy and conservative pain management (includes medical emergencies and suspected severe pathology), and active foot ulcerations - Past or scheduled amputations - Advanced renal failure - Recent physical trauma and suspected fracture - Severe recurrent depressive episodes - Carpal Tunnel Syndrome diagnosis or symptoms as the only source of neuropathic pain - Unable or unwilling to provide consent for study participation - Unable to attend in-person appointments at treatment site (for health reasons) - Unable to attend in-person appointments at treatment site (for organisational reasons) - People who know other participants signed up to the study
Note: Explicitly *not* excluded are people with:
- Blindness - Other pain in addition to neuropathic pain, such as musculoskeletal pain or mixed pain presentations - Any medication use, including opioids

Recruitment

Participants will be recruited through online and in-print advertising, using Google Ads, social media, and the local newspaper 'Southwark News'. Further, the author is in consultation with the charities Diabetes UK and Know Diabetes to employ their local networks for recruitment. Further, local podiatry and diabetes services will be used for recruitment (depending on whether NHS ethics will be obtained).

Randomisation

Blocked randomisation with permuted block sizes will be used, generated with the web-based electronic data capture system Castor (https://www.castoredc.com/). Participants will be allocated to treatment and control groups in a 1:1 ratio. Participants will not be stratified, and no minimisation scheme employed. Group allocation will be performed by the consent-obtaining researcher after obtaining consent, communicating a participant's group allocation to the treatment provider using a coded label to ensure blinding of participants and clinic administration staff to participants' group allocation.

Providers

Trial providers will be fully qualified and GOsC-registered osteopaths, specifically trained to deliver *NeuOst*. They will have a minimum of two years' clinical experience since graduation but no more than minimal clinical experience with people with peripheral polyneuropathies prior to the trial.

Interventions

NeuOst intervention

Patients will receive six sessions of the *NeuOst* intervention, delivered by qualified osteopaths in a 1:1 in-person format. Treatment will consist of manual therapy, exercise, ACT-informed elements, and diabetes management and education, following a pre-defined protocol. Individual intervention components are outlined in Table 30, and a protocol describing intervention content per treatment session is available in Appendix 5: Supplementary information for *NeuOst* feasibility trial.

In brief, each treatment session will last 45-60 minutes and include manual therapy components, exercise, ACT-informed activities, targeted communication, education, and clinical processes relevant for diabetes management and complication screening. Physical exercises will be introduced and practiced in these in-person sessions, and patients encouraged to follow a gradually intensifying home-based programme of aerobic, anaerobic, and sensorimotor exercises, allowing for pacing and exercise individualisation. Medical case history taking, patient-centred communication, and clinical testing are also employed.

Control intervention protocol

As discussed initially, the aim of the control intervention will be to facilitate the study of the efficacy of *NeuOst*'s manual therapy component. The control intervention was thus designed to omit the manual therapy component and its hypothesised mechanisms of action but otherwise resemble the *NeuOst* intervention in all features listed in the CoPPS Statement (such as session duration, structure, addressed body areas etc.; see Table 3 in Hohenschurz-Schmidt et al., 2023d). The control intervention will be delivered by the same set of providers as the *NeuOst* intervention, who will be specifically schooled. Through this high-similarity design, the control

intervention is hoped to blind participants to their group allocation and thus balance expectancy and learning effects between groups.

The manual therapy components in *NeuOst* are soft tissue techniques, joint mobilisation, passive muscle stretching, active-resisted muscle activation, and neurodynamic techniques. These will be replaced with intermittent soft touch from providers, passive less-than-mid-range joint movement, manoeuvres that mimic neurodynamic techniques, and active joint movement by patients. As per CoPPS Statement, this avoids hypothesised treatment mechanisms mediated by the activation of muscle spindles, tendon organs, mobilisation of nerves in their connective tissue sheet, local effects on muscle tissue, and the activation of the descending pain inhibitory systems through nociceptive stimulation. Effects of the control intervention through other mechanisms cannot be ruled out, but these mechanisms are not typically hypothesised as important in manual therapy. (Such mechanisms include proprioceptive input from mid-range stimulation of tendon and muscle mechanoreceptors, C-tactile and other skin mechanoreceptor-mediated effects, centrally mediated effects on the autonomic nervous system and psychological 'stress' arising from soft touch and other pleasantly perceived stimuli, and effects mediated by increased skin temperature). Manual therapy elements in the control intervention will be described as 'subtle hands-on techniques' and a specific mechanistic rationale for these will be provided to patients. Communication around this control element will be practised with providers. Further, a treatment protocol will suggest wording to providers and clearly define which communication, education, and manual techniques are in and out of scope in each study arm.

Apart from the CoPPS Statement and the above scientific considerations, specific stakeholder engagement informed the design of the control intervention. Two people living with painful polyneuropathy were interviewed who also had experienced the *NeuOst* intervention and/or other osteopathic treatment. Further, two osteopaths were questioned about their thoughts on the proposed *NeuOst* protocol, one of whom had participated in a recent efficacy trial of paediatric osteopathy where they had provided a soft-touch control treatment.

The interviewed patient partners understood the scientific need for a trial and approved of a participant-blinded design, providing there was informed consent. They felt that omitting only the manual therapy components was preferable to omitting all or multiple 'active' intervention components. This was due to two considerations:

1. Patient partners felt that some 'value' should be provided to control group trial participants. Interviewees had made positive experiences with, e.g., foot exercises which led them to recommend including those.
2. They had concerns that omitting more components might compromise the credibility of the control intervention.

All PPI participants, including osteopaths, approved of the proposed description of the control intervention as 'a treatment that is the same as the tested *NeuOst* intervention but omits one component that the trial aims to study'. This description is aligned with the recommended wording in the CoPPS Statement (Hohenschurz-Schmidt et al., 2023d). Interviewees also highlighted the need to carefully communicate the purpose of the trial to participants, highlighting that even if participants themselves felt no improvement in their symptoms, taking part and persisting with the study might contribute to the development of a helpful intervention in the future. This notion reflects findings from the patient interviews conducted for the CoPPS Statement, where patients also perceived a 'higher purpose' of taking part in research (Section 3.5, Hohenschurz-Schmidt et al., 2023).

An interviewed osteopath related their unease during a previous trial when they noticed that patients had not understood that they might be randomised to a control treatment, highlighting how important a well-conducted informed consent process may be to providers. They were also concerned that they had to act unnaturally in the control arm, for example if they were asked not to speak to patients. In the trial they had taken part in previously, they had also felt uncomfortable withholding a treatment from patients in which they felt emotionally and professionally invested (or, in the words of the CoPPS Statement, had high 'allegiance' to (Section 3.5, Hohenschurz-Schmidt et al., 2023)). They appeared more at ease after discussing the notion of equipoise and that there was currently very limited evidence to suspect manual therapy components were effective for people with pDPN. They were also reassured by being able to

provide many other 'active' treatment elements, mirroring the thought process of interviewed patient partners.

Finally, all interviewees were aware of the pivotal role of trial providers in maintaining participant blinding and delivering interventions with high fidelity (further discussed below).

The CoPPS checklist for control interventions in PPS trials has been completed for the *NeuOst* feasibility study (Table 32), documenting and making explicit the reasoning behind control-related design choices.

Table 32: Completed CoPPS Checklist for the control intervention used in the *NeuOst* feasibility trial. Checklist of recommendations to observe during the development and implementation of a control intervention in efficacy and mechanistic trials of physical, psychological, and self-management interventions (CoPPS Checklist). According to the CoPPS Statement, these items represent a core of best-practice recommendations that apply in any controlled efficacy and mechanistic trial of PPS interventions. Using columns 3-4, the completed checklist documents our decision-making during the planning phase of the *NeuOst* feasibility trial.

Essential recommendation item	Reference to CoPPS section	Implemented (Y/N)	Description of how implemented or justification for not implementing
Design and development stage			
Ethical arguments for and against performing an efficacy or mechanistic trial with a control intervention were considered, including from a trial participant's perspective.	2.3.	Y	*As no high-quality evidence of efficacy exists for the NeuOst intervention, conducting an efficacy-focused trial is considered ethical. Testing trial feasibility before embarking on a full-scale trial is ethically warranted to avoid wasting resources and participants' time on a full-scale trial* (Blatch-Jones et al., 2018). *Participants will not forgo another intervention that is known to be efficacious, and there are no limits on medication use as part of participant eligibility.* *The risk to patients from assignment to the control intervention is expected to be minimal because a) all involved and modelled treatments have a low risk profile, b) patients will be under the supervision of qualified healthcare providers, and c) safety monitoring processes will be in place.* *The research burden to participants is considered acceptable even when assigned to the control intervention, as the control intervention will contain active therapeutic elements and may thus provide benefits through those mechanisms and through placebo effects. This view was explicitly confirmed by stakeholder interviews. The burden of research-related processes such as outcome assessments is minimised through careful selection of employed outcome measures and through using a telephone interview rather than in-person visits for follow-up. Participants will also be reimbursed for their travel expenses. Patient partners considered the proposed blinding procedures acceptable during preparatory stakeholder interviews.* *Careful communication during the informed consent process will outline the ethical considerations for and against taking part in this study, clarify oversight processes and that the control intervention was developed and tested with patient partners, and highlight participants' ability to withdraw themselves and their data from the study at any point and without providing a reason. These processes and documents will be reviewed by an institutional ethics committee. A deceptive design as defined by CoPPS (section 2.3; E&E document) is not employed.*

The objectives of the control intervention in the context of the research question at hand were defined.	2.1. & 3.1. This is to include considerations of the blinding of any involved parties.	Y	*The aim of the control intervention is to facilitate the study of the efficacy of NeuOst's manual therapy component. Through compliance with the CoPPS 'similarity principle' and thus resemblance to the tested intervention, the control intervention is hoped to achieve the objectives of blinding participants to their group allocation and thus to balance expectancy and learning effects between groups.* *Intervention providers will not be blinded as their collaboration is required to deliver the control intervention with high fidelity. Deceiving qualified practitioners as to the intention of the control intervention procedures was also considered unrealistic and ethically difficult. Research staff obtaining outcome data and clinical administrative staff will be blinded to participants' group allocation. This is a secondary objective of the control intervention and should be facilitated by participants not being able to discern tested and control interventions.*
A literature review of comparable control interventions and their available blinding data was performed.	3.1.	Y	*In a systematic review of control interventions in PPS trials for people with pain* (Hohenschurz-Schmidt et al., 2022b)*, 62 manual therapy trials were identified. Forty-two (68%) of those employed a manual (i.e., touch-based) control intervention, 16 (26%) employed detuned devices, and 5 (8%) were categorised as 'other'. To develop the NeuOst control intervention, we examined the touch-based control interventions more carefully, acknowledging the CoPPS principle of similarity and the objectives of our control intervention which ruled out non-manual control interventions. Of those 42 touch-based control interventions, 22 (52%) simulated a manual therapy technique or manoeuvre, 19 (45%) employed soft touch only, and techniques or simulated manoeuvres were applied to non-affected body parts in two cases (0.5%). Blinding effectiveness data were available from 7 of these trials; percentages of control intervention participants believing they had received the 'real' treatment ranged from 45%* (Chen et al., 2008) *to 90%* (Panagopoulos et al., 2015) *(Mean 62%, SD 15.9%). All trial authors considered blinding to have been successful according to their own criteria, but Bang's blinding index for an independent assessment could only be calculated in one study* (Bialosky et al., 2009)*. Overall, blinding data were either not available or it was not possible to discern approaches that were more or less effective because of little variability in the reported. Reporting bias possibly plays a role, with ineffective blinding likely underreported.* *Assuming that in trials reporting any blinding-related data, control intervention design was more rigorous and more effective, these trials' control intervention design was assessed first. In this subgroup, the following features relevant for the NeuOst trial were identified:* - *Neural tension or neurodynamic techniques were exclusively controlled for by replicating the technique's limb movement through a smaller range of motion to avoid tension on the nerve and its surrounding soft tissues* (Bialosky et al., 2009; Fernandez-Carnero et al., 2019; González et al., 2021)*.* - *Control intervention manoeuvres not targeted at nerves also involved avoidance of full ranges of joint motion, occasionally simply placing a limb into a certain position and holding it there.*

			- *In about half the cases, soft-touch control interventions involved some movement but always limited pressure.* *Irrespective of whether blinding data were reported, the following features were also noted. Soft-touch control interventions without any movement of the practitioner's hands were more common in this group than in the allegedly effectively blinded trials:* - *To control for 'therapeutic intent', practitioners were occasionally instructed to perform a distracting mental task, such as counting backwards* (Kutner et al., 2008; Aghabati et al., 2010). - *In spinal manipulation and articulatory techniques, control interventions regularly involved a slight displacement of the therapists' hands away from the intended joint, thus guiding forces into bone and not joint or soft tissue structures* (Surenkok et al., 2009; Lewis et al., 2010; Hudson et al., 2018). - *A similar approach was used by Rocha and Sanchez (2012) for trigger point treatment, pressing the muscle more gently and not directly on identified trigger points.* - *In a few cases, control interventions were made more 'active' by instructing patients to move their spine or limbs or to breathe in a particular way* (Eguaras et al., 2019). - *Papa et al. (2012) used an extended postural examination and body palpation as a control intervention.* - *As one of few author groups, Fraser et al. (2020) reported how the control intervention was communicated to patients ("participants […] were told that they were to receive a gentle soft-tissue technique similar to massage"). In doing so, they likely set expectations that were then met by the control intervention application.* *Based on these previous experiences, the NeuOst control intervention was decided to include the following features, later evaluated against their potential to affect treatment mechanisms (see next cell):* - *A 'sham' neurodynamic sequence as per above.* - *'Sham' mobilisation involving less than 50% or a joint's range of motion and, where possible, directed not at the joint line but adjacent bone.* - *Control soft tissue techniques that involve little force but some superficial skin movement.* - *Techniques will be described as deliberately 'gentle', 'precise', or 'subtle' by the providers.* *In addition, a mechanistic rationale for individual techniques will be created for patients.*

The mechanism(s) of interest of the test intervention were defined.	3.1.	Y	*The component of interest was the manual therapy. According to Bialosky et al. (2017), who reviewed possible mechanisms of manual therapy, manually applied mechanical force is an inherent element of any manual therapy intervention, although they find no evidence to support any specific mechanical parameters as the mechanism leading to the clinical effects of complex manual therapy. They further postulate that manual therapy acts on multiple mechanisms, from local tissues to neurophysiological, immunoendocrine, and cognitive changes mediated by the peripheral and central nervous system. They also acknowledge 'nonspecific responses', which are not the mechanism of interest in the proposed efficacy trial. Bialosky et al. (2017) do not mention low-force touch as a possible mechanism of manual therapy, but this has certainly been postulated by other authors, especially in the paediatric, cransiosacral and therapeutic touch manual therapies (e.g., McGlone et al., 2017). Instead, Bialosky et al. (2017) only refer to studies with considerable mechanical force, such as massage or spinal manipulation techniques. Studying the underlying neurophysiology, a paper by Proske 2012 helps to identify receptors that would be activated through, for example, soft tissue or joint mobilisation techniques. The control intervention will minimise mechanical stimulation or relevant receptors to omit the 'chain of neurophysiological effects' produced by transient mechanical stimulation of peripheral tissues* (Bialosky et al., 2017a).
The components of the test intervention thought to act on the above mechanism(s) were specified.	3.1.	Y	*Relevant components thought to act on the above-identified receptors and mechanisms are:* - *End-range muscle stretches* - *Fast movement* - *End-of-range joint positions* - *High pressure on tissues* - *Active muscle contraction against resistance* - *Gentle or caressing superficial skin movement* - *Pain or uncomfortable sensations for patients* - *End-range stretches and positions that could tense a neurovascular bundle, specifically the sciatic nerve's*
The control intervention replicates as many components of the test treatment as possible, other than the component(s) being tested.	2.1. Also see Table 3 for a list of influential components	Y	*It replicates all components other than those listed in the previous cell. These are replaced with elements not thought to act on hypothesised treatment mechanisms. Namely:* - *Intermittent soft touch from providers* - *Passive less-than-mid-range joint mobilisation* - *Manoeuvres that mimic neurodynamic techniques but avoid nerve tension* - *Active joint movement by patients without resistance*

It was ensured that the control intervention is inert for the studied mechanism(s) and does not include the component(s) of interest.	2.1.	Y	*A topical review of potential mechanisms of action of manual therapy was conducted. Based on this, the following manual therapy features were decided to be avoided, to minimise stimulation of certain afferent fibres:* - *Avoid end-range muscle stretches* - *Avoid fast movement* - *Avoid end-of-range joint positions* - *Avoid large pressure on tissues* - *Minimize gentle or caressing superficial skin movements* *Ensuring that these mechanical forces were avoided was part of the development and will be implemented in the provider training; Here, providers will be coached to apply limited force and only to certain body parts. However, there is considerable uncertainty about the relative contribution of various mechanisms to any efficacy of manual therapy, mechanisms have not been explored in people with pDPN, and it will not be possible to measure involvement of relevant receptors and neurophysiological mechanisms as part of the* NeuOst *trial. All these factors will have to be acknowledged as limitations in a trial report.* *Further, based on the neurophysiological review, it cannot be excluded that the control intervention influences patients' experiences through the following mechanisms:* - *Effects from mid-range stimulation of tendon and muscle mechanoreceptors (provision of proprioceptive input)* - *C-tactile and other skin mechanoreceptor-mediated effects (mainly on pain inhibition and the limbic system)* - *Centrally mediated effects on the ANS and stress levels arising from soft touch and pleasantly perceived stimuli* - *Effects mediated by increased skin temperature*
The control intervention was tested in a feasibility or validation step, ensuring that certain quality criteria were met (see CoPPS statement).	Note the quality criteria and additional suggestions in sections 3.2. & 3.3.	Ongoing	*This is part of the* NeuOst *feasibility trial, were intervention credibility, blinding effectiveness, participant expectancy, and control intervention acceptability to patients and providers will be tested. As part of the early control intervention development, stakeholder engagement was performed, and a practical PPI session is planned where a patient partner can comment on the control intervention.*

The risk of group contamination was considered and mitigated if required.	3.2.	Y	*Patients will be scheduled to visit the clinic at different times according to their group allocation. At the screening interview, patients will also be excluded if they know other people participating in the trial. Providers will be trained to maintain blinding throughout. Administrative staff will be blinded to participants' group allocation and coached to recognise the need for maintaining blinding, for example discouraging them from discussing treatment experiences with patients.*
Providers were specifically trained to deliver the control intervention (if applicable).	See similarity items in sections 2.1. (Table 3) and considerations for training in section 3.3.	In preparation	*This will be part of the immediate trial preparation, and training content is in planning. Considerations of equipoise and potential barriers to the delivery of the control intervention were explored in PPI interviews and will be specifically addressed.* *Further, pre-trial provider training will include the following topics as suggested by the CoPPS Statement:* - *Discussing the concept of equipoise and the need to adhere to trial procedures* - *Educating providers about the complex nature of the tested intervention, the replication of most of its components in the control arm, and the potential benefit to participants through the placebo effect* - *Discussing the similarity between conditions and the many remaining components of the control* - *Reminding providers of the ethical justifications of control interventions in clinical research and the purpose of RCTs* *Providers' expectation of benefit will be assessed at regular intervals throughout the study. Patient-provider interactions during the trial will be partially scripted, and fidelity assessed intermittently.*
Staff (not just treatment providers) were educated to recognise the importance of maintaining effective blinding (if applicable).	3.3.	Planned	*This will be done as part of a 1-hour staff training session prior to study commencement. The Head of Clinic and Chief of Reception have already been engaged in two preparatory meetings and welcomed such staff training. Research team members will visit the study site frequently to evaluate and address any problems that may arise.*
Outcome assessors were blinded.	3.4.	Planned	*The trial protocol includes a plan for blinded outcome assessment through a) blinded patients entering their outcome data directly into a data capture system or on paper questionnaires, and b) outcome collection during a follow-up telephone interview by a member of the research team blinded to participants' group allocation.*
The roles of treatment providers and outcome assessors were separate (applicable if providers cannot be blinded).	3.4.	Y	*Treatment providers will not be involved in the collection of outcome data.*

Statistical analyses were blinded.	3.4.	Y	*The trial protocol includes a plan for blinded statistical analysis.*
Trial conduct stage			
Providers' fidelity to intervention protocols and scripts was monitored (if applicable).	4.1.	Planned	*Providers will complete a fidelity checklist after each* treatment session (see Appendix 5: Supplementary information for NeuOst feasibility trial). Intervention fidelity will also be assessed through random audio-recording and observation of a subset of treatment *sessions.*
Participants' adherence to and compliance with intervention protocols was monitored (if applicable).	4.1.	N/A	*This is particularly relevant for the home-based exercise programme, not the control intervention which is largely provider-delivered.*
Provider expectations of benefit from the control versus the test treatment were evaluated (if applicable).	4.2.	Planned	*This will be evaluated through questionnaires after each treatment session (together with protocol fidelity).*
Participants' expectations of treatment benefit were assessed at baseline and after starting treatment sessions.	4.2.	Planned	*Treatment expectancies will be assessed with the Stanford Expectations of Treatment Scale* (Younger et al., 2012) *at baseline and after the first treatment session, and both interventions' expectancy and credibility will be assessed after session 1 with the Devilly and Borkovec Credibility/Expectancy Questionnaire* (Devilly and Borkovec, 2000).
Participant blinding was assessed (if applicable).	4.2	Planned	*At the 4-week follow-up assessment, blinding effectiveness will be examined by means of a blinding effectiveness assessment questionnaire, asking participants to guess whether they received the* NeuOst *or the control intervention (answer options:* NeuOst*, control, and Don't know).*
Reasons for participants' withdrawal from the study were documented.	4.3	Planned	*Participants who withdraw from the study will be contacted via email, text, or phone to obtain reasons for withdrawal. Using multiple-choice or one-word answer options is hoped to increase the chance of obtaining this information.*

Intervention flexibility, adherence, and fidelity

The *NeuOst* intervention protocol contains obligatory and optional components per treatment session (Appendix 5: Supplementary information for *NeuOst* feasibility trial). Similarly, exercise components can be modified, allowing providers to adapt the treatment to patient preferences, abilities, and practical requirements. At the same time, having a list of obligatory core components will ensure a level of intervention standardisation. Protocol adherence will be promoted through pre-trial training and within-trial supervision. Alongside fidelity monitoring through provider self-report and intermittent fidelity checking through study staff, this ensures a necessary level of treatment homogeneity.

Adherence to the home-exercise component in the treatment arm will be measured through weekly patient self-report, obtained through a brief questionnaire prior to each treatment visit (administered by blinded staff and information explicitly not conveyed to treatment provider to avoid performance bias).

Provider training

Providers will be trained in the *NeuOst* intervention during a 2-day course. The course was developed by the *NeuOst* working group and its members deliver the course content. Pre-course reading and recorded lectures will involve the pathophysiology of diabetes and DPN, musculoskeletal changes in diabetes, and a lecture on psychosocial factors in pDPN as well as basic ACT concepts. The in-person training modules include a practical session about neurological bedside assessment for neuropathies led by a neurologist, a Q&A on managing people with diabetes led by a diabetes nurse, a Q&A with a patient partner, practical training in basic psychologically-informed approaches (session led by either a clinical psychologist or an osteopath with training in ACT), training in patient-centred communication, and practical sessions on the intervention's manual and exercise components and how these can be combined in treatment sessions.

To deliver the control intervention, providers will be specifically trained in a similar manner, following CoPPS guidance (Section 3.3, Hohenschurz-Schmidt et al., 2023). Trial-specific provider training will include practical training sessions as well as discussions of equipoise, allegiance, blinding, protocol adherence needs, research standards, and ethical or practical concerns. During the trial, weekly supervision sessions will be facilitated by the study lead or clinical specialists as needed. Provider expectations about treatment effectiveness will be evaluated as part of the self-reported fidelity questionnaire at each session.

Concomitant treatments

Concomitant treatments will be permitted but participants are encouraged to not alter their analgesic medication schedule during the trial. Changes in medication and other concomitant treatments will be monitored as per below.

Blinding

The following stakeholders will be blinded to participants' group allocation:

- Trial participants
- Clinical support and administration staff
- Outcome assessors (collecting data at end-of-treatment and follow-up)
- Trial statistician

Blinding will be attempted through a specifically designed control intervention (sometimes called a 'sham' intervention), that replicates the NeuOst treatment in all components other than the ones whose effect a definite trial would be designed to study. Communication of this blinded trial design to trial participants will emphasise the need for such scientific comparisons and that both treatments are similar, apart from some elements. In doing so, the

terms ‘placebo’, ‘sham’, ‘dummy’ or similar will be avoided as per current CoPPS guidance (Hohenschurz-Schmidt et al., 2023d).

Provider blinding will not be possible as providers will have to actively deliver both the test and the control intervention. However, extensive pre-trial training and supervision during the trial will ensure that providers understand the purpose of the control intervention and are supported to deliver it with high fidelity. In doing so, the CoPPS Statement will be followed in detail, as described above.

Outcomes

The primary study outcome measures will be feasibility outcomes specified in Table 33. Progression rules to a full-scale trial are also specified there. Methods of collecting and analysing the feasibility outcomes are specified in the detailed trial protocol (Available at: https://osf.io/sbqce/).

Table 33: Primary feasibility trial outcomes and progression rules. Bang's blinding index = BI; Note that the index is scaled to an interval of -1 to 1, 1 being complete lack of blinding, 0 being consistent with perfect blinding and -1 indicating opposite guessing. A secondary progression rule regarding blinding effectiveness will be defined, also incorporating participant guesses from the test intervention group in the form of a ratio of BI between both groups.

	Green (Go)	Amber (Amend)	Red (Stop)
Recruitment	20 participants recruited within 2 months	15-20 participants recruited within 4 months	<15 participants recruited within 4 months
Consent	≥40% of potentially eligible participants	20-39% of potentially eligible participants	<20% of potentially eligible participants
Blinding	BI of -1.0 to 0.1 in sham group or >80% of participants in control group unsure about allocation or believing "real intervention" allocation.	BI of 0.1 to 0.5 or 60 to 80%	BI of 0.5 to 1.0 or <60%
Intervention fidelity	>80% of fidelity tests judged as protocol-compliant intervention	50% to 80% of fidelity tests judged as protocol-compliant intervention	<50% of fidelity tests judged as protocol-compliant intervention
Data Completeness	<10% missing data at 4-week follow-up	10-30% missing data	>30% missing data
Participant Retention	<15% attrition at 4-week follow-up	15-40% attrition at 4-week follow-up	>40% attrition at 4-week follow-up
Adverse events	Occurrence of minor trial-related adverse events (e.g., post-treatment transient muscle soreness or symptom aggravation)	Occurrence of a serious trial-related adverse event or excessive minor side effects (>80% of patients)	Occurrence of multiple serious trial-related adverse event

Due to the feasibility nature of the trial, clinical outcomes will be secondary, also for evaluation of suitability as a primary outcome measure in a full-scale trial. These are patient-reported pain severity (measured by the Brief Pain Inventory, short form, also using the pain interference subscale); pain 'bothersomeness' and tolerability (measured as the number of days with 'bothersome' and 'intolerable' pain in past week (as per Patrick et al., 1995; Grøvle et al., 2010; Markman et al., 2020); disability (Oswestry Disability Index, ODI), neuropathy-specific quality of life (measured with the Neuropathy- and Foot Ulcer-Specific Quality of Life Instrument, NeuroQoL (Vileikyte et al., 2003)); generic health-related quality of life (measured by the SF-36 (Ware and Gandek, 1998; Ware Jr., 1999)); sleep disturbance (on the PROMIS Sleep Disturbance Short Form (8a)); physical activity levels (via patient self-report using the Low-level Physical Activity Recall, LOPAR). Exploratory outcome measures are satisfaction with social roles and activities (PROMIS Satisfaction with Social Roles and Activities Short Form (8a)); global impression of change (as per Dworkin et al., 2005); and analgesic medication and concomitant treatment use (via an adapted provider-completed NIH NCCIH Prior and Concomitant Medications and Interventions Form); and depression and anxiety symptoms (Hospital Anxiety and Depression Scale; HADS). Selection of these outcomes was informed by the available literature and patient interviews. Surrogate measures will be avoided, but if self-reported activity levels are found to be potentially affected, activity tracking by means of wearables can be considered for a full-scale trial.

Qualitative information on the *NeuOst* intervention and trial-related aspects will be obtained from a subset of trial participants via interview after follow-up and unblinding. Semi-structured interviews will be conducted with the aims of:

1. Gathering additional information for the interpretation of the most important trial outcomes (incl. contextualisation of feasibility outcomes)
2. Understanding the perceived benefits/harms from *NeuOst* in more detail, also with a view on deciding most appropriate clinical outcomes for a full-scale trial
3. Adjusting or confirming intervention theory and content

Quantitative outcomes will be evaluated as between-group differences in mean changes from baseline to follow-up and the end of the treatment period.

Data collection methods

Baseline data will be collected at the first clinical visit through a researcher prior to randomisation (see study flowchart in Figure 13), including demographic data and clinical questionnaires. Prior to treatment sessions 2-6, participants will be asked by blinded clinic administration staff to complete blinding- and expectancy-related questionnaires and the Brief Pain Inventory. After treatment session 6, outcome data will be collected by a blinded member of the study team, using paper forms or capturing data directly in the electronic data capture system during an interview with the participant. At a single follow-up assessment at 4 weeks after the end of the treatment phase, outcome data will be collected during a telephone interview by a blinded member of the study team.

To evaluate the acceptability of the test and control interventions, trial participants will be invited to participate in optional semi-structured telephone interviews regarding their experience on the trial.

Sample size and power

The study will aim to recruit twelve patients per group. The sample size for the feasibility RCT was chosen based on practical considerations and funding constraints. Since the primary outcomes are feasibility outcomes, a formal sample size calculation was not required. Regardless, taking the Brief Pain Inventory as an example, and assuming a clinically meaningful difference in BPI score as the commonly reported reduction of worst pain ratings of 2 points (effect size of 0.67) (Mathias et al., 2011), a standard deviation of 3 points, and an alpha level of 0.05, the feasibility trial has approximately 35% power to detect this between-group difference. The study is thus, as expected, underpowered. In other words, the study is unlikely to be able to detect a

clinically meaningful difference in pain reduction between the *NeuOst* intervention group and the control group, even if such a difference exists. Therefore, no conclusions about efficacy should be drawn. However, the outcome data can still be used to inform sample size calculations for a larger, full-scale trial, which would have more statistical power to detect the targeted effect size (Eldridge et al., 2016b).

Statistical methods

The feasibility outcomes will be analysed descriptively, and the proportions of participants meeting the progression rules will be calculated.

The analysis of clinical outcomes will be the comparison of the mean changes from baseline to follow-up between the *NeuOst* and control group using an analysis of covariance (ANCOVA) model, adjusting for baseline values. Other analysis models will be explored in preparation of a full-scale trial, such as comparing per-protocol and ITT analysis, and using different imputation methods for missing data. Cluster effects per intervention provider will be explored, and so will be the potential influence of provider and participant expectancy ratings and participants' characteristics (such as disability levels, age, gender).

Study monitoring

Trial steering and data and ethics monitoring will be performed by a single committee (the TSC/DEMC), constituted of two independent researchers and a member of the public. The TSC/DEMC will also invite observers, including representatives of the sponsor and research network to meetings. The terms of reference for the *NeuOst* TSC/DMEC were specified in a separate document and were formally agreed upon between involved parties, including arrangements for harm reporting and monitoring. Regular meetings are scheduled. A statistician of the sponsor institution will prepare data reports for data monitoring.

Ethics

Ethics approval will be sought from the sponsor institution's Research Ethics Committee.

Usability evaluation of IMMPACT and CoPPS statements

The development of the trial protocol and the control intervention were guided by the IMMPACT (Hohenschurz-Schmidt et al., 2023a) and CoPPS Statements (Hohenschurz-Schmidt et al., 2023d). The experience of employing these guidelines to the *NeuOst* project were evaluated from a user perspective, with results presented in Table 34 for the IMMPACT statements, and Table 35 for the CoPPS Statement. In brief, the IMMPACT statements were more difficult to work with. Especially the first IMMPACT statement offers little in terms of step-by-step guidance or checklists for researchers; multiple 'considerations' require readers to weigh off design options in the context of their own study. Doing so may require substantial knowledge of trial methodology and is not helped by conceptual ambiguity inherent in the field of pragmatic trials. At the same time, however, IMMPACT is clear in its core tenets and draws on existing frameworks to prevent low-quality trial design. The CoPPS Statement was found to be a strong and clear guide through the research process. It also provides comprehensive additional documentation and support to inform and structure decision-making.

Further, it was found that both documents complemented one another, mainly for the formulation of the study question and for basic design choices such as deciding the comparator arm. The IMMPACT statement's recommendation to consider each methodological decision for its impact on internal *and* external validity was applicable also to a generally efficacy-oriented study, where such reasoning in the design phase will facilitate the reporting of potentially generalisable aspects of the study as well as the delineation of the limits of generalisability.

Table 34: Usability evaluation of the IMMPACT statements in the context of the *NeuOst* trial protocol development. Adapted from Nielsen 1994. The following rating system was used: 0= the guideline does not fulfil the criterion at all; 1 = may superficially address some aspect of the criterion; 2 = addresses a small part of the criterion, but there are major aspects of the criterion that are not addressed; 3 = addresses a major part of the criterion, but there are some aspects of the criterion that are not addressed; 4 = fairly complete addressing of the criterion, but there is still more to be covered; 5 = complete addressing of the criterion. This evaluation refers to the published IMMPACT statement on pragmatic trials (Hohenschurz-Schmidt et al., 2023a) and the second paper on the topic as described in Chapter 5 of this thesis.

Usability heuristic	Rating (0-5)	Experience / Reflection
1. **Visibility of research steps:** The guideline should provide clear information on the stage of the research process to which recommendations apply. The user should be able to easily understand what stage of the research they are at, what needs to be done next, and what the outcomes are.	3	*The IMMPACT statement is clear in that it only is to be used during the planning phase of pragmatic trials, not later. There could be more clarity about the sequence in which decisions should be made or when informing decision-making by stakeholder engagement is most effective. In* NeuOst*, I began asking patient partners about the trial design after basic decisions were made, seeking their input mainly regarding the control intervention design. I do not think this affected the design negatively, but there was no guidance on project planning and sequencing in the IMMPACT statement.*
2. **Match between guideline and the real world:** The guideline should use terminology and concepts that are familiar and relevant to the user, and the research process should align with real-world research practices (including feasibility in light of practical constraints of clinical trials). The guideline should not introduce unnecessary jargon or complex terminology that might confuse the user.	3	*The IMMPACT statement includes a glossary of key terms but may not always be easy to follow for the non-invested reader. However, terms such as 'assay sensitivity', although used in expert consensus discussions, were deliberately avoided to enhance readability.* *The first IMMPACT statement presents alternative designs that are not currently used in pragmatic trials of pain interventions (such as adaptive or Large Simple Trials). Here, implementation of the guidance may be limited by commonly awarded funding, research expertise, and access to research Collaboratories able to support such large and complicated trials. For* NeuOst*, with its current support network, additional expertise would be required to conduct such a trial even if a large-scale funding should be obtained. In many other PPS*

		interventions, often lacking access to industry funding, this may also be unrealistic.
3. **Consistency and standards:** The guideline should use consistent terminology, formatting, and labelling throughout the document. There should be no conflicting messages. This consistency makes it easier for the user to understand the guideline and use it effectively.	3	*The IMMPACT statements are clear about what trials they refer to but suffer from the terminological and conceptual lack of clarity that exists regarding pragmatic trials at large* (see Pawson, 2019)*. This is, for example, evident in the discussion of how 'pragmatic' trials with placebo controls are, and readers will wonder why a sham-controlled 3-armed trial is discussed as exemplary (The reviewer certainly did).* *There is one instance where the first IMMPACT statement could have been clearer; namely when declaring that sufficient and high-quality evidence of efficacy should exist before pragmatic trials are conducted, while also mentioning that these "are regularly conducted in the absence of high-quality efficacy research". It requires careful reading to understand that this is an attempt to acknowledge current research practice but not an endorsement of it. In making the fundamental decision about whether to design* NeuOst *as a generally pragmatic or efficacy-orientated trial, I found the IMMPACT statement very useful and clear enough to come to a well-reasoned, unambiguous decision.*
4. **Recognition rather than recall:** The guideline should present information in a way that is easy for users to remember and apply.	2	*IMMPACT does not use metaphors or graphics to convey key messages. However, I found that some key messages are worded very succinctly, and they are also repeated multiple times: First, IMMPACT cites Ford and Norrie (2016) in that a pragmatic attitude to trials should not be equated with a "*laissez-faire approach to trial conduct*". Second, it is stated that "*Allowing and measuring heterogeneity where necessary but reducing it where possible is important.*" While I find that these sentences pointedly synthesise the statement's core messages, I note that many other recommendations are so detailed and nuanced that users will have to refer to the publications repeatedly and carefully to implement the guidance. Even as the guidelines' lead author, I had to revisit the documents many times to apply them to the* NeuOst *trial.*

5. **User control and freedom:** The guideline should provide the user with options to review and modify their research plan as needed. Users should be able to make choices about how to approach their research.	3	*IMMPACT primarily calls for a justification of research design decisions and for authors to make explicit their rationale underlying design choices. Most points are also framed as 'considerations' rather than 'recommendations', acknowledging the complexity of pragmatic trials and leaving room for autonomous decision-making. Therefore, if justified well, authors are free to design their trial as desired. Especially regarding the more detailed methods (such as choice of eligibility criteria, treatment standardisation, outcome measures etc.), I found IMMPACT helped me to decide which level of alignment with real-world clinical practice was appropriate, while enabling me to make decisions that suited my own study's context.* *Regarding the fundamental decision of whether to design a trial as generally more towards one or the other pole of the explanatory-pragmatic continuum, the IMMPACT statement will restrict the current practice in the pragmatic trials, if implemented correctly (by discouraging pragmatic trials in the absence of high-quality efficacy research). I found this clarity useful to guide an otherwise difficult choice, but it could be considered an infringement on 'freedom' as per this criterion.*
6. **Error prevention:** The guideline should be designed to anticipate and address common errors users might make, such as missing key steps in the research process. The guideline should provide clear instructions and guidance to help the user avoid mistakes.	4	*'Mistakes' in this context are difficult to ascertain. The IMMPACT statement seeks to reduce research waste by making trials more interpretable, through more control over or monitoring of heterogeneity. As safety mechanisms, IMMPACT refers to the PRECIS-2 tool on one hand – to guide design decisions on the explanatory-pragmatic continuum – and the Cochrane RoB tool on the other, to ensure that basic bias control is considered. Finally, core recommendations are listed in a one-sentence format. Together, these tools should ensure that no major design decision is made without considering its relationship with the trial's research question, its impact on generalisability, and on a trial's risk of bias. With the caveat that some 'mistakes' are likely to be spotted when conducting and analysing the* NeuOst *feasibility trial, I found this guidance sufficient for the development process of the study.*

7. **Help and documentation:** The guideline should provide clear and accessible documentation that explains the research process and how to use the guideline. This documentation should be easily accessible and provide answers to common questions and problems that users might encounter	1	*The IMMPACT statements do not provide a checklist or additional explanation documents. At many points, they also refer to external references for guidance on more complicated methods. Implementing the statement also requires a good understanding of the PRECIS-2 tool, but basic concepts of pragmatic trials are introduced as they refer to the content of the statement.*
8. **Other features.** As the heuristics framework was newly adapted and is not usually used to evaluate research guidelines, a 'other' field was included to allow for evaluation of unanticipated challenges.		*N/A*

Next, the thesis author's experience of using the CoPPS Statement to develop the control intervention of the NeuOst feasibility trial is presented (Table 35).

Table 35: Usability evaluation of the CoPPS Statement in the context of the *NeuOst* trial protocol development. Adapted from Nielsen 1994. The following rating system was used: 0= the guideline does not fulfil the criterion at all; 1 = may superficially address some aspect of the criterion; 2 = addresses a small part of the criterion, but there are major aspects of the criterion that are not addressed; 3 = addresses a major part of the criterion, but there are some aspects of the criterion that are not addressed; 4 = fairly complete addressing of the criterion, but there is still more to be covered; 5 = complete addressing of the criterion.

Usability heuristic	Rating (0-5)	Experience / Reflection
1. **Visibility of research steps:** The guideline should provide clear information on the stage of the research process to which recommendations apply. The user should be able to easily understand what stage of the research they are at, what needs to be done next, and what the outcomes are.	5	*CoPPS follows a clear timeline, reflected in its subheadings and infographic that cover the conceptual development of a control intervention, its pre-trial testing or practical development, the conduct of a trial, and the analysis, interpretation, and reporting of findings.* *Having focused mainly on the conceptual and practical development as part of this thesis, it was clear to me throughout where in the CoPPS Statement to find the relevant information and how information related to the steps of a research process. I also did not feel that details about important sequences of events were omitted.*
2. **Match between guideline and the real world:** The guideline should use terminology and concepts that are familiar and relevant to the user, and the research process should align with real-world research practices (including feasibility in light of practical constraints of clinical trials). The guideline should not introduce unnecessary jargon or complex terminology that might confuse the user.	4	*The guideline does not align with current practice, as demonstrated by the preparatory systematic review (Hohenschurz-Schmidt et al., 2023b), but postulates a new best practice. However, it addresses a common real-world challenge of trial design, as attested by the considerable interest in the project by involved experts, at conferences, or online. In doing so, it provides hands-on guidance on how to approach a complex issue, often acknowledging the inherent complexity and practical challenges. The CoPPS manuscripts were reviewed by a professional editing service, and I found all documents easy to read.* *CoPPS also engages in the difficult task of changing terminology, discouraging the terms 'sham' and 'placebo' control and encouraging the use of descriptive language and the phrase 'control interventions in efficacy and mechanistic trials'. As such, first-time readers might not immediately recognise what the*

		statement refers to, although terminology is clarified early on, including in the abstract. *Applying CoPPS to the NeuOst control intervention, I found it addressed all problems that arose. It also anticipated any concerns that were brought up in stakeholder interviews. Implementing CoPPS is, nonetheless, a complex and lengthy process, for example involving a review of the literature for existing control interventions but also hypothesised treatment mechanisms. As such, CoPPS may prolong the planning phase of PPS trials.*
3. **Consistency and standards:** The guideline should use consistent terminology, formatting, and labeling throughout the document. There should be no conflicting messages. This consistency makes it easier for the user to understand the guideline and use it effectively.	4	*I have not encountered instances where CoPPS recommendations were contradictory. The various documents are labelled in the same way and sections have consistent numbering for cross-reference. The large number of documents provided, however, may in itself limit the effective use of the guideline.*
4. **Recognition rather than recall:** The guideline should present information in a way that is easy for users to remember and apply.	3	*CoPPS distils the basic principle to the so-called 'similarity principle' for control interventions. I found this notable, and this one-sentence principle leaves little room for non-similar control interventions (such as detuned ultrasound for physical therapy). Further, however, CoPPS contains a multitude of detailed recommendations and considerations which will require constant reference to the document for implementation.*
5. **User control and freedom:** The guideline should provide the user with options to review and modify their research plan as needed. Users should be able to make choices about how to approach their research.	1	*Readers who follow the CoPPS Statement have most freedom in deciding which intervention components their trial wants to study. Having decided the objectives of the control intervention, however, the CoPPS Statement is largely prescriptive, making dozens of definite recommendations. Complying with those required considerable reading and contemplation, and I did not feel I had much room for manoeuvre if I wanted to comply with these recommendations as well as possible.*

6. **Error prevention:** The guideline should be designed to anticipate and address common errors users might make, such as missing key steps in the research process. The guideline should provide clear instructions and guidance to help the user avoid mistakes.	5	*The CoPPS checklist ensures the core recommendations are met and facilitates documentation of development processes. The infographic, and dedicated boxes (such as about provider training or stakeholder involvement) also highlight key steps.* *Unambiguous recommendations about the importance of measuring and reporting blinding should also mitigate the publication of poorly blinded studies, especially if journals adopt CoPPS as a go-to reference for relevant trials.*
7. **Help and documentation:** The guideline should provide clear and accessible documentation that explains the research process and how to use the guideline. This documentation should be easily accessible and provide answers to common questions and problems that users might encounter	4	*The Explanations & Elaborations document is directly accessible as a supplement to the main CoPPS Statement. It provides further detail and examples that were highly relevant when developing the NeuOst control intervention.* *Possibly, CoPPS will require additional clarifying documentation once it has been tested independently.*
8. **Other features.** As the heuristics framework was newly adapted and is not usually used to evaluate research guidelines, a 'other' field was included to allow for evaluation of unanticipated challenges.		*N/A*

Discussion and Limitations

The *NeuOst* intervention is the preliminary result of a multidisciplinary evidence-informed approach to intervention development. *NeuOst* aligns with other contemporary attempts to develop multimodal and psychologically informed interventions that can be delivered by manual or physical therapists (Carnes et al., 2017; Keefe et al., 2018). *NeuOst* also promises to broaden the scope of osteopathic practice, caring for people with a more disabling condition than most musculoskeletal pain, multimorbidity, considerable complication risk, and often socioeconomic disadvantages. In doing so, *NeuOst* stands in the tradition of a proud portfolio of community clinics at the University College of Osteopathy, where students have worked with people who are homeless or are HIV positive. Importantly, the *NeuOst* intervention is a response to the need for nonpharmacological options for people with neuropathic pain, reported clearly in the literature and by interviewed patients. Clinicians and academics showed considerable interest in the *NeuOst* project (for example at the German Pain Congress 2022) as did the public (as testified by an invitation to speak at the Royal Society of Medicine in March 2023). For people with pDPN, *NeuOst* promises the rigorous development and evaluation of a multimodal nonpharmacological treatment. Academically, the project is closely aligned with current research priorities in the pain field. For example, the Veterans Health Administration (VHA) called for "*determining optimal approaches for those with comorbid mental health or complex medical conditions*." (p. S13, Becker et al., 2018). If the *NeuOst* intervention was shown to be efficacious, osteopaths may be able to play a bigger role in collaborative multimodal models of care; another priority for development in pain research (Becker et al., 2018). Having been called to serve on the 2023 IASP Global Year Task Force on Integrative Pain Care (https://www.iasp-pain.org/advocacy/global-year/integrative-pain-care/), experiences from *NeuOst* will inform the thesis author's contributions to the task force.

Despite widespread enthusiasm, any discussion of *NeuOst* must acknowledge that the project is in its infancy: To date, there is no evidence of efficacy or safety, and whether a large-scale trial is possible is only just being investigated. Even if found to be efficacious, the *NeuOst* intervention would be challenging to implement in current osteopathic private practice models.

NeuOst may also not be cost-efficient by current NHS standards. A positive development for the intervention at this current stage would be a successful feasibility trial or one that can help optimise the intervention; funding to support a large-scale study is the best-case scenario. The exploratory implementation as a community clinic at the University College of Osteopathy is also planned, which would enable the research team to gain further insights into the intervention and develop the infrastructure for a larger RCT.

The protocol for the *NeuOst* feasibility RCT conforms with the definition of a randomised pilot or feasibility study, as per consensus definition from (Eldridge et al., 2016b), which are to test all or most elements of a future definite trial design on a smaller scale to see if the trial is possible.[12] An important part of the planned design is the feasibility and acceptability of the control intervention. Further, the study will provide important insights about the practicality of other design choices, including how their positioning along the explanatory-pragmatic continuum affects trial feasibility, generalisability, and internal validity. As illustrated in the supplementary Table 43 (Appendix 5: Supplementary information for *NeuOst* feasibility trial), the *NeuOst* feasibility trial falls mid-way onto the pragmatic-explanatory continuum (with an average rating of 3.25 out of 5 per PRECIS-2 domain), despite the efficacy-oriented primary question of a definite trial. As such, the feasibility study will also be useful to detect potential sources of heterogeneity in the data, allowing the evaluation of whether each source of heterogeneity can be accepted and monitored or whether it poses an unwarranted risk to a definite trial's internal validity (as recommended by the IMMPACT statement (Hohenschurz-Schmidt et al., 2023a)). For instance, it may become apparent that the *NeuOst* intervention is not suitable for people with high disability levels, for example because these people cannot attend sessions regularly, respond negatively, or cannot participate in the planned exercises. To study efficacy, the target population and the intervention theory may then have to be redefined. Other design choices to be critically evaluated include recruitment pathways, the intensity of the intervention, flexibility and

[12] Note that pilot and feasibility studies are another area of terminological ambiguity, the discussion of which is beyond the scope of this thesis.

adherence requirements, the permissiveness of concomitant treatments, and the insightfulness of the explored outcome measures.

Developing the *NeuOst* feasibility protocol with the IMMPACT and CoPPS frameworks was a good opportunity to practically apply these recommendation frameworks. Indeed, as a multimodal and very complex intervention, *NeuOst* was a realistic and a highly challenging scenario for application; one that embodies the doubts commonly levered against efficacy trials in PPS: Blinding in such interventions is said to be "notoriously difficult" (p. 943, Mataix-Cols and Andersson, 2021), and efficacy studies have been argued to be clinically meaningless (Fischer et al., 2014). While the ability to blind participants in the control intervention remains to be demonstrated, a high-similarity control intervention with clearly defined objectives was successfully designed thanks to the CoPPS Statement. It would also be clinically meaningful if *NeuOst* was found to not be efficacious in a well-powered efficacy trial, as such a result would discourage the clinical practice of manual therapy for people with pDPN. If efficacious, one could then embark on the comparative study of effectiveness, harms, and cost against the current standard of care.

Testing the usability of the IMMPACT and CoPPS frameworks through their first author is, of course, an approach with considerable limitations. Uninitiated readers will require more time to orientate themselves in these documents, they might not know the referenced papers and their content, and many readers will also be less immersed in the contemporary debates around clinical trial methods for PPS interventions. Conversely, experienced triallists will be more aware of the practical opportunities and challenges of clinical trials, again perceiving these guidelines and their usability differently. Independent feedback on the content and usability of the developed guidance statements will be required. For a start, the author is supervising two groups of osteopathy students at the *Haute École de Santé Fribourg* (Switzerland) who are applying the CoPPS Statement to two hypothetical trial scenarios of osteopathic techniques. Preliminary results indicate considerable challenges for novice researchers in working with the guideline. These students are also employing a different method of usability testing (a modified 'SWOT'

analysis) and their experiences will be helpful to inform future evaluation methods of guideline usability and content (which are currently lacking). In the coming years, developing a transferable approach to guideline usability evaluation and collecting feedback from independent users, appears to be the most rigorous approach to prepare updates of the CoPPS and IMMPACT statements.

For this thesis, however, the modified Heurisitc usability evaluation method provided an opportunity to reflect critically on the guidelines' strengths and challenges. It became clear that the more structured development of CoPPS has indeed resulted in a guideline that is easier to use, more comprehensive, and more aligned with the challenges on the ground of efficacy trials in PPS research. Extensive and structured stakeholder engagement and the dedicated investment of over 50 subject experts were crucial in facilitating this. The IMMPACT statements are less user-friendly and require cross-referencing with the PRECIS-2 tool (see supplementary Table 43 for an evaluation of design choices with PRECIS-2). However, especially the first IMMPACT statement was framed as a contribution to current debate rather than as a definite guideline (e.g., using the term 'consideration' rather than 'recommendation'). In retrospect, while the IMMPACT statement cannot be used in isolation to guide trial design, this approach appears more appropriate given the lack of empirical evidence regarding major points of contention, ongoing conceptual debate (Pawson, 2019), the continuous and possibly multidimensional nature of the pragmatic-explanatory spectrum, and the dependency of pragmatic trial design on the underlying research question. Different from the CoPPS Statement in challenge and ambition, the IMMPACT statements are nonetheless clear regarding fundamental considerations of study design (for example regarding the appropriateness of an explanatory of more pragmatic attitude to trial design) and provided a good overview of methodological challenges to consider during the design of the *NeuOst* feasibility trial. For example, the approach to align with clinical practice where there was little risk of design decisions impacting on internal validity, was highly relevant and led to the use of realistic eligibility criteria and patient-oriented clinical outcomes. The fact that the IMMPACT statements were able to guide decisions in a largely efficacy-oriented study further attests to the statements' broad scope.

In the usability evaluation, CoPPS was perceived as much more prescriptive. Although ambitious, the thesis author felt that CoPPS' clear-cut recommendations strengthened the reasoning during the *NeuOst* protocol development. An example is the recommendation to define treatment mechanisms and potential mechanisms through which the control intervention could affect outcomes. These CoPPS checklist items motivated the author to carefully review the manual therapy literature to identify current evidence on supposed mechanisms. In return, this knowledge enabled the researcher to decide which manually applied forces and techniques should be avoided in the control intervention. Also having identified gaps in the evidence of manual therapy mechanisms, the description of the control intervention was very clear about tested mechanisms and potential limitations of the chosen approach. Although the prescriptiveness of CoPPS resulted in a low usability rating for one domain, too much 'freedom' in the design of control interventions in efficacy trials may be what has led past researchers to sidestep difficult decisions or remain ambiguous in their reporting.

Overall, the *NeuOst* project was closely aligned with current priorities in pain research. For this thesis, the project enabled the important practical contextualisation of the CoPPS and IMMPACT frameworks. Planning an RCT was an important learning experience, as will be the study's conduct and analysis. For people with pDPN and for osteopathic practice, *NeuOst* explores new opportunities.

Discussion Section

Chapter 7: General Discussion, Limitations, and Outlook

Chapter summary

The project produced two sets of evidence-based guidance to support improvements in PPS trials across the pragmatic-explanatory spectrum. This guidance was found to be applicable to the design of a study of a newly developed multimodal intervention for those with painful diabetic neuropathy.

In this discussion chapter, the project's new contributions to knowledge are discussed in the context of other methodological literature, and its relationship to broader frameworks of complex and nonpharmacological intervention research and evidence-based practice is described. Additionally, the chapter situates the project outputs within current research priorities and identifies areas for future investigation. Limitations of the project are also addressed, and alternative approaches are discussed. Finally, the chapter reflects on the author's personal and academic learning journey throughout the project.

Discussion of main project outputs

Summary of main contributions

This PhD project started out with the question: *Which trial design is best for the study of physical, psychological, and self-management interventions for people with persistent pain;* having the ambition to develop '*best-practice trial methods across the pragmatic-explanatory continuum*'. During the project, the question was deconstructed into individual workstreams focusing on each end of the pragmatic-explanatory spectrum. This approach resulted in a competent discussion of methodological considerations across the spectrum, the shared outcome being that trial design choices must be informed by the underlying research question. The hypothetical triallist who, in the thesis introduction, was faced with the task to study the effects of a given PPS intervention, now has two published documents at hand to decide whether a comparative effectiveness or an efficacy- or mechanism-focused RCT is more appropriate. Once the investigation's direction is clarified based on what is known to date about their intervention of interest, the triallist can formulate the basic design. Whether deciding on a pragmatic or an efficacy-focused study, the IMMPACT and CoPPS recommendation frameworks then provide the researcher with detailed guidance on influential aspects of trial design, conduct, and reporting (Hohenschurz-Schmidt et al., 2023d, 2023a). More clarity was also provided by defining relevant terminology and optimising language with regards to control interventions. When applied, it is hoped that each recommendation framework can contribute to trials that are better able to answer their research question in a reliable manner. The interpretation of trial results is discussed in both recommendation frameworks, emphasising the need for trial authors to explicitly state their intentions, and justify design choices while considering internal *and* external validity. Additionally, the articles call for detailed reporting of methods and ancillary findings, including potential sources of heterogeneity in pragmatic trials such as concomitant treatment use, intervention adherence, and fidelity. In efficacy trials, there is a need for detailed reporting of control interventions and participant blinding data, now reinforced by the CoPPS checklists.

The two guidance frameworks were each informed by their own systematic review of trial methods and consensus-finding processes with subject experts and stakeholders. Both reviews showed that current methods were highly variable (Hohenschurz-Schmidt et al., 2022c, 2023b). In pragmatic and comparative effectiveness trials, methods occasionally conflicted with some authors' conceptualisation of 'pragmatism'. In efficacy trials, it was shown that the similarity between test and control interventions was linked to trial results. During consensus discussions, subject experts often debated the appropriate scope of each guideline. Some experts were inclined to highlight the potential value of the trials they were most familiar with, at pains not to devalue 'their' approach to the benefit of the other. Infusing pragmatic trial designs with internal validity considerations proved challenging, but the final guidance reflects successful implementation of this approach. When employing the frameworks to inform the design of the *NeuOst* feasibility trial, the IMMPACT statement was found to be relevant for a larger part of the trial design spectrum, the simultaneous consideration of internal and external validity being fundamental for most design decisions in trials. Conversely, the scope of the CoPPS guideline for control interventions is carefully defined as efficacy and mechanistic trials. At this end of the pragmatic-explanatory spectrum, CoPPS fills an important gap in guidance which, in itself, may have contributed to a preponderance of comparative effectiveness studies in some PPS areas (Hohenschurz-Schmidt et al., 2022c, 2023b)

Contextualising the main contributions within the literature

The IMMPACT and CoPPS Statements add to a large body of methodological literature and reporting guidelines for clinical trials. Noteworthy publications include the CONSORT Statement and its extensions for nonpharmacological and pragmatic trials (Zwarenstein et al., 2008; Boutron et al., 2017) which will continue to guide trial reporting. Trial interventions are best reported following the TIDieR reporting checklist (Hoffmann et al., 2014), for which a version for 'placebo controls' was published as the present project was ongoing (Howick et al., 2020). Reflecting the need to acknowledge unique challenges in complex intervention and

nonpharmacological trials, the TIDiER-Placebo checklist was expanded by the CoPPS Statement to reflect those challenges.

Numerous recent methodological publications have attempted to guide researchers in designing control interventions that address the challenges of binding and expectancy matching in nonpharmacological trials. In the field of manual therapy, D'Alessandro et al. (2021) proposed a 'touch equality assumption' for matching of touch-related features between study groups, also specifying how touch-based control interventions should be reported. Braithwaite et al. (2020b) Delphi survey led to recommendations for blinding in physical therapy trials, with 'indistinguishability' of interventions, matching of intervention components, and provider training at the heart of their framework. Most of these proposals build on early work by Baskin et al. (2003), who proposed the concept of 'structural equivalence'. In psychotherapy, recent authors continue to disregard participant blinding to group allocation and propose other ways to optimise blinding, such as blinding to study hypotheses and emphasising blinding of study staff (Mataix-Cols and Andersson, 2021). The CoPPS Statement is aligned with most tenets of these publications, but it is more broadly applicable, benefitting from a unifying principle, and it is accompanied by detailed additional documentation to guide decision-making and enforce rigour. To further explore the concerns in the psychotherapy field, a dedicated research project for the application of CoPPS is being planned with the psychology department at Aarhus University.

There is abundant methodological guidance available in the field of pragmatic trials, most prominently the PRECIS-2 tool and related publications (Loudon et al., 2015a; Zwarenstein et al., 2020). Despite the lead authors claiming otherwise (Zwarenstein et al., 2017), theirs and many other publications (Treweek and Zwarenstein, 2009; Zwarenstein and Treweek, 2009; Rowbotham et al., 2013; Gordon et al., 2020a) convey the impression that 'pragmatism' is always a desirable goal for RCTs.[13] Relatedly, discussing pragmatic trials without describing how to maintain the internal validity of trails conveys the impression that generalisability and internal

[13] This impression arises not least from publication titles such as 'Making trials matter' (Treweek and Zwarenstein, 2009) or that speak of 'trials with purpose' (Loudon et al., 2015a; Gordon et al., 2020a).

validity are always unrelated. This depiction oversimplifies trial design, which is a complex endeavour where a multitude of design choices can be made, only sometimes independently from one another. While the authors of the PRECIS-2 tool acknowledge that issues of internal validity lie outside its scope (Zwarenstein et al., 2017), this may be a weakness of the framework. The IMMPACT statements are among the few publications that consider each design decision freely on the entire pragmatic-explanatory spectrum. As opposed to PRECIS-2 but aligned with calls by Nicholls et al. (2020), the IMMPACT statements instead emphasise the need to consider which research question a trial aims to answer and how its objectives can be achieved in a reliable manner. In line with Pawson's work (2019), the project has not found evidence of a distinct and well-defined entity called 'pragmatic trial'. IMMPACT raises concerns about elevating 'pragmatism' to an ideal in order to generate clinical trial findings that are both interpretable and useful. While it is challenging to determine a clear trend, some recent publications not associated with the PRECIS group, such as Zuidgeest et al. (2017), Wright et al. (2021) and Keefe et al. (2022), also highlight the significance of internal validity in the context of pragmatic trials.

Conversely, the fact that IMMPACT, an organisation closely tied to the safety and efficacy-focused FDA and with links to several pharmaceutical companies, chose to focus on pragmatic trials is noteworthy for two reasons. Firstly, it acknowledges that well-conducted effectiveness-focused trials have value even in drug research, where (comparative) effectiveness is not a formal requirement for market approval (Kalkman et al., 2018). Secondly, nonpharmacological trials are increasingly important in pain research, and pragmatic trials are particularly prevalent in this area, as indicated by this project's review (Hohenschurz-Schmidt et al., 2022c). IMMPACT's efforts to improve methods in pragmatic trials can be viewed as recognition of this trend.

Contextualising project outputs as evidence-based practice approaches

Apart from their direct relationship with other methodological guidance, the developed guidelines can be situated within broader frameworks of evidence-based practice and their underpinning philosophical stances. For example, in their underlying attitude to nonpharmacological intervention research, the IMMPACT and CoPPS Statements are related to the MRC framework for complex intervention development and evaluation, and many other frameworks that describe complex intervention research as circular (Walach et al., 2006; Skivington et al., 2021) or integrative (Jonas, 2005). In these frameworks, different types of knowledge are acquired from a multitude of complementary research designs (Fischer et al., 2012), for example knowledge about treatments' mechanisms, efficacy, safety, effectiveness, cost, personalisation, meaning, and implementation. The current PhD project focused on enhancing the quality of clinical trials that address mechanistic, efficacy and effectiveness questions and guides decision-making about trial design in the context of a given intervention. However, it is essential to acknowledge that areas not targeted by the CoPPS and IMMPACT statements, such as mechanistic experiments, observational studies, implementation research, and evidence synthesis, may require similar efforts to assess methodological quality and identify areas in need of improvement. By systematically improving research designs across different areas of complex intervention research, a more comprehensive understanding of interventions and their impact can evolve, and research waste minimised.

In evidence-based practice, clinical practice is seen as compatible with scientific experimentation (Leach, 2006). However, some authors and practitioners from the 'Complementary and Alternative Medicine' field (Keshet, 2009) or adherents of 'traditionalist' modes of practice (Bordoni, 2019; Simpson and Young, 2020) argue that 'holistic' interventions are not amenable to 'reductionist' research methods, or that their personal experience outweighs scientific evidence. In contrast, the present PhD work takes an empiricist stance, which asserts that scientific knowledge should be derived from systematic observation or experimentation. Positivism emphasises the importance of objectivity, rationality, and logical

analysis in scientific inquiry (Guba and Lincoln, 1994; Shariff, 2015). In clinical trials, this means that researchers strive to eliminate biases and confounding factors that could affect the validity of their results – an endeavour clearly reflected in the outputs of this PhD project. The goal is to produce objective knowledge that can inform clinical practice, also in the field of PPS interventions. Nonetheless, as demonstrated by the stakeholder engagement activities of this project, the subjective experience of patients can guide research endeavours at many key points.

Both empiricism and positivism are underpinnings of the evidence-based practice (EBP) movement, which supports the use of current best evidence to guide clinical decision-making (Leach, 2006). Showing the compatibility of (post)positivism and constructivist philosophies (Lee, 2012), EBP models of care nonetheless value the subjectivity of patients and advocate for well-informed, shared decision-making as best practice (Hutting et al., 2022). As Leach (2006) argues, EBP should also attempt to bridge the "*research-practice divide*" (p. 248). While this is also one of the main motivations for pragmatic trials (Ford and Norrie, 2016; Gordon et al., 2020a), implementing evidence-based practice remains challenging due to the sheer complexity of clinical practice (Djulbegovic and Guyatt, 2017). Decision-making in the face of uncertainty will always require heuristics, and thus be vulnerable to bias (Croskerry, 2009; Costa et al., 2022). Personal preferences of patients and practitioners are important parts of the therapeutic relationship; and we make sense of the world through experience, not usually by reading. Weighing preference and experience against scientific 'gold standard' evidence can be ethically and personally challenging (Straus et al., 2018). Rather than discarding scientific evidence, however, there are contemporary philosophies that attempt to reconcile the complexity of clinical practice with the tenets of evidence-based medicine. For example, dispositionalism describes causation as highly complex and context-sensitive. Taking a dispositional stance has been proposed as a suitable for PPS practitioners, enabling them to deliberately use different types of evidence for individual patients and make sense of the complexity of clinical practice (Rocca and Anjum, 2020; Thomson and MacMillan, 2023). While dispositionalism may be a suitable model to approach complexity at the individual level, however, clinical trial research requires larger groups and the use of group-averages to ensure the validity of results and a

certain level of generalisability.[14] Even when analysing *sub*groups, clinical trials thus inevitably miss some of the complexity, but their results can still inform clinical practice (Djulbegovic and Guyatt, 2017). Pragmatic trials and other research designs discussed as part of the IMMPACT publication may be able to reduce the research-practice divide, but they will not obliterate it.

Contextualising the project with current priorities in pain and nonpharmacological intervention research

Having situated the PhD outputs within scientific, conceptual, and philosophical contexts, it is worth evaluating how the project outputs addressed current research priorities in the pain field. This will also facilitate the discussion of remaining and future avenues for investigation.

Bushnell et al. (2021) list several challenges and opportunities for the nonpharmacological treatment of pain; and so do other contemporary author groups (Becker et al., 2018; Kerns et al., 2019). Proposed research agendas and challenges include:

- Improving terminology and public understanding around 'nonpharmacological' treatments for pain (Bushnell et al., 2021)
- Developing better control interventions for clinical trials (Becker et al., 2018; Bushnell et al., 2021)
- Improving mechanistic understanding of interventions, including the relevance and interactions of constituent components in the context of individual patients or patient subgroups (Bushnell et al., 2021)
- Promoting and investigating the implementation of therapies "*with strong efficacy/effectiveness evidence*" (p. S12), focusing on barriers to uptake and impact of various delivery methods (Becker et al., 2018; Kerns et al., 2019)

[14] Single-case experimental designs (a type of single-subject or 'N-of-1' studies) may enable the granular appreciation of an individual's unique response to an intervention, but it is questionable whether they can produce insights about cause and effect with the certainty and relative generalisability that randomised controlled trials can (Cook et al., 2002; Kratochwill et al., 2010; Tate and Perdices, 2020).

- Integrating of technological advances to increase access to therapies and support research, in particular using telehealth, virtual reality, wearables, and mobile devices (Becker et al., 2018; Bushnell et al., 2021)
- Developing and studying multimodal nonpharmacological and integrated pain management programmes, including collaborative and stepped care models (Becker et al., 2018; Kerns et al., 2019)
- Supporting tailoring approaches through prognostic factor research and the use of innovative trial designs (Becker et al., 2018)

And finally:

- Exploring the role of nonpharmacological modalities in the care of understudied populations, mainly people with multiple and other chronic conditions, non-musculoskeletal pain, children, the elderly, and people from minorities and disadvantaged communities (Becker et al., 2018; Kerns et al., 2019)

From this list, it is apparent that the present PhD project was well-aligned with research priorities in this rapidly growing field: Primarily, the project made important contributions to addressing the problem of control interventions in PPS trials and explored pragmatic trials from a critical perspective, making suggestions that should enhance the rigour of such trials. Beyond that, the work on control interventions bears relevance to an improved mechanistic understanding of interventions. As described in Appendix 7: Outreach activities and additional contributions, several side-projects also explored placebo and nocebo effects as important factors in PPS interventions. The thesis author has also explored novel modes of treatment delivery, both through facilitating the adoption of remote consultations by manual therapists and through engagement with the digital therapeutics industry (Appendix 7: Outreach activities and additional contributions). The author's work on developing a multimodal intervention for people with pDPN conforms with the research priority to explore nonpharmacological treatments in underserved groups of people with pain and comorbidity. Having begun in the final year of the project, this research will be a focus of the immediate post-doctoral phase.

Future opportunities and remaining challenges for PPS research

From the above list, it becomes apparent that there remains room in the author's future research agenda to incorporate other research methods for an enhanced understanding of treatment mechanisms and to explore personalised approaches. Mechanistic research may include reviving a collaboration with the pain neuroimaging group at the Institute of Psychiatry, Psychology and Neuroscience at King's College London; or making use of the Rice lab's expertise in Quantitative Sensory Testing (QST). Single Case Experimental Designs which explore personalisation options of *NeuOst*, or adaptive elements in a definite RCT, are also possible. Finally, the thesis author is strengthening his connections amongst researchers of the placebo effect, which may also contribute to a better mechanistic understanding of physical and psychological interventions, but also allows the exploration of patient preferences, ethical issues, and interpersonal factors in nonpharmacological healthcare.

Apart from personal opportunities for future research, the PhD project has made apparent which developments in PPS intervention research and practice are required to successfully enhance access to nonpharmacological pain care and promote research quality.

Fundamental to achieving these aims is a clear understanding and an addressing of differences between drug and non-drug research; eventually enabling decision-makers to incorporate these considerations into their judgement of which treatments should be made available. The project has made inroads towards the clear methodological delineation of drug and non-drug research. For example, the project produced guidance on control interventions that is specific to complex PPS interventions. It highlighted nonpharmacological interventions as distinct from drugs in the IMMPACT statements and proposed solutions for pertinent challenges in pragmatic trials. While these frameworks may promote better-quality research in the PPS field and may reduce research waste (Buchbinder et al., 2020), to fully address the problem, changes at other levels are also required, including at policy level. First, the problem of (non)comparability between drug and non-drug trials needs to be better described and quantified. As described in the CoPPS and IMMPACT statements, there is conflicting evidence on how effect sizes may differ

between these types of trials. How differences in effects relate to interventions' safety profiles, cost, and patient preferences, including in subgroups, also needs to be explored. Similarly, the relative contribution of the placebo effect in PPS trials needs to be investigated and compared to drug studies. To fill this gap in knowledge, the author is currently leading a review of three-armed RCTs of PPS interventions (See Appendix 7: Outreach activities and additional contributions). Second, evidence synthesis methods need to be developed that acknowledge necessary differences between research on different types of interventions but also allow for the direct comparison of trial quality across drug and non-drug studies. This could be achieved, for example, by grading each trial against a best-practice standard that is specific to the type of trial, not against an absolute standard as is currently the case with risk of bias tools. Thirdly, when a better appraisal of trial methodology is possible, evidence synthesis for clinical and policy decision-making will also be enhanced. Given that such guidelines often incorporate not just evidence-based considerations but also expert opinion (Blazey et al., 2022) and healthcare political aspects, education of these stakeholders regarding best-practice approaches in PPS research may also be required. Similarly, methodological rigour of nonpharmacological research is also essential for the integration of PPS approaches with standard medical care for patients with complex conditions and multimorbidity; an area where integrated publicly-available care programmes remain an exception (Whitty et al., 2020). Finally, and paralleling drug research in this instance, methodological challenges are not the only source of poor research quality. The project has not touched on several other concerns relevant to rigour in clinical trials, such as spin in trial reports or publication bias, but other authors have (Boutron et al., 2019). Another important topic not addressed in this PhD project is a potential lack of research integrity of individual researchers, recently publicised effectively by O'Connell et al. (2022) in the psychotherapy field. Such lack of trustworthiness may include the deliberate falsification of studies. Relatedly, authors increasingly publish in predatory journals, which is a concern especially in the smaller PPS professions with little resources for research training and governance (Mills and Inouye, 2021). Not only does this undermine scientific standards (Paci et al., 2022) and distort evidence in systematic reviews (Hayden, 2020), but predatory journals also enable the growth of a pseudoscientific 'parallel' research community (Beall, 2016). As exemplified by the osteopathic field, some authors avoid

more rigorously peer-reviewed journals (e.g., Bordoni and Escher, 2023; Consorti et al., 2023), likely for a variety of reasons (Mills and Inouye, 2021). To investigate the extent of this problem in manual therapy and develop solutions, the thesis author and the National Council for Osteopathic Research have recently obtained dedicated project funding to study the veracity of published trials (See Appendix 7: Outreach activities and additional contributions).

Individual PPS professions and their higher education institutions could drive positive responses to these challenges. Notably, they could foster research collaboration and mentoring, support pan-professional frameworks for better research methods such as CoPPS, cultivate evidence-based practice and critical appraisal skills amongst professionals, engage stakeholders, lobby policy-makers to address stakeholder needs and values, and develop innovative treatment approaches that address needs and ensure equitable access to high-quality person-centred care (Finucane et al., 2023). The growing interest in nonpharmacological pain treatments benefitted the present project enormously. Potentially, it might also facilitate these additional positive changes.

Limitations and alternative approaches

The project's main limitation is the absence of a practical testing phase prior to the publication of each of the developed guidance frameworks. This was likely not feasible as part of a three-year PhD project, although recommended in the development of clinical practice guidelines (Brouwers et al., 2010). More practical testing would have likely been useful to identify gaps in the developed guidance, feasibility problems, or incompatibility with research regulations. The successful application to the *NeuOst* feasibility trial is reassuring, and the large number of experienced triallists involved in both endeavours mitigates these concerns somewhat, but structured and independent evaluations will be required in the preparation of guidance updates.

It can also be argued that by not focusing on an individual intervention, generalisations were made that disregard the distinctiveness of individual therapies. For example, during the application of the IMMPACT and CoPPS frameworks to the development of the *NeuOst* protocol, the supposed mechanisms of manual therapy interventions had to be researched separately. Conversely, however, the supposed broad applicability of both frameworks is a strength because it may promote research quality across different research fields and to the same standard, thus also facilitating comparisons between different PPS intervention trials.

Another limitation applicable to all project outputs is the minimal involvement of stakeholders from non- 'Western' countries. Despite a patient partner from sub-Saharan Africa in the CoPPS project and the inclusion of a few studies from East Asia in the systematic reviews, it cannot be guaranteed that the CoPPS and IMMPACT statements are compatible with research environments, stakeholder values, and therapeutic particularities in places other than the US, Europe, and Australasia. Retrospectively, this should have been done better, for example by the targeted involvement of the International Association for the Study of Pain's various global chapters or by facilitating involvement through translation of the research into other languages.

The project produced two, eventually three guidance documents, arguably adding twigs to what already is an almost impenetrable thicket of guidelines and methods papers. The biggest obstacle to the effective use of methods papers is likely whether they are found by the person who requires them. While libraries of reporting guidelines such as the EQUATOR database improve searchability somewhat, a comparable collection of methods statements and papers does not exist. The visibility of this project's guidance may however be relatively high, because the frameworks were published as open-access articles and under the name of a renowned organisation such as IMMPACT or, respectively, in *The BMJ*.

Alternative approaches to methodological research would also have been feasible. For example, rather than surveying current methods and producing guidance based on expert consensus, practical experiments with different blinding methods could have been conducted.

Similarly, 'Studies Within A Trial' (SWATs) (Treweek et al., 2018) could have evaluated the impact of different design decisions in pragmatic and other RCTs. Although they would likely not have resulted in comprehensive guidance, these methods may have been less prone to simply replicating current practice or to 'expert' bias (Noting, however, that both the CoPPS and IMMPACT frameworks deviate considerably from current standard practice and recommendations are largely backed up by empirical evidence). Overall, the systematic evaluation of current methods and stakeholder-driven development of guidance likely had the largest potential to find broadly applicable solutions to the multitude of challenges in PPS trials.

Reflecting critically, the IMMPACT statements could have dismissed the arbitrary distinction between explanatory and pragmatic trials more clearly, for example by not using the term 'pragmatic trial' due to its lack of conceptual clarity (Pawson, 2019; Taljaard et al., 2022). This was, however, not possible given the opposing views in the author group and may also have adversely affected the paper's recognisability. The CoPPS Statement was bolder in this regard, discarding 'sham' and 'placebo control' terminology (including from its title). This may, in turn, mean that some authors do not immediately recognise the guidance as relevant to their research problem. Additional limitations specific to individual aspects of the project have been discussed in respective chapters, including shortcomings of the systematic reviews or the employed methods particularly in the development of the IMMPACT statements.

Learning journey and reflections

The main personal lessons learnt from this PhD project are:

1. Attention to methods is a fundamental research skill.
2. Methodological research is sought after, important, and has the potential to bring together researchers from different fields.
3. Collaboration is a powerful and hugely rewarding way of conducting research, but requires elaborate managerial, organisational, and communicative know-how.
4. People with lived experience of a studied phenomenon contribute invaluable insight to a research project. Ideally from conception to dissemination, working with patient partners and other stakeholders will always be a fundamental building block of my future research endeavours; and one that brings a lot of joy and learning. At the same time, PPI needs to be costed for, as remuneration is fundamental to facilitate access and for demonstrating to patient partners that their expertise is valued.
5. Producing research that is applicable to the realities of clinical practice is challenging. I do not want to miss my personal clinical experience which regularly reminded me of the complexity of clinical practice and that the ultimate beneficiaries of academic endeavours must be human beings, if not humanity at large.

I am conscious of the privileged position from which I commenced my research journey. Not only does this refer to my personal circumstances and background, but also to what is best described by the German word *Vertrauensvorschuss*: literally translated, an advance portion of trust; My supervisors, at a stage when I had barely begun working with them, introduced me to the IMMPACT network and later supported my ambition to recruit high-profile experts into the CoPPS collaboration. I have learnt from this that students need to be given opportunities and a *Vertrauensvorschuss*, irrespective of the possibility that their endeavour might not succeed.

In the future, I would like to focus more on projects that address socioeconomic inequalities and advocate for disadvantaged populations, having realised, partly thanks to this

project, that questions of health and social structures are inseparable. Researchers can improve methods and interventions to enhance access, and clinicians can reduce barriers to care provision, but a truly biopsycho*social* approach to healthcare must include advocacy work and transferring power to disadvantaged communities (Marya and Patel, 2021).

The apparent challenge of involving people from non – 'Western' countries and local minorities in research has become starkly apparent in the formation of my research networks for this project. In the future, I will have to develop connections that can help bridge cultural and language barriers if the aim of producing truly broadly relevant research is to be achieved. There are people within the International Association for the Study of Pain to learn from, and strong chapters from outside Europe, the US, and Australasia. My involvement in the Global Year Task Force 2023 (See Appendix 7: Outreach activities and additional contributions) may prove to be a stepping stone for such networking, especially due to the focus on integrative pain care and thus the involvement of other PPS clinician-academics. When conducting clinical trials, local partners are needed to facilitate equitable involvement.

Overall, I learnt many academic and personal lessons which made the PhD project a hugely rewarding experience.

Conclusion

The PhD project provided a nuanced answer to the question of "*Which trial design is best for the study of physical, psychological, and self-management interventions for people with persistent pain?*", enabling researchers to decide whether to employ an efficacy-focused or more pragmatic attitude to trial design. Detailed methodological guidelines for randomised control trials across the entire pragmatic-explanatory continuum were developed. Notably, the CoPPS Statement filled the gap of best-practice considerations for control interventions which has been a main point of critique for efficacy trials of PPS interventions in the past. The IMMPACT statements contribute a strong focus on the balance between internal and external validity to the discussion of pragmatic trials and provide detailed considerations for pragmatic trials in pain research. Guidance was formulated in relation to current research practice and by drawing on empirical evidence and various stakeholders' expertise. In a first evaluation, these guidance frameworks were found applicable to and useful in the design of a trial studying a multimodal intervention for people with painful diabetic polyneuropathy. By drawing on the newly developed guidance, a control intervention was developed, although it remains to be tested whether the control intervention achieves its objectives. While the project had some limitations, such as insufficient expert involvement from regions outside the US, Europe, and Australasia, it aligned well with contemporary research priorities in the field of pain and nonpharmacological intervention research. This is not least demonstrated by the considerable interest from the scientific community and several publications in high-profile journals. Future research aims include the practical validation of the guidance frameworks. For the field at large, identified research needs include the improved understanding of treatment mechanisms, the adaptation of evidence appraisal processes to the particularities of nonpharmacological intervention research, and more projects that enhance access to high-value care for people from underserved populations.

Acknowledgements

I am forever grateful to all the colleagues, supervisors, friends, and family members who supported, inspired, and suffered me on this quest to becoming an academic. Specifically, I would like to thank Prof Andrew Rice, Dr Jerry Draper-Rodi, Dr Whitney Scott, and Prof Lene Vase for their expert supervision, which was both intellectually stimulating and fun. Each member of the supervision team contributed their own skillset and personality, which made the project an incredibly enriching experience. Thank you also to the Pain Group members, especially Drs Harriet Kemp, Nadia Soliman, Jan Vollert, Matt Evans, Donna Kennedy, and Sara Dixon-Smith. You are an unbelievably smart bunch and I love working with you! Next, there is the Research Department at the University College of Osteopathy, notably Dr Oliver Thomson, Steve Vogel, and Andrew MacMillan. Digging into the particularities of osteopathy will likely be a never-ending story and it is good to know that there are such trustworthy and critical travel companions. A big HELLO also to baby Annabelle Grace, Andrew's daughter who will forever have been born on the day this thesis was submitted: Here's your first citation! Thanks also to Sascha Smith for sharing my PhD student struggles and the support (and coffees) when needed. I am also grateful to other researchers and young academics who supported me along the way, notably the IMMPACT and CoPPS collaborators, the ICTU team, and Profs Robert Dworkin and Dan Cherkin. Dan, thank you for the kind academic and personal mentoring, and Bob, thank you for trusting me with this project!

The early stages of this project lived and thrived thanks to a wonderful 'reviewer crowd', consisting of many individuals with names and needs who nonetheless supported this project. It was great fun and I hope you all gained as much from the experience as I did. A shout-out to Jules Phalip and Jessica Chan. You are great people and I look forward to seeing where your journeys will take you! A sincere thanks also to all *NeuOst* collaborators for saying 'Yes' to a slightly different project. I look forward to taking this further! Thank you to all the clinicians and patient partners who trusted me and allowed me to learn from their experience. Amongst those, a huge

thanks goes to Liz Pigott for your kindness and support – I learnt more from you than from many a professor!

Thank you to Amira Korkor and Ali Ansaripour for hosting me in Oxford for a much-needed writing retreat, the rest of the NeuroGeek crew, and Dr Naiana Suter for the formatting tips. Mardi Jamieson, thank you for coping with such an unreliable clinician. I needed that and thank you for making my life a little easier by being flexible and relaxed. Cheers to my friends for pretending you were interested – at least sometimes – and for the distraction when needed.

My parents Uschi und Ralph, and my brother Christoph saw me through to this point. Never doubting or questioning and always supportive in so many ways. Absolute bedrocks.

Finally, my wife Greta, and my two sons Anton and Emil: Thank you for being there and for everything that you are: loving, kind, fun, smart, ambitious, understanding, and definitely mad. I love you.

Appendices

Appendix 1: Systematic review and crowd-related content

In this appendix, an example of the collaborative working style via Microsoft Teams is provided in Figure 14.

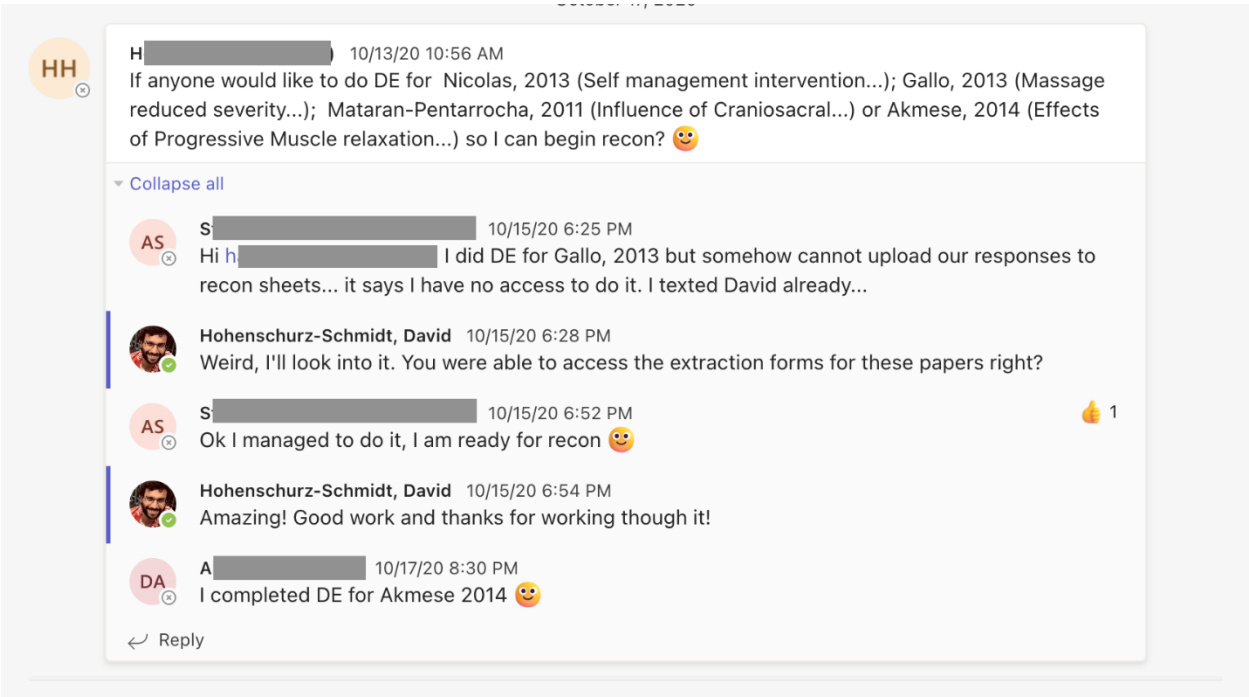


Figure 14: An example of the use of Microsoft Teams chat space for the organisation of data extraction between crowd-members. Anonymised screenshot. DE = data extraction, recon = reconciliation of data extractions.

The email excerpts below illustrate the experiences of crowd members, sent in response to the notification that our systematic review had been accepted in the journal PAIN.

"*Congratulations everyone and thank you for the opportunity to work with you all! Hopefully, we have many more to come. This is a massive achievement so thank you David for leading this project and all the training and persistence*."

"*WELL DONE YOU! You've been amazing at managing the whole team, it must have been really tough, but you did a fantastic job. It was inspiring to work for you, and I hope that I'll be able to write a paper of this quality one day! I'm looking forward to celebrating these TWO publications around a beer with you one day*."

"*I am so pleased to hear this!!! Thank you so much for the amazing opportunity, it was such a pleasure to work with you and the team. Really looking forward to seeing these publications, and I hope to work together again at some point in the future :) Hope all is going well with you and have a great weekend!*"

"*It was an amazing experience to work with the team, and you are a charismatic leader!*"

"*Thank you David and congratulations to such fantastic publications! I am very happy to have been a part of this."*

Appendix 2: Consensus methods, additional content

This appendix presents additional detail on consensus development methods for the CoPPS Statement as well as participant feedback.

Round 1 Delphi questionnaire

The Delphi's round 1 survey items can be grouped as follows:

- Introduction and glossary (not rated)
- Demographic data of panellist
- General study design
- Blinding
- Conceptualisation of 'placebo interventions'
- Similarity between control and study intervention across 29 criteria, characterizable into the following domains:
    - 'Non-specific' or contextual factors
    - Treatment structure and application
    - Therapist characteristics
- Other methods to enhance blinding
- Assessment of control intervention credibility and blinding success
- Control intervention development and testing
- Suitability for trial purposes
- Reporting of control interventions
- Open-ended questions

Open-ended questions were:

- *What else do you recommend for the development of a placebo control intervention?*
- *Which other characteristics should a placebo control intervention have, compared to the active intervention under investigation?*

- *What else do you recommend for carrying out a placebo control intervention during a trial (i.e., what is important when implementing the protocol and running a trial?)*
- *What else do you recommend in order to enhance participant blinding in a trial?*
- *What else do you recommend for the reporting of a placebo control intervention?*

Note that the consensus process resulted in a change of terminology around 'placebo' and 'sham' controls further discussed below, so that these terms were avoided in the guidance statement. In this thesis, I have similarly used the description 'control interventions in efficacy or mechanistic trials of PPS interventions' instead. In the Delphi survey, however, the term 'placebo control' was still used, with a shared understanding that we were discussing so-called 'placebo', 'sham' and 'attention controls'.

Patient interview script

Table 36 Patient consultation interview script, including hypothetical clinical trial scenarios ('vignettes') that were created based on Delphi-derived consensus items.

Demographic detail
Demographic questions based on UK national census 2021, see online methods file for detail: https://osf.io/jmyhq/ What is your experience with pain? Do you have any personal previous experience as participant in clinical trials: Y/N Please provide any other information you would like to share and consider important:
Introduction
Please tell me a little bit about yourself. Which treatments or self-help approaches have you tried for your condition?

Do you have any previous experience with clinical trials? If so, what? If not, what is your understanding of clinical trials? (Provide explanation of placebo / sham-controlled RCT if required)

Vignettes

1st Scenario (Sham controlled trial, general)

Please consider the possibility of participating in a clinical trial of a physical, psychological, or self-management pain therapy (for example, exercise, manual therapy, or cognitive behavioural therapy). Before deciding to take part, you are told by the research team that you will be assigned either to a group that receives the treatment or to a group that receives a "sham" or "placebo intervention".

What are your thoughts, hopes or concerns?

Further explanation: A sham treatment is an intervention that copies the real treatment of the trial but lacks some important aspects – those aspects of which the researchers think that they make the patients better. (Provide relatable example if needed / desired)

How does that make you feel?

What would you like to comment?

What would be potential barriers to you participating in such a trial? What are enablers?

2nd Scenario (Blinding)

Once you signed up to the trial, you are randomly allocated to the real treatment or the sham intervention. You are asked to attend several sessions and some pre- and post-trial assessments (highlight the potential time burden), but nobody will tell you which group you are in (Chances are usually 50:50).

What are your thoughts, hopes or concerns?

Would you attempt to find out in which group you are? If so, how would you go about it?

What do you think could tell you / give it away to you that you are in the intervention or sham control group?

3rd Scenario (lack of provider blinding)

In trials of physical or psychological interventions, the therapy providers usually know whether they are providing the real or the sham treatment to you. That is not the case in drug trials, where doctors usually do not know in which group their patients are.

How do you feel about that?

What are the implications for you?

4th Scenario (Testing blinding effectiveness)

During the trial or at the end of it, researchers may ask you to guess whether you think you received the real or sham treatment.

What are your thoughts, hopes or concerns?

5th Scenario (Unblinding)

Imagine that at some point during the trial you become convinced that you are not receiving the real treatment but the sham.

What would be the implications of this for you?

After the trial, do you prefer to be told in which group you were or not? Why?

At the end of the trial, the researchers tell you that you have received the sham intervention.

How do you feel? What does this mean to you?

Would you participate in future studies of this kind? Why?

6th Scenario (Sham development and testing) *Developing a credible sham intervention in nonpharmacological trials is challenging for researchers because it needs to feel like a real treatment to patients but lack certain elements that the scientists want to study in the trial – just like a placebo pill looks and tastes like a real drug but lacks the pharmacological substance.* What do you think about involving people with lived experience (patients) in the development and testing of a sham intervention *prior to using it in a real trial*? Potential follow-up question: Why do you (not) support stakeholder involvement? What are potential benefits / drawbacks?
Conclusion
Is there something you would like to add?

Summary of consensus discussions

Table 37: Summary of discussion points and main arguments made during the consensus workshops. Exemplary anonymised quotes from meetings participants are included.

Topic	Main arguments
Glossary / terminology	The terms 'sham' and 'placebo intervention' were considered problematic by most attendees. Potential solutions put forward were to 1) simplify language by speaking of 'control interventions' rather than 'sham', and 2) change the sequence of presenting terms and explanations to patients, i.e., first describing the control intervention and its concepts and then potentially informing participants that this is also known as 'sham' or 'placebo'.

	Quote: "…*move away from using this kind of imprecise terminology, and instead just describing what you do and what you have done*." There was support from a patient representative for speaking of 'control interventions' and focussing on describing content.
Scope of the final guideline	Attendees agreed that this guidance applied after researchers had decided that a 'sham'-controlled RCT is appropriate. The focus being on efficacy and mechanistic trials, the manuscript should mention possible research questions. For other potential comparator arms and how to choose the right type of trial, it was agreed to provide suitable references only. Also, it was deemed out of scope to make recommendations for the study of placebo effects. Quote: "*the sham or placebo part is not necessarily ineffective and very often will have effects (…) so it's really to stress that it is the additional effect of the active ingredients of interest that is being studied here*."
Principles of control interventions in efficacy and mechanistic trials	Attendees agreed that the proposed concept of control interventions in efficacy and mechanistic trials must acknowledge multiple intervention components. As a design principle, we agreed to postulate that a control intervention should consist of most of those components but not the one(s) that the trial sets out to investigate.
Objectives of a control intervention in an efficacy or mechanistic RCT	Objectives of control interventions in efficacy or mechanistic trials were defined as: 1) To test the mechanism of intervention components (acknowledging the potential necessity to conduct iterative trials and dismantling designs) 2) To balance expectations between trial arms (mainly through blinding and by producing similar contexts) It was agreed that it was not usually an objective to study real-world effectiveness.

	Attendees wanted to clarify in the manuscript that a main criticism of complex intervention RCTs is the challenge of (patient and provider) blinding. It was agreed that this guidance aimed at overcoming these challenges by accepting that perfect blinding is often not possible but postulating a way to still produce comparable expectations between groups with regards to the therapeutic benefit / outcomes. Specifically, the proposed design principle to achieve this is to make tested and control interventions as similar as possible on every aspect but the component(s) to be studied. Quote: "*Yes, blinding is essential, but I question if it is a principal objective. The principal objective is to balance expectations*." (The attendee then explained this notion in relation to placebo science.)
Fear of devaluing other kinds of trial designs	There were fears that our guidelines could lead to overly 'effective' control interventions that might lead to increasing amounts of unsuccessful trials, thus 'devaluing' trials. It was agreed that there should be nuanced explanations around how to interpret these trials (out of fear of delegitimising contextually rich conservative pain treatments). Some participants saw a danger that 'sham'-controlled RCTs may be unduly advised. They asked for it to be highlighted that this is not the only paradigm that can be recommended, and that efficacy or mechanistic trials can only be recommended for certain kinds of research questions. Quote: "*How to avoid that we take the whole field into something that is that is not helpful in the end?*"
Positioning nonpharmacological efficacy research in relation to drug trial standards	Discussions under this topic touched on the 'non-inertness' of control interventions in PPS trials, context-specific effect sizes, the common lack of provider blinding, and the impact of provider expectancies and problems with their assessment. Quote: "*You just can't assume that it's an inert control*."

	The proposed conclusion was to state in the manuscript that drug and non-drug trials are different, e.g., in the above characteristics, and to discuss expectancy measurement and the implications of a lack of provider blinding. However, it was agreed to refrain from a discussion of effect sizes and how these might compare between drug and non-drug RCTs due to a lack of empirical data (Note that this topic was later picked up on during peer-review and additional comments added to the manuscript).
(Lack of) provider blinding	Multiple arguments around the impact of a lack of provider blinding were made: Attendees highlighted the importance of ensuring providers can deliver both treatments with equipoise. It was recommended to address equipoise as part of the provider training for the trial and assess perceived equipoise of trial providers. Provider allegiance to tested interventions should also be pointed out as a potential risk to trial validity, and attendees made several suggestions on how to address this. There were disagreements regarding the use different sets of providers for each study arm: Some attendees pointed out that this may be beneficial due to variations in allegiance, others highlighted uncontrollable additional factors if different providers were used for each arm. It was concluded that these issues should be mentioned but no absolute statements made to not limit research creativity with our guidance.
Presentation of guidance	Meeting attendees considered how the consensus meetings' nuanced discussions could best be conveyed and what to do with items that did not reach consensus during the Delphi but might have relevance in certain scenarios. It was agreed that results from the Delphi survey should be published in detail separately or as supplement.

	There was agreement that a checklist of core recommendations and an infographic were desirable. It was considered whether a summary of stakeholder interviews should be provided as a supplement or in the manuscript. To highlight the importance if stakeholder involvement, the former was agreed. Further, it was agreed that findings from patient partner interviews could not be framed as 'recommendations' but should be presented for consideration only. (Note that the findings from stakeholder interviews were later moved into the E&E document)
Relation of developed guidance to the TIDieR-Placebo reporting guideline	It was concluded that the TIDieR-Placebo reporting guideline (Howick et al., 2020) was a good starting point but that more detailed reporting was required regarding the content and reasoning of control interventions. We therefore agreed to recommend general compliance with the TIDieR-Placebo checklist but put forward additional items in our manuscript. These items were also informed by the preparatory systematic review and meta-analysis (Hohenschurz-Schmidt et al., 2023b, 2022b).
Real-world effectiveness and consideration of additional control groups (no-treatment / waitlist controls)	It was argued that adding a no-treatment arm can move a trial further towards the pragmatic end of the trial design spectrum (Thorpe et al., 2009; Loudon et al., 2015a), where conclusions about real-world effectiveness become easier. However, some attendees felt that even with a no-treatment arm, the implications of a non-significant between-group comparison of specifically designed control intervention vs. tested treatment are not any clearer. Another attendee retorted that a no-treatment arm in addition to a control intervention would help do away with the undue notion that a 'placebo' has no effect. Attendees discussed examples from post-traumatic stress disorder (PTSD) trials and the case of acupuncture in NICE guidelines. The latter was deemed to illustrate the implications of incorrect conceptualisation of intervention mechanisms. Attendees highlighted the need to be clear

	that our recommendations are about mechanism and efficacy, not necessarily effectiveness. The following caveat was also agreed: If trials are set up to study efficacy and even mechanism, no conclusions about effectiveness can be made based on within-group changes without a no-treatment control. It should be stressed that conclusions need to be aligned with the research question.
Outcome measures	After some controversy about the susceptibility of 'objective' versus 'subjective / patient-reported' outcome measures to context-dependent effects, the discussion was reconciled as follows: We want to be sure not to disregard patient-reported outcome measures, as these represent the lived reality of trial participants and patients. Also, 'objective' measures are sometimes not available or desirable. Advocating for certain types of outcome measures in our guidance could amount to recommendations about what research questions people should ask, and that is not the purpose of guidance like ours. Quote: "…*add a sentence or two about the fact that the outcomes have to be linked to the hypothesis, but not going too much into it, because otherwise, as you say, we're going to dictate what type of research people can do and what questions they can ask*."
Unclear treatment mechanism(s)	This topic was summarised and agreed as follows. Quote: "*What do you do if your if your treatment mechanism is pretty much in the dark and unclear? That's a common scenario, so we need to anticipate that in in some way. Where we ended up with those discussions was to raise that [issue of unclear mechanisms], but to defer to the to the individual investigators to clarify what it is that they are looking at. Each therapy will at least have some theories about the supposed treatment mechanism. If that's correct or not is exactly the question that might be investigated with a trial like this, but there are ways of informing that either from previous mechanistic studies or*

	through the theory of the person who's developed the respective treatment – that's going back to our point earlier that you need to be clear on what you're testing and be careful in interpreting that (…). And then as a quality control or as a as a safety check, most researchers need to ensure that the control is inert for the mechanism that they are looking at. So, if you phrase it those two ways around, then I think you can give people something they can work with."
Complex interventions	To elucidate the effects and / or mechanisms of complex PPS interventions, attendees proposed the study of multiple components in one single trial or across dismantling / iterative trials. It was also noted that mediation analyses are common in the psychotherapy field. A participant cautioned not to "destroy" research by calling for unrealistically complicated trials. Others added that the scientific community may over-simplify the search for mechanism, with the identification of 'active' ingredients being difficult or wrong and their interactions often over-simplified or disregarded. Another participant abstracted the discussion to say that we should clarify early in the manuscript that the drug-trial placebo controlled RCT model cannot be seamlessly transferred to complex intervention research. They asked for the guideline to refer to the Medical Research Council framework for complex intervention development and testing (Skivington et al., 2021). Quote: "*There will be no comparably neat, clear design*." Another participant cautioned that the difficulties in designing an adequate ('sham') control intervention should not be used as an argument for uncontrolled studies or solely comparative effectiveness trials. Specifically designed control interventions control for rituals and extra-therapeutic factors, which other designs don't. Another attendee commented on the framing in the guidance manuscript: The aim is to

	help the research field, not to postulate an "ideal trial" and undermine the research even further. I summarised the sentiment of the discussion in that the final recommendations needed to be realistic and presented with humility.
Ethics	During the discussion of ethical considerations, controversial points included: - The ethics of perpetuating treatments with no evidence of efficacy - Pros and cons of offering the trial intervention after trial completion - Balancing the ethics of concealing a trial hypothesis from participants (i.e., the existence of a control intervention) with the ethical need for high-quality internally valid trials (i.e., ensuring blinding) The agreed conclusion was to refer to general ethical considerations of clinical research and specifically designed control interventions and placebos in trials, as well as to the ethics of 'deception'.
Blinding assessment and the assessment of participant expectations	On the importance of assessing the effectiveness of blinding, it was agreed that blinding is a main objective of control interventions in efficacy and mechanistic trials, and thus should be assessed. Attendees stated that such assessment is a question of proportion (of guesses between trial groups). Results should be reported with raw data to enable independent calculation of blinding indices. There was no agreement on the best timepoint of assessment; on one hand, early assessment was argued to potentially affect clinical outcomes and blinding, later assessment, on the other, might be confounded by experienced symptom changes. Assessment at the end was finally agreed to be the safest option, but there was a perceived need for more research on the influence or feasibility of blinding assessment at different timepoints throughout longer trials.

	It was further agreed that patient expectation and satisfaction should not be used as proxies for blinding. However, expectations were still seen as potentially important. Some attendees thus advocated for expectation assessment, also to enable stratification approaches for subsequent effectiveness questions.
Input form patient interviews and issues of stakeholder involvement	A patient representative stated that a patient's main concern is accessing and understanding the trial results. There were also discussions around offering treatment at the end of a trial. Quote: "*There's a giving a report to the patient, which should be mandatory of sending our papers out with a thank you note or something. And then there's the other thing of giving the treatment to everybody. I think that's mostly in our interest because, as was said, it's very motivational for the patients to come into the study. If they're sure that they're going to get the intervention, that's why they're joining the trial. But it might not always be possible for different reasons, and so I think that should be an option and something that is discussed*." Further, there were calls for stakeholder representation when developing the consent process. It was agreed to strengthen the recommendations on stakeholder involvement in the piloting and pre-trial testing of control interventions. The manuscript was also seen to need more information about participants' potential fear of adverse events from control interventions and the potential benefit of a 'reality-check' from patients. These issues were to be framed as considerations for the consent process and pilot work, not as formal recommendations (as not assessed during the Delphi, and potentially very situation-specific).

Analysis of patient partner interviews

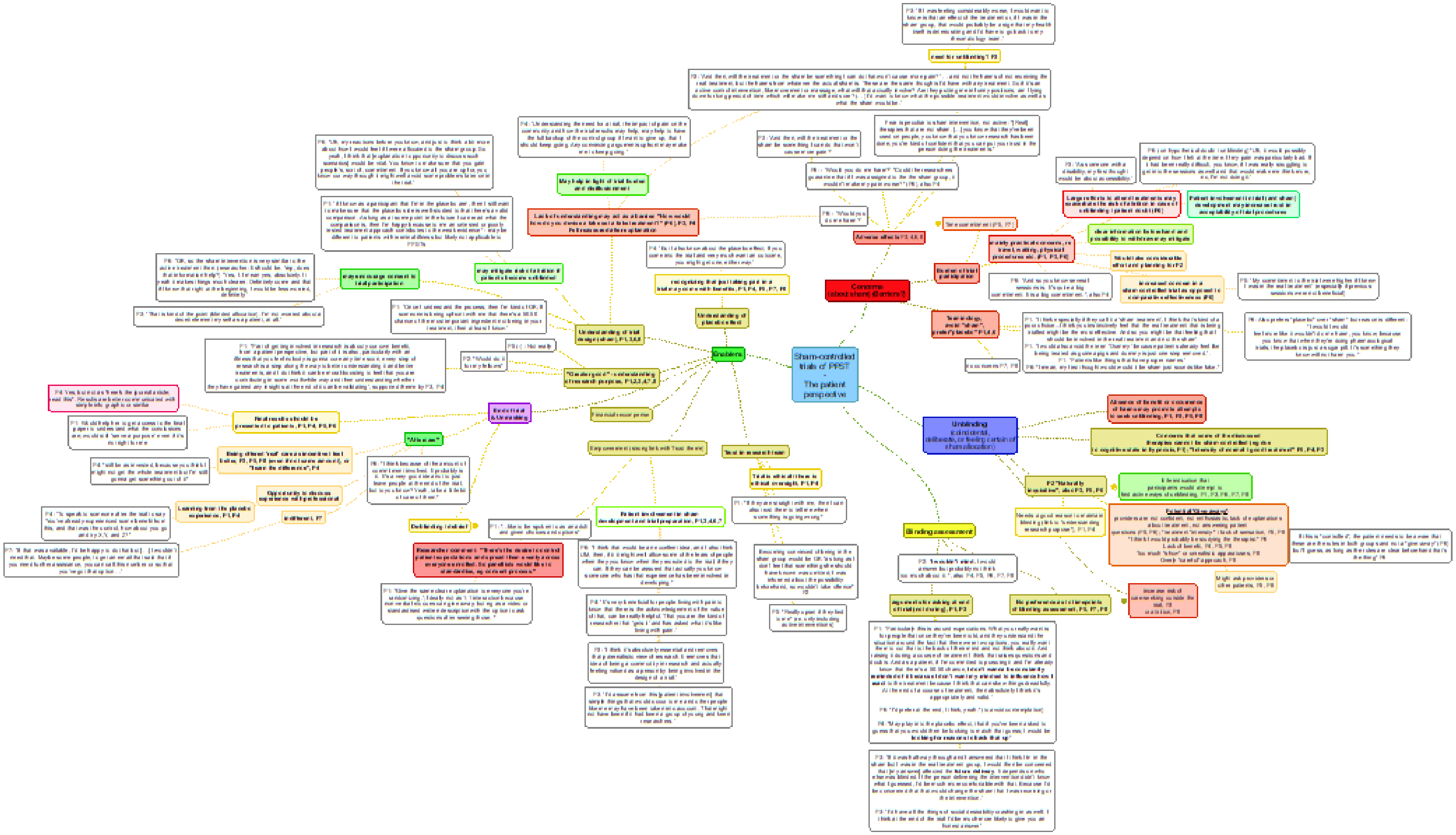


Figure 15: Analysis map of patient partner interviews

Barriers and enablers to trial participation for patient partners

Table 38: The lived-experience perspective on potential barriers and enablers for participation in efficacy or mechanistic trials of PPS interventions with a specifically designed control intervention. Content in the left-hand column is based on interviews with eight people with various pain-related experiences, during which hypothetical trial scenarios were discussed. The right-hand column details potential solutions and ways of incorporating the identified themes into control intervention development and implementation; these were developed during interviews and refined during consensus meetings.

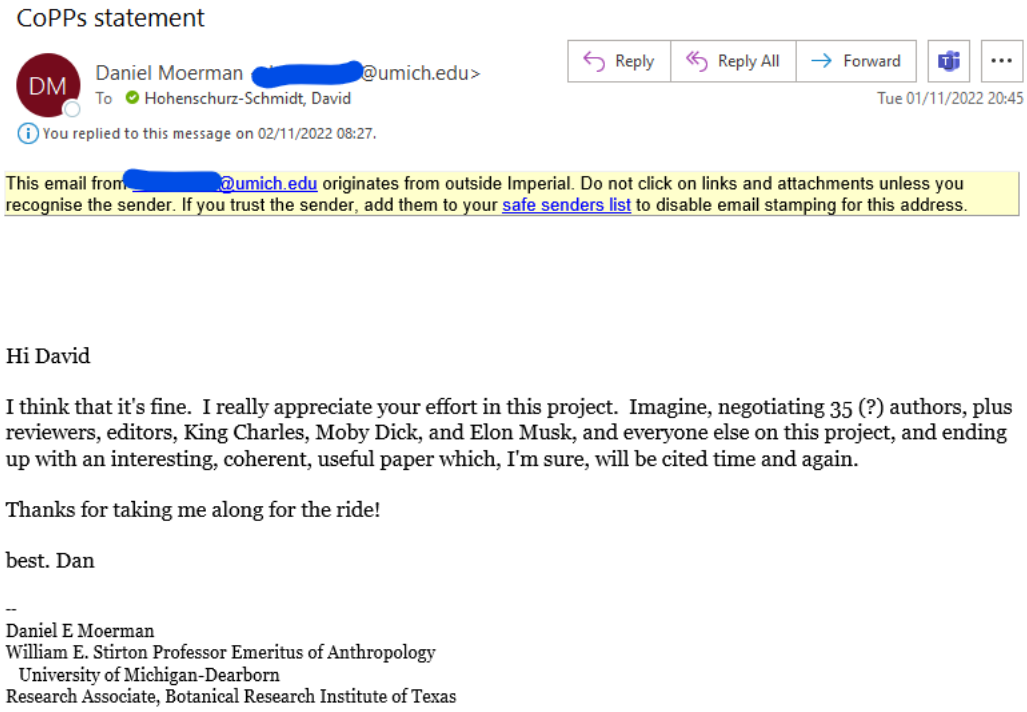
Theme and explanation (Based on interviews)	**Potential incorporation in trial development and conduct** (Based on interviews and consensus meetings)
Barriers	
Fear of adverse effects Some people were worried about being assigned to a 'sham' or otherwise not commonly used control intervention, fearing it might aggravate their symptoms. Others were equally apprehensive about assignment to any trial arm for the same reason. *"Could the researchers guarantee that if I was assigned to the sham group, it wouldn't make my pain worse?"*	Fear of adverse effects can be partly mitigated by involving other people with experience of the studied condition in the development (and testing) of the control intervention and communicating this involvement to potential trial participants. Fear can also be addressed in the informed consent process, clearly communicating the purpose of the control intervention and its resemblance to the test intervention. Communicating about trial interventions should not, however, unbalance participant expectations of benefit or harm.
Burden of trial participation	

People living with long-term or very disabling pain noted mainly practical concerns; for others, the time commitment of participating in a trial was seen as potentially burdensome. The possibility of receiving a "fake" treatment exacerbated this concern.	Concerns about the burden of trial participation may possibly be mitigated by patient involvement during development of the intervention.
Terminology: Avoid 'sham', prefer 'placebo' "*I mean, my first thought would be… sham just sounds like fake*."	Implications of the interviewees' preference for the term "placebo control" were discussed during consensus meetings and had to be reconciled with and the research experts' concerns regarding the lay understanding of the term placebo. Eventually, it was agreed to recommend describing the control intervention as neither 'sham' nor 'placebo' but instead in relation to the tested treatment. This approach was hoped to improve understanding and mitigate all parties' concerns. For more detail, see the Terminology and Communication section (Section 1) of the final guideline and its E&E document.
Lack of understanding of control interventions Understanding of the concept of control treatments was largely vague, contributing to concerns about adverse events and the burden of taking part in a trial. Explaining the concept was reassuring. "*How would you devise a fake treatment*?"	Education about control interventions and the purpose of controlled trials can be included in the informed consent process.
Enablers	
Understanding of the placebo effect	

Many participants felt that just taking part in a trial may come with benefits, often explicitly attributed to the placebo effect. "*But I also know about the placebo effect; If you come into the trial and very much want an outcome, you might get one, either way*."	The placebo effect could be discussed during the informed consent process. Balancing of expectations should again be the aim, emphasising that both groups will experience comparable placebo effects.
Understanding of trial design and control principles A good understanding of the principles of clinical trial research and the value of control groups was deemed desirable by participants. "*If I know as a participant that I'm in the placebo arm, then I still want to make sure that the placebo side is well-studied to that there's a valid comparison*."	An improved understanding of trial design and control principles may mitigate the risk of attrition in cases of unblinding or doubt and may facilitate enrolment. As above, education about control interventions and the purpose of controlled trials can be included in the informed consent process.
"Greater good" – Understanding of research purpose Most participants reported that they saw a higher purpose in contributing to research; some even felt this was their main motivation to participate. *"I would do it for my fellows."* "*I do think it can be morale-boosting to feel that you are contributing in some worthwhile way and then understanding whether [the researchers]*	The sense of purpose in taking part in research could be enhanced and used as a motivational factor when informing potential participants about a trial. The participants' commitment should be respected and validated at the end of a trial (see description of 'aftercare' below).

have gained any insights at the end of it can be validating."	
Trust in the research team Interviewees hoped for a trustworthy collaboration between trial participants and the research team. "*Becoming convinced of being in the sham group would be OK as long as I don't feel that something she should have known was omitted; I was informed about the possibility beforehand, so wouldn't take offence.*"	The distinction between blinding and lying may have to be clarified. The ethical justification for the trial and trial oversight arrangements may be useful to communicate. Unblinding at the end of a trial can be offered.
Patient involvement in control development and trial preparation Knowing that patient representatives were involved in the development of a trial may mitigate many of the concerns presented above. "*I'd assume from this [patient involvement] that simple things that would occur to me and other people like me may have been taken into account. That might not have been if it had been a group of young and keen researchers*." "*It's very beneficial for people living with pain to know that there is the acknowledgement of the value of that [involving patients]. That you are the kind of researcher that 'gets it' and has asked what it's like living with pain*."	People with lived experience can anticipate many barriers to the successful conduct of a trial, including concerns peculiar to a controlled (as opposed to a comparative effectiveness) trial. Involvement should commence early in the trial planning stage and involve people from the respective target population. From a trialist's perspective, this also offers the opportunity to anticipate and test potential 'give-aways' for unblinding in the control arm.

Panellist feedback



CoPPs statement

DM Daniel Moerman [redacted]@umich.edu>

To Hohenschurz-Schmidt, David

Reply | Reply All | Forward | Tue 01/11/2022 20:45

You replied to this message on 02/11/2022 08:27.

This email from [redacted]@umich.edu originates from outside Imperial. Do not click on links and attachments unless you recognise the sender. If you trust the sender, add them to your safe senders list to disable email stamping for this address.

Hi David

I think that it's fine. I really appreciate your effort in this project. Imagine, negotiating 35 (?) authors, plus reviewers, editors, King Charles, Moby Dick, and Elon Musk, and everyone else on this project, and ending up with an interesting, coherent, useful paper which, I'm sure, will be cited time and again.

Thanks for taking me along for the ride!

best. Dan

--
Daniel E Moerman
William E. Stirton Professor Emeritus of Anthropology
University of Michigan-Dearborn
Research Associate, Botanical Research Institute of Texas

Figure 16: Feedback from a Delphi panellist and CoPPS co-author

Questions to consider before using blinding

Table 39: Questions to consider before using blinding. The left column is adapted from (Anand, et al., 2020) who in turn adapted from (Miller, 2004); the right column is the thesis author's own contribution.

Questions to consider before using blinding in any clinical trial (according to (Anand, et al., 2020))	Considerations and applicability to trials of Physical, Psychological, and Self-Management interventions for pain
• Is blinding needed for a scientifically sound result? (Will the intervention have a placebo effect which needs to be separated from its true effect?)	PPS interventions are contextually rich, making the placebo effect a likely contributor to clinical effectiveness (Bialosky et al., 2017b; Miciak et al., 2012; Rossettini et al., 2020). While this is supported by mechanistic studies around communication and other contextual factors (Colloca, 2019b; Rossettini et al., 2020), direct clinical or meta-analytical studies regarding the proportion of placebo effects in PPS interventions, or their interactions with 'specific' therapy components, are lacking. Evidence from related fields suggests a moderate effect size (Sham shockwave therapy for heel pain (Gerdesmeyer et al., 2017); Psychological placebos for depression (Furukawa et al., 2014)). Under these premises, the matching of contextual factors and intervention components serves the purpose of balancing expectancy and learning effects between groups, which we have shown to influence outcomes irrespective of blinding (Hohenschurz-Schmidt et al., 2023b). Under this topic, one should also consider the option of employing an open-label placebo (OLP) control, i.e., a supposedly inert intervention of which participants are being made aware (unblinded allocation). Before administering an OLP, participants are usually educated about the placebo effect and the possibility of experiencing symptom reductions from inert treatments; OLPs have shown moderate-to-large effects across several conditions (Carvalho et al., 2016; von Wernsdorff et al., 2021). We have criticised a recent prominent OLP trial of 'pain reprocessing therapy', but largely on the grounds of the large structural differences between OLP and tested intervention (comparing a one-off open-label saline injection with high-intensity psychotherapy and education) (Ashar et al., 2021; Ashar and Wager, 2022; Hohenschurz-Schmidt et al., 2022a). Whether placebo effects can be adequately controlled for in RCTs with structurally largely equivalent OLP control interventions remains to be tested.
• How likely is it that patients or clinicians will behave differently if they know the intervention and would this change in behaviour bias the results?	This item requires further elaboration: First, 'Changes in behaviour' could mean the seeking of concomitant treatments or dropping out of the study. Both events can occur due to 'resentful demoralization' which occurs when one study treatment is systematically perceived as less desirable. In PPS interventions, common active comparators are 'treatment as usual' or structurally different attention controls, both of which likely less desirable for patients, so that blinding may guard

	against bias. The use of another active specific treatment as comparator, likely less demoralising, should not primarily be decided based on the appeal of blinding to trial designers, but on the research question (Keefe et al., 2022).
• Are the potential harms to patients of using blinding excessive?	Here, we need to distinguish between direct adverse effects from control interventions, and indirect harms such as the withholding of an effective treatment alternative. Direct adverse effects from PPS control interventions are likely rarely serious – as they are in the tested interventions (Carnes et al., 2010; Niemeijer et al., 2020). As for indirect harms, again, context matters: If the test intervention has been shown to be efficacious in repeated well-conducted efficacy trial, a blinded comparison against a specifically designed control that omits efficacious intervention components is not ethical (Emanuel and Miller, 2001; Millum and Grady, 2013). To date, however, this is not a common scenario in PPS clinical research. Instead, efficacy trials are skipped in the intervention development cycle in favour of comparative effectiveness trials against other low-evidence interventions or ill-defined 'usual care' (for an example see CBT for headaches (Williams et al., 2020, 2021)). Given the potential scientific insights, outweighing the low risk from control intervention in PPS, controlled efficacy trials are ethical in many areas of PPS (Miller and Crouch, 2008; Annoni, 2018). A further harm – interference with a more efficacious mainstay medical treatment - is also uncommon with blinded efficacy trials of PPS. Usually, these trials add test or control interventions to patients' usual care, and do not targeted patient populations that suffer from acute or rapidly deteriorating medical conditions for which effective medical treatments are available.
• Does the anticipated social value of the study results justify any potential harms of blinding?	More broadly than above, harm can be conceptualised to include wasting a patient's time or public resources through enrolment in the control arm of efficacy trials. The counterarguments are, of course, that large pragmatic trials are more expensive to conduct than smaller well-conducted efficacy trials, and that unproven interventions waste more time and resources in clinical practice. Contrary to research settings, in clinical practice this may come at the expense of otherwise effective treatments (including the use of 'complementary medicine' to cure life-threatening disease (Johnson et al., 2018)). Therefore, in the case of lacking high-quality efficacy data for a given PPS intervention and assuming a low risk profile of the control intervention, blinded
• Does the financial cost of blinding compromise spending on other methodological aspects of trial integrity?	This question is again rather unspecific. Ideally, RCTs are costed for in a manner that the research question can be answered in a scientifically rigorous manner. Even in surgery, where blinding procedures are indeed very laborious and costly, this scientific-financial-ethical problem is increasingly answered in favour of blinded sham-controlled efficacy trials. Control interventions in PPS efficacy trials are comparatively

	inexpensive, also recognising that budgets may be smaller. Binding is, in most trials, not the costliest element but can be fundamental to obtaining valid results. On the other end of the explanatory-pragmatic continuum, a major criticism of comparative effectiveness research is the inability to ascertain whether the comparator intervention did produce the desired effect in a given trial; which is why the US Food and Drug Administration recommends a third no-treatment arm for non-inferiority trials (Food and Drug Administration, 2016). This 'methodological aspects of trial integrity' is almost always compromised on, likely due to the cost involved.

Appendix 3: IMMPACT Statement, additional content

Table 40: Alternative designs to parallel-group RCTs, potentially suitable for pragmatic research questions. Based on (Hohenschurz-Schmidt et al., 2023a).

Alternative Design	Description
Large, simple trials (LSTs)	Large, parallel-group trials with minimal data collection requirements, broad eligibility criteria, and objective, often routinely collected outcome measures. LSTs are cost-efficient, maintain scientific rigor, and interfere little with clinical practice. They provide high-quality information even for rare outcomes and could provide valuable insights into drug interactions (Reynolds et al., 2011; Roehr, 2013).
Cross-over design	Cross-over designs may be used for short-acting, non-disease modifying drugs to answer pragmatic research questions. They remove between-patient variance, reducing sample size requirements, and can assess the effects of switching between treatments. Individual or multiperiod cross-over designs are adaptations geared towards clinical decision-making.
Stepped wedge cluster randomized design	A variant of a cluster randomised trial. Starting with a non-exposure period for all study clusters, this design involves "random and sequential crossover of clusters from control to intervention until all clusters are exposed." (Hemming et al., 2015). The phased baseline period acts as the internal control condition and all clusters contribute to both study conditions. Stepped wedge trials provide more statistical power than parallel cluster designs when clusters are heterogeneous and/or large.
Enrichment designs	Randomising only patients with a supposed increased likelihood of treatment success or other characteristics. Targeted pre-identification during eligibility screening, including the use of biomarkers, clinical diagnostics, and enrichment phases can be pragmatic if these methods address a clinically relevant question and are feasible in routine clinical practice. Although this process allows for pragmatic research questions, it certainly reduces generalisability by excluding a certain number of potential treatment recipients.
Adaptive and other designs responsive to accumulating trial information	Adaptive designs enable researchers to respond to interim safety and efficacy data. They modify the trial's course in accordance with pre-specified rules (Pallmann et al., 2018). Adaptive designs are challenging and have multiple limitations, but can be efficient for comparing multiple treatments and doses.

Appendix 4: NeuOst intervention, additional content

Evidence overview of intervention components considered for *NeuOst* intervention

The below table provides an overview of potential intervention components, amenable to osteopathic practice, and with the potential to be combined in *NeuOst*. At the intervention development stage, this table served as the basis for discussion amongst the research team and during stakeholder engagements. The final table was then used to develop the intervention protocol for the *NeuOst* feasibility RCT.

Intervention elements can be split into manual and non-manual groups and pre-defined 'techniques' in each group (following the UK Back pain Exercise And Manipulation (UK BEAM) trial approach (Harvey et al., 2003)). ACT-informed techniques were based on the unpublished 'ACT OPEN Workbook', used to accompany online ACT treatment delivered for people with HIV-related neuropathy (Scott et al., 2021) and similar to the workbook used in Kioskli et al.'s (2020) study with people with pDPN.

Table 41: Intervention components of NeuOst. Manual therapy components and techniques based on Harvey et al. (2003) classification, with 'subtle' manual therapy components added.

Intervention Type	Intervention Components	Treatment targets / Mechanisms relevant to pDPN	Evidence of Efficacy and / or Effectiveness *(population, study design, comparator, n total, main findings)*	Potential role in Multimodal Intervention	Expected interaction with other components
Manual therapy-based components					
Soft tissue components	Massage of painful and non-symptomatic areas; deep vs superficial; painful vs pleasant techniques	counter-pain mechanisms; overall / ANS relaxation; muscle relaxation; expectation effects	A review that included (six) manual therapy trials found Thai foot massage, Gua Sha, reflexology and aromatherapy to be more effective on pain and function parameters	List of eligible techniques that can be applied freely depending on	Forming the basis of 'treatment' during in-person sessions; Hoped to facilitate engagement in more active self-

	(Techniques: cross-fibre stretch, longitudinal stretch, direct pressure, deep friction)		than education controls, TAU, or no treatment. Also included trials on CTS in people with DM (Hernández-Secorún et al., 2021). Largest trial apparently Chatchawan et al. (2015) on six sessions of Thai foot massage vs educational control (total n = 60), finding benefits in functional and sensation tests. Looking at primary pain-outcomes and >3-week follow-up in neuropathies, Moisset et al. (2020) identified no eligible studies from most non-drug fields, including physiotherapy, therapeutic patient education, Tai Chi or Chinese acupuncture. For spinal manipulation for peripheral (poly-)neuropathies, no human studies could be found. (See Song et al., (2016); Onifer et al., (2018) for rodent studies) Here you can cite the literature on immediate neurophys effects of spinal mobilisation, see also our (now rather old Systematic review in manual therapy 2008). Mostly short term effects though, but pain modulatory plus sympathetic NS.	provider choice and patient preferences.	management components through short-term relief. Arguably enhancing the therapeutic relationship / alliance through touch (McGuirk, 2012; Kerr et al., 2019).
Articulatory / mobilisation components	(Techniques: low-through high-amplitude passive movements of lumbar spine and sacroiliac joints (and necessarily hips); flexion, extension, rotation, side-bending, manual traction; oscillation)				
'Subtle' Manual Therapy components	(Myofascial, craniosacral, and visceral techniques) Gentle hands-on approaches, usually based on structural models of health and disease or palpation of tissue tension.	ANS relaxation; muscle relaxation; expectation effects			
Thrust components ('manipulations')	(Techniques: high or low velocity; low amplitude; direct or leverage; directed at central lumbar, zygapophysial or sacroiliac joints; unilateral or bilateral; at one or more locations) Thrust force; Cavitation; joint loading and tension; sensory experience	muscle relaxation; descending pain inhibition; temporarily enhanced joint mobility; expectation effects			
Soft tissue component 2 (Neurodynamic techniques)	Manual techniques that aim at mobilising a nerve in its peripheral course.	Usually applied to suspected entrapment neuropathies. Role in	Single trial on "nerve mobilisation" (Singh et al., 2012): n = 30, DM2, 21 days tibial nerve stretch (Dorsiflexion	Introduced during in-person treatment sessions and provided	Potentially reinforcing benefits of manual therapy; Linking in with activity-increasing aspect

	Can also be provided as home-programme.	pDPN to be explored, but preclinical study suggesting mitigation of axon degeneration (Zhu et al., 2022). Techniques for sciatic and distal nerves likely most relevant.	+ Eversion + SLR, 10 mins per day) compared to no treatment. Found sign. b/w group differences in changes in vibration thresholds. High RoB and 4/11 PEDro score according to (Hernández-Secorún et al., 2021). In chemotherapy-induced neuropathic pain, a recent exploratory RCT suggests usefulness of upper extremity home-based programme (daily 5-10 minutes starting prior to chemotherapy) (Andersen Hammond et al., 2020)	as part of home-exercise programme	of home-exercise programme.
Non-manual therapy-based components					
Physical activity / Exercise components					
Exercise, Sensorimotor / balance advice and prescription	Advice and provision of exercise plans, including detail on form, sets, and repetitions, frequency, and individualisation. Type of exercises: **Sensorimotor / balance training**; regimen provided as training plan (see Appendix in Ahmad et al. [2020])	Sensorimotor / balance training: Supposed altered motor programming and movement patterning through repeated proprioceptive and somatosensory input (Ahmad et al., 2020). Fall prevention and avoidance of pressure areas on feet (risk reduction for ulcerations) (Hernández-Secorún et al., 2021).	Sensorimotor / balance training: Hernández-Secorún et al. (2021) reviewed 12 moderate-to-high-quality trials of various balance programmes in DPN, one with vibration elements. They report largely significant benefits on various proprioceptive, some neurophysiological and EMG, and few functional or disability outcomes, compared to attention or TAU controls. E.g., some indication from small attention-controlled RCT (n = 38) that proprioception, EMG, and potentially nerve conduction velocity may be improved (Ahmad et al., 2020).	Could be either provided to those people who express an interest in a home-based exercise plan, or as generic take-home programme. Equipment would have to be provided or improvised or exercises adapted; Initial practice within clinic session; Safety concerns?	May interact with below advice on general activity, for example by setting the basis for starting some form of physical activity or allowing for movement even during very painful episodes, or by improving proprioception and making daily activities easier. Also interact with mindfulness and pain education components, trying to reframe movement as beneficial and non-threatening. Interaction

			Ankle-reaching task + virtual obstacle crossing (2x per week, 4 weeks; n = 39, elderly) showed some effects on balance over no-treatment control (Grewal et al., 2015). Note on **dosage**: All studies have session schedules of at least 40-60mins once per week, usually more frequent.		with manual components: providing additional sensory input.
Exercise (Physical, Stretching)	Static and / or dynamic lengthening of muscle groups of varying duration and sequences.		No studies found that examined solely stretch exercises, but stretches were included in most regimes reviewed by (Hernández-Secorún et al., 2021)	Can be provided at very low intensity and may thus be suitable for people with high levels of pain or high kinesiophobia as introductory exercises. Practiced during in-person sessions and provided as "homework".	May be facilitated through short-term relief from manual therapy and enabled through cognitive changes from CBT-based and pain education components (incl. goal setting techniques). Could be combined with elements of mindfulness or controlled breathing.
Exercise (Physical, anaerobic / Strength)	Strengthening exercises with body-weight, resistance bands, or weights. Variable repetitions; load parameters; addressed muscle groups; May include social components.	Autonomic; muscle strength; coordination; proprioception; endorphins	Reviewing 4 trials in DPN, all of which deemed strength training effective in balance and strength variables compared with no-treatment control groups (Hernández-Secorún et al., 2021).	Provided depending on patient preference and disability levels; Potentially practiced during in-person sessions and provided as "homework";	May be facilitated through short-term relief from manual therapy and enabled through cognitive changes from CBT-based and pain education components (incl. goal setting techniques).
Exercise (Physical, aerobic)	Cardiovascular exercises from fast walking to running / cycling / swimming. Variable in duration; intensity; social; whole-body vs legs only	Cardiovascular and autonomic; proprioception; endorphins; (Diabetes-related mechanisms?); Weight loss	Hernández-Secorún et al. (2021) reviewing 7 aerobic-type training programmes, finding most of them to be effective than no-treatment controls for function, pain, and balance in DPN. Positive findings for both HIIT and moderate intensity	Provided depending on patient preference, access to equipment and training facilities, and disability levels.	Interactions with dietary components; May be facilitated through short-term relief from manual therapy and enabled through cognitive changes from CBT-based and pain education

			training. Also, treadmill compared favourably to balance training. No meta-analysis and primary studies not always great.		components (incl. goal setting techniques and value-driven action).
Breathing exercises	OR breathing re-training. Controlled diaphragmatic breathing, usually with specific duration of in- and outbreaths.	Stress-reduction (ANS); prefrontal control; Control over physical reaction to pain.		Potentially practiced during in-person sessions and provided as "homework"	Can be incorporated in all physical exercises and forms part of many mindfulness techniques; May increase ability to engage in movement through altered attention focus (from pain experience to breathing)
Communication, Education, and ACT-informed components					
Pain self-management advice	Advice on rest, ice, heat, over-the-counter medication, and physical activity.	Pharmacological analgesia during flare-ups; Local and central effects on nociceptive processing.		Providers have access to list of best-practice advice for MSK and neuropathic pain flare-ups. Can be provided verbally and / or via an information sheet.	May complement intervention sessions by easily applicable, active, and healthy coping strategies. Basis of more specific exercise / regular engagement in activities. Links with pain education to understand value of active coping strategies and rationale for pharmacological pain relief in facilitating movement and (possibly) preventing chronification.
General physical activity guidance	Recommendations on starting or modifying physical activities. May include daily walking, taking stairs, interrupt sedentary periods, or	- slow muscle atrophy - maintain gait mechanics - reduce risk of falls	Activity promotion studied often alongside dietary intervention in diabetes risk prevention (Balk et al., 2015). Recent 3-armed RCT found little benefit on metabolic function from pedometer-based	Aim could be to build on patient's preferences and existing habits. Needs to be individualised; "Advice" would be	Hoped to be enhanced through osteopaths' training in CBT-informed techniques.

	exercise programmes such as gym, yoga, swimming, cycling, speed walking running, etc.; Could consider use of accelometre / pedometric app. Technique: Can be delivered in 5A's manner (Estabrooks et al., 2003), a structured communication approach to evaluate and promote physical activity in primary care.	- cardiovascular benefits and enhanced glycaemic control Statement of the American Diabetes Accoc.: "*Exercise improves blood glucose control in type 2 diabetes, reduces cardiovascular risk factors, contributes to weight loss, and improves well-being. Regular exercise may prevent or delay type 2 diabetes development. Regular exercise also has considerable health benefits for people with type 1 diabetes (e.g., improved cardiovascular fitness, muscle strength, insulin sensitivity, etc.).*" Should be tailored to individual needs (Colberg et al., 2016)	activity promotion, with or without additional group support for behaviour change and diabetes specialist 1:1 support, and any effects waning after intervention termination (Rossen et al., 2021). Reviews conclude that pedometers increase activity levels, with less evidence to suggest metabolic improvements (Baskerville et al., 2017; Franssen et al., 2020), and some indication that combination with in-person consultation may add benefit (Hodkinson et al., 2021).	provided using below CBT-informed communication strategies.	
Mindfulness-based Stress Reduction, MBSR	Meditation, Rest, Breathing …	Stress-reduction (ANS); prefrontal control; empowerment	High-intensity 8-week group programme improved function, better health-related quality of life, and reduced pain intensity, pain catastrophizing, and	Self-management component; Instructed through provider / other resources	

			depression compared to those receiving usual care (individual RCT, pDPN, n = 66) (Nathan et al., 2017).		
Education (Pain Education)	Learning about pain mechanisms; experiential and individualised components (incl. reflection on own experience); Neuropathic pain-related information	Reducing kinesiophobia and catastrophizing. Possibly reframing threat value of pain.	Systematic review finding improvements in kinesophobia and catastrophising but not pain intensity or quality of life (Watson et al., 2019). Also see (Ashar et al., 2021) regarding these cognitive mediators.	As part of provider interaction at outset (to set the scene for improved pain self-management); As (optional) component provided through external resources (online)	Possibly allowing for better engagement in physical activity and / or manual therapy through reductions in negative pain connotations.
Sleep Management (Education)	Learning about best-practice strategies to improve sleep	Addressing one of the most salient QoL disturbances (as identified during PPI and congruent with literature (Ferini-Strambi, 2017; Naranjo et al., 2019))		Provided as simple hand-out and discussed with provider	Possibly buffering the impact of NeuP on sleep and addressing relevant effect of symptom; thereby making intervention more logical and hopefully effective.
Acceptance and Commitment Therapy (ACT) – informed components	Techniques / exercises (from Scott & McCracken, n.d., ACT OPEN workbook), selection: 1. Focus on your breathing (Mindfulness) 2. Passenger on the bus Metaphor and reflective questions (Shifting attentional focus) 3. Notice five things (Awareness)	Mechanisms: Learning to tolerate distressing experience and to learn self-acceptance (Lau and McMain, 2005) More control over emotional aspect of pain experience	Moisset et al. (2020) included five open-label RCTs of psychological studies in their review on pain in (any) peripheral neuropathy, concluding: There is a weak recommendation for the use of CBT and mindfulness as an add-on therapy for neuropathic pain (moderate final quality of evidence). Another review (Racaru et al., 2021) included nine RCTs with both TAU and attention control comparators. Meta-analyses of 2-4 studies each showed	Several: 1. Combined mindfulness-based exercises during manual treatment 2. Dedicated mindfulness exercises during treatment session, such as body scans etc. 3. Homework-based practice possible for all (if programme inflexible) / some patients (if individualised)	Promoting an 'aware mode' characterised by a freedom of choice (Lau and McMain, 2005), possibly allowing for increased activity and engagement despite pain, fear of pain, or associated negative thoughts. May interact with pain and psychoeducation, "about [pain]-related thoughts and symptoms

	4. Unwanted party guest Metaphor and reflective questions (Openness) 5. Observe-Breathe-Open Up Meditation (combination of skills) 6. Walking Meditation		significant effects of interventions over controls for short- and medium-term pain intensity, pain interference, and depressive symptoms, and for short-term quality of life.	In training programme: Cultivate accepting and mindful provider state of mind (provider practice?)	can facilitate earlier detection of these experiences, thereby increasing the chance of timely interventions like a previously created relapse-prevention plan." (Lau and McMain, 2005), p. 865
Relaxation (PMR)	Guided sequence of contraction and relaxation of subsequent muscle groups.	Stress-reduction (ANS)	(Izgu et al., 2020) 3-arm parallel-group trial of 12-week self-managed PMR vs mindfulness vs single-session educational attention control (T2DM with NeuP, n total = 77), finding sign. b/w-group differences in favour of PMR and MM over control for pain intensity but not QoL.	Step-by-step protocol provided as audio-file onto patient's phone (see Izgu et al. supplement). Could be standard component or according to patient preference.	
Mindfulness meditation	Guided mindfulness tasks involving sustained awareness of bodily sensations, breathing, etc.	Stress-reduction (ANS); prefrontal control;	(Izgu et al., 2020), see above. Evidence synthesis of *brief* mindfulness-based interventions on pain-related outcomes is inconclusive (McClintock et al., 2019). Longer programmes seem not to have consistent effects on pain, QoL, or mental health in acute or chronic pain (Ball et al., 2017); Effects on physical function in chronic pain are questionable (Jackson et al., 2019). No evidence of effects from n=20 pilot by Teixeira (2010): After 4-week pre-recorded meditation programme compared to nutrition diary for neuropathic		Mindfulness elements could also be integrated into manual and exercise-based intervention components. Refocus of attention from pain experience to, eg, breathing may facilitate engagement in physical activity despite pain. Supported by positive effects of pain education and CBT / goal setting.

			pain, NP-related quality of life, and sleep quality.		
Diabetes management components					
Diabetes complication screening and management	Providers trained to routinely screen for diabetes complications and doing so at regular intervals. Providers trained to recognise hypo and deliver hypo treatment in emergency situations	Prevention of complications and / or aggravation.	Part of NICE guideline for Type 2 Diabetes (NICE, 2015). Further guidelines exist for individual complications.	Incorporated into baseline assessments and regular check-ups.	Hoped to enhance therapeutic relationship through professionalism and reduce risk of disease complications alongside more active components. Comply with best-practice guidelines; role as first contact practitioner; Raising patients' awareness to signs and symptoms of potential complications (linked to education components).
Education (Diabetes and its complications)	Providers trained in pathophysiological knowledge and clinical presentations;		Systematic review of 42 RCTs finding that self-management education reduces all-cause mortality risk compared to usual care (He et al., 2017)	Could be short checking in with patient regarding their disease understanding and addressing of misconceptions if present	Basic knowledge to facilitate communication and management, and enhance therapeutic relationship.
Dietary components					
Dietary (Advice or Planning)	Intervention could range from simple encouragement to change eating habits, to online education session or peer-support group, to individualised dietary plans.			Basic checking of dietary habits and existing management strategies; qualified comment or further advice only if needed. Could also be (optional) online or group-based component.	Similar to above, but could play larger role and be easier / more relevant to deliver

Integration with standard medical care					
Integration with pharmacological and other medical DM management	Acknowledgement of medication list; Provider awareness of common side effects and prescribing guidelines. Communication with prescribing clinician if side effects suspected and desired by patient. Providers aware of standard medical management, incl. medications and side effects. Providers know the standard clinical steps / stages of patient management and can communicate with respective specialists.	Potential removal of side-effect-producing medication.		Part of initial consultation. Need to acknowledge professional boundaries and competence of osteopaths. Medication changes could be considered as exploratory outcome measure in feasibility study, will have to be monitored as potential confounder / mechanisms during definite trials.	Hoped to enhance patients' trust in provider competency as well as ability to communicate with other medical professionals if needed. Knowledge of common medications and their potential side-effects relates to education and choice of self-management advice; Pain relief from other sources maybe allowing for reduction in medication.
Integration with pharmacological and other medical <u>pain</u> management					
Excluded elements					
	Justification				
High-velocity thrusts to the neck (Cervical manipulation with a rotary component, taking the neck beyond its normal range**)**	To limit any risk of serious adverse effects and because not indicated for pDPN				

Printed or digital materials outside the dedicated NeuOst material	To reduce heterogeneity during a trial and avoid possible contradiction with NeuOst management approach				
Conservative treatment modalities occasionally provided by osteopaths: Acupuncture, bed rest, biofeedback, electrotherapy.	To reduce heterogeneity during a trial and avoid possible contradiction with NeuOst management approach				
Appliances and devices: Belts, Straps, Braces, Splints, Tape, Massage guns, Laser and Ultrasound, Digital Therapeutics, Wellness smartphone or PC applications. (**Not excluded** is NeuP-relevant advice on e.g., compression gloves or socks, heated or low-sensory mittens, wrist splints for Carpal Tunnel Syndrome, desensitization and stereognosis exercises (Andersen Hammond et al., 2020))	To reduce heterogeneity during a trial and avoid possible contradiction with NeuOst management approach To maintain pragmatic treatment provision				

Logic model of the *NeuOst* intervention

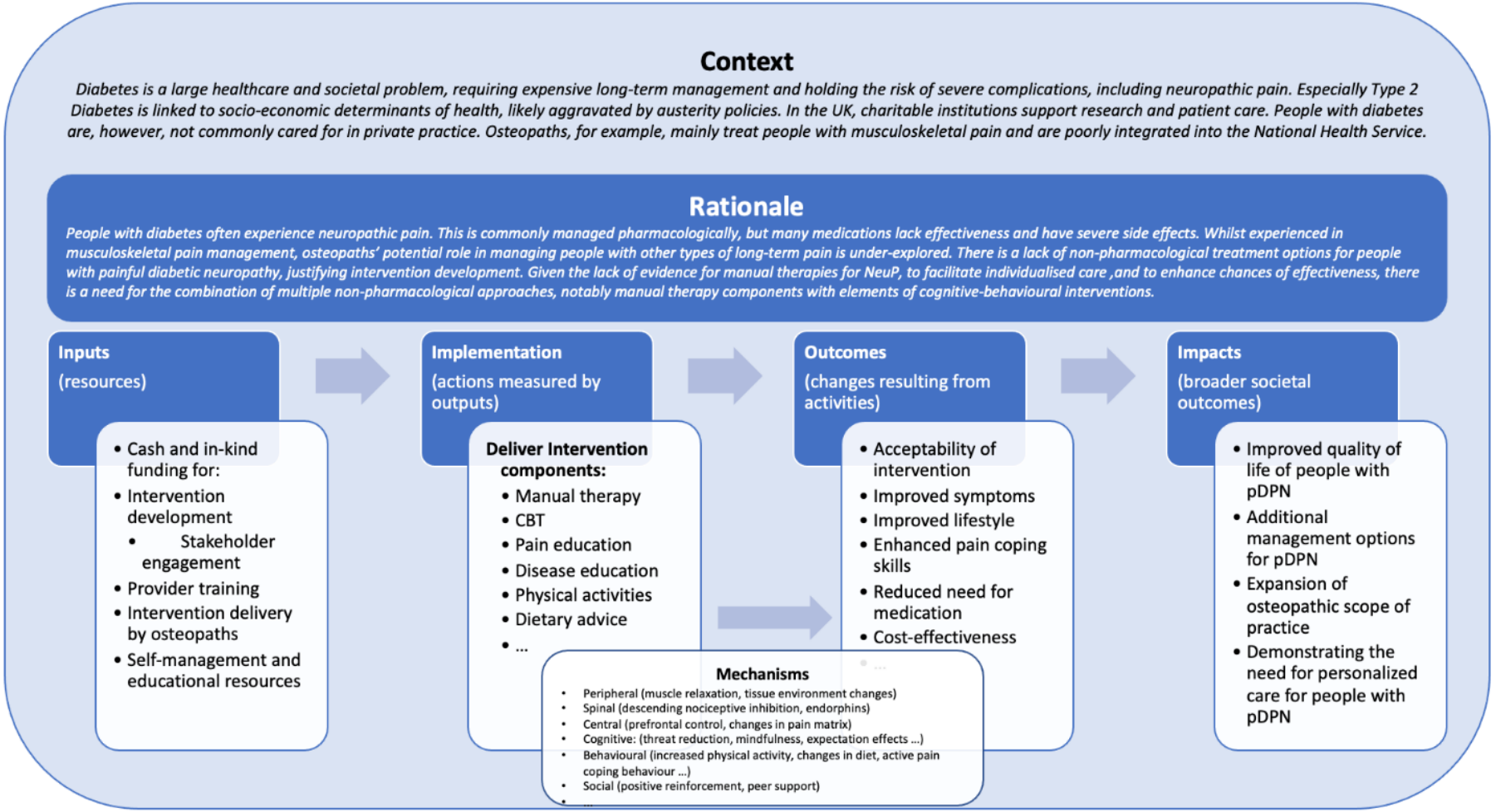


Figure 17: Logic model of NeuOst complex intervention. Design and concept based on: https://www.gov.uk/guidance/evaluation-in-health-and-wellbeing-creating-a-logic-model and https://www.strategyunitwm.nhs.uk/sites/default/files/2018-03/Logic%20models%20and%20complex%20programmes%20-%20a%20brief%20guide.pdf

Appendix 5: Supplementary information for *NeuOst* feasibility trial

Treatment protocol for *NeuOst* feasibility trial

The intervention components will be delivered in a step-wise programme, delivered over six sessions. This approach was chosen to provide practitioners with an initial structure to deliver NeuOst. It will also facilitate high-fidelity intervention delivery during the feasibility trial. Clinically, graded exposure to physical activity, a step-by-step approach to psychological interventions, and repeated sessions of manual therapy are also common.

Table 42: *NeuOst* intervention protocol and fidelity checklist.

Intervention content (providers decide the sequence per session) *Components marked 'optional' are non-obligatory; all others must be delivered as part of the respective treatment session. Elements written in grey recur in from the previous session(s).*	**Fidelity checklist**	**Notes**
Session 1		Session duration up to 1.5 hours
Communication and ACT-informed elements: • Active listening to patient narrative • Initial case history • Introduction to the therapeutic objectives and structure of NeuOst • Acknowledgement of medication list • Conversation about medications (optional) • Communication with primary care / specialist care provider (optional) • Signposting to DM/NeuP-related care options (optional) • Exploration of patient beliefs (regarding pain and disease) (optional) • Education about pain and neuropathic pain, including management strategies, sleep management, and devices (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	Up to 30 mins

• ACT-informed: Exploration of currently employed pain control strategies: Which strategies are you employing? What is their benefit and/or cost? How are these strategies aligned with living the kind of life you want? • ACT-informed: Physical activity and values exploration: Is physical activity important to you? Why / why not (values)? How will doing this physical activity make your life better and more satisfying? What else will it help you do in your life?	☐ ☐	
Manual therapy: • Assessment (inspection, active & passive joint, palpation) • Peripheral joint articulation (feet & ankles at least 2mins per side end range-of-motion; then knees & hips any duration; all others optional) • Spinal joint articulation (optional) • Spinal joint manipulation (optional) • 'Subtle' osteopathic techniques (optional) • Peripheral soft tissue manipulation (gastrocnemius at least 1min per side, all others optional) • Paraspinal soft tissue manipulation (optional) • Active resisted muscle activation (lower extremity) • Neurodynamic technique lower extremity (active); 10 reps • Integration of patient breathing and / or relaxation into manual treatment (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Neurological: • Provider baseline neurological examination (in addition to study baseline testing; not used as outcome)	☐	5 mins
Physical activity: • Raise the topic of physical activity with permission (use motivational interviewing techniques throughout) (5As': "Ask") • Evaluation of current physical activity levels ("Assess") • Explore barriers to more engagement in physical activity • Educate about benefits of different types of physical activity ("Advise") • Values-based goal setting for aerobic exercise activity (patient choice of activity, use Borg scale for exertion level) ("Agree") • Practice session 1 home-based sensorimotor, stretch, strength (SSS) programme content • Decide on suitable modifications for SSS programme (optional) • Provide session 1 handout / link ("Arrange")	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	10 mins

• Provide "wobble cushion" (provided as part of trial)		
• Diabetes complication screening (red flags for foot health, vision changes, kidney disease etc.) • Conversation about dietary DM management (optional) • Education about DM (optional) • Education about DPN and other complications of DM, including management strategies (optional) • Provision of NeuOst educational materials (optional)	☐ ☐ ☐ ☐ ☐	
Home programme content: Includes checklist for self-monitoring ("Assist") *Cardio:* (Duration or distance, and level of exhaustion agreed in session and documented on handout; grade difficulty using Borg Rating Scale of Perceived Exertion to achieve "somewhat hard" to "hard" physical activity [13-15/20]) • 3x p.w.: Engage in agreed amount of aerobic activity *SSS (sensorimotor, stretch, strength) exercises:* (Possible modification: support with handhold on desk etc if needed – safety; Eyes open vs. closed for difficulty; With or without wobble board) • 2x p.d.: 10x slump 'tensioner' neurodynamic movement • 1x p.d.: 10x sit-to-stand • 30-60 secs: Wobble cushion exercise (modifiable from simple two-legged standing to single-leg reaching)		
Session 2		45 mins duration
Communication and <u>ACT-informed elements</u>: • Active listening to patient narrative • Follow-up case history	☐ ☐ ☐	15 mins

• Value and personal goals exploration (incl. brief info about what values are and why they are useful) • Notice Your Breathing exercise • Encouragement to do breathing exercise at home • Conversation about medications (optional) • Communication with primary care / specialist care provider (optional) • Signposting to DM/NeuP-related care options (optional) • Exploration of patient beliefs (regarding pain and disease) (optional) • Education about pain and neuropathic pain, including management strategies, sleep management, and devices (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
Manual therapy: • Follow-up assessment (inspection, active & passive joint, palpation) • Peripheral joint articulation (feet & ankles at least 2mins per side end range-of-motion; then knees & hips any duration; all others optional) • Spinal joint articulation (optional) • Spinal joint manipulation (optional) • 'Subtle' osteopathic techniques (optional) • Peripheral soft tissue manipulation (gastrocnemius at least 1min per side, all others optional) • Paraspinal soft tissue manipulation (optional) • Active resisted muscle activation (lower extremity) • Neurodynamic technique upper and / or lower extremity (passive); 10 reps • Integration of patient breathing and / or relaxation into manual treatment	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Neurological: • Follow-up neurological examination (if indicated, optional) (in addition to study baseline testing; not used as outcome)	☐	
Physical activity: • Revisit experiences with SSS programme • Reinforce importance of exercise (optional) • Explore barriers to adherence and provide advice on overcoming those (optional) • Practice session 2 home-based sensorimotor, stretch, strength (SSS) programme content • Decide on suitable modifications for SSS programme (optional) • Provide session 2 handout / link ("Arrange")	☐ ☐ ☐ ☐ ☐ ☐	5 mins
Diabetes management and education: • Foot care education and run-through	☐	10 mins

• Provide foot care handout • Conversation about dietary DM management (optional) • Education about DM (optional) • Education about DPN and other complications of DM, including management strategies (optional) • Provision of other NeuOst educational materials (optional)	☐ ☐ ☐ ☐ ☐	
Home programme content Week 2: *Cardio:* • 3x p.w.: Engage in agreed amount of aerobic activity *SSS (sensorimotor, stretch, strength) exercises:* • 2x p.d.: 10x slump 'tenser' neurodynamic movement • 1x p.d.: 10x sit-to-stand • 30-60 secs: Wobble cushion exercise (modifiable from simple two-legged standing to single-leg reaching) • 1x p.d.: 10x heel-to-toe raises		
Session 3		45 mins duration
Communication and ACT-informed elements: • Active listening to patient narrative • Follow-up case history • Passenger on the bus Metaphor and reflective questions • Provision of guided meditation recording • Conversation about medications (optional) • Communication with primary care / specialist care provider (optional) • Signposting to DM/NeuP-related care options (optional) • Exploration of patient beliefs (regarding pain and disease) (optional) • Education about pain and neuropathic pain, including management strategies, sleep management, and devices (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Manual therapy: • Follow-up assessment (inspection, active & passive joint, palpation)	☐	15 mins

• Peripheral joint articulation (feet & ankles at least 2mins per side end range-of-motion; then knees & hips any duration; all others optional) • Spinal joint articulation (optional) • Spinal joint manipulation (optional) • 'Subtle' osteopathic techniques (optional) • Peripheral soft tissue manipulation (gastrocnemius at least 1min per side, all others optional) • Paraspinal soft tissue manipulation (optional) • Active resisted muscle activation (lower extremity) • Neurodynamic technique upper and / or lower extremity (passive); 10 reps • Integration of patient breathing and / or relaxation into manual treatment (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	
Neurological: • Follow-up neurological examination (if indicated, optional) (in addition to study baseline testing; not used as outcome)	☐	
Physical activity: • Revisit experiences with SSS programme • Reinforce importance of exercise (optional) • Explore barriers to adherence and provide advice on overcoming those (optional) • Practice session 3home-based sensorimotor, stretch, strength (SSS) programme content • Decide on suitable modifications for SSS programme (optional) • Provide session 3 handout / link	☐ ☐ ☐ ☐ ☐ ☐	5 mins
Diabetes management and education: • Conversation about dietary DM management (optional) • Education about DM (optional) • Education about DPN and other complications of DM, including management strategies (optional) • Provision of NeuOst educational materials (optional)	☐ ☐ ☐ ☐	5 mins
Home programme content Week 3: *Cardio:* • 3x p.w.: Engage in agreed amount of aerobic activity *SSS (sensorimotor, stretch, strength) exercises:* • 2x p.d.: 10x slump 'tenser' neurodynamic movement		

• 1x p.d.: 10x sit-to-stand • 30-60 secs: Wobble cushion exercise (modifiable from simple two-legged standing to single-leg reaching) • 1x p.d.: 10x heel-to-toe raises • 1x p.d.: 10x per side standing knee raises or step-up exercise		
Session 4		45 mins duration
Communication and ACT-informed elements: • Active listening to patient narrative • Follow-up case history • Leaves on a stream exercise (ACT OPEN manual p. 26) • Conversation about medications (optional) • Communication with primary care / specialist care provider (optional) • Signposting to DM/NeuP-related care options (optional) • Exploration of patient beliefs (regarding pain and disease) (optional) • Education about pain and neuropathic pain, including management strategies, sleep management, and devices (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Manual therapy: • Follow-up assessment (inspection, active & passive joint, palpation) • Peripheral joint articulation (feet & ankles at least 2mins per side end range-of-motion; then knees & hips any duration; all others optional) • Spinal joint articulation (optional) • Spinal joint manipulation (optional) • 'Subtle' osteopathic techniques (optional) • Peripheral soft tissue manipulation (gastrocnemius at least 1min per side, all others optional) • Paraspinal soft tissue manipulation (optional) • Active resisted muscle activation (lower extremity) • Neurodynamic technique upper and / or lower extremity (passive); 10 reps • Integration of patient breathing and / or relaxation into manual treatment (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Neurological: • Follow-up neurological examination (if indicated, optional) (in addition to study baseline testing; not used as outcome)	☐	
Physical activity: • Revisit experiences with SSS programme • Reinforce importance of exercise (optional) • Explore barriers to adherence and provide advice on overcoming those (optional)	☐ ☐ ☐	5 mins

• Practice session 4 home-based sensorimotor, stretch, strength (SSS) programme content • Decide on suitable modifications for SSS programme (optional) • Provide session 4 handout / link	☐ ☐ ☐	
Diabetes management and education: • Conversation about dietary DM management (optional) • Education about DM (optional) • Education about DPN and other complications of DM, including management strategies (optional) • Provision of NeuOst educational materials (optional)	☐ ☐ ☐ ☐	5 mins
Home programme content Week 4: *Cardio:* • 3x p.w.: Engage in agreed amount of aerobic activity *SSS (sensorimotor, stretch, strength) exercises:* • 2x p.d.: 10x slump 'tenser' neurodynamic movement • 1x p.d.: 10x sit-to-stand • 30-60 secs: Wobble cushion exercise (modifiable from simple two-legged standing to single-leg reaching) • 1x p.d.: 10x heel-to-toe raises • 1x p.d.: 10x per side standing knee raises / step ups • 1x p.d.: 5x getting up from floor		
Session 5		45 mins duration
Communication and <u>ACT-informed elements</u>: • Active listening to patient narrative • Follow-up case history • <u>Notice five things (Awareness exercise)</u> • <u>Clinician to check how the participant is applying/practicing the mindfulness exercises from earlier sessions. Clinician can discuss barriers to practice.</u> • Conversation about medications (optional) • Communication with primary care / specialist care provider (optional) • Signposting to DM/NeuP-related care options (optional) • Exploration of patient beliefs (regarding pain and disease) (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins

• Education about pain and neuropathic pain, including management strategies, sleep management, and devices (optional)	☐	
Manual therapy: • Follow-up assessment (inspection, active & passive joint, palpation) • Peripheral joint articulation (feet & ankles at least 2mins per side end range-of-motion; then knees & hips any duration; all others optional) • Spinal joint articulation (optional) • Spinal joint manipulation (optional) • 'Subtle' osteopathic techniques (optional) • Peripheral soft tissue manipulation (gastrocnemius at least 1min per side, all others optional) • Paraspinal soft tissue manipulation (optional) • Active resisted muscle activation (lower extremity) • Neurodynamic technique upper and / or lower extremity (passive); 10 reps • Integration of patient breathing and / or relaxation into manual treatment (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Neurological: • Follow-up neurological examination (this week **not** optional) (In addition to study baseline testing; not used as outcome)	☐	
Physical activity: • Revisit experiences with SSS programme • Reinforce importance of exercise (optional) • Explore barriers to adherence and provide advice on overcoming those (optional) • Practice session 5 home-based sensorimotor, stretch, strength (SSS) programme content • Decide on suitable modifications for SSS programme (optional) • Provide session 5 handout / link	☐ ☐ ☐ ☐ ☐ ☐	5 mins
Diabetes management and education: • Conversation about dietary DM management (optional) • Education about DM (optional) • Education about DPN and other complications of DM, including management strategies (optional) • Provision of NeuOst educational materials (optional)	☐ ☐ ☐ ☐	5 mins
Home programme content Week 5: *Cardio:* • 3x p.w.: Engage in agreed amount of aerobic activity		

SSS (sensorimotor, stretch, strength) exercises: • 2x p.d.: 10x slump 'tenser' neurodynamic movement • 1x p.d.: 10x sit-to-stand • 30-60 secs: Wobble cushion exercise (modifiable from simple two-legged standing to single-leg reaching) • 1x p.d.: 10x heel-to-toe raises • 1x p.d.: 10x per side standing knee raises / step-ups • 1x p.d.: 5x getting up from floor • 1x p.d.: 10x per side Star-steps (front, side, back large steps)		
Session 6		45 mins duration
Communication and ACT-informed elements: • Active listening to patient narrative • Follow-up case history • Setting longer-term values-based goals for physical activity/lifestyle changes/diabetes management. Discussing how the participant can draw on skills learned from the treatment to respond to barriers that will show up moving forward. • Conversation about medications (optional) • Communication with primary care / specialist care provider (optional) • Signposting to DM/NeuP-related care options (optional) • Exploration of patient beliefs (regarding pain and disease) (optional) • Education about pain and neuropathic pain, including management strategies, sleep management, and devices (optional)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins
Manual therapy: • Follow-up assessment (inspection, active & passive joint, palpation) • Peripheral joint articulation (feet & ankles at least 2mins per side end range-of-motion; then knees & hips any duration; all others optional) • Spinal joint articulation (optional) • Spinal joint manipulation (optional) • 'Subtle' osteopathic techniques (optional) • Peripheral soft tissue manipulation (gastrocnemius at least 1min per side, all others optional) • Paraspinal soft tissue manipulation (optional) • Active resisted muscle activation (lower extremity)	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐	15 mins

• Neurodynamic technique upper and / or lower extremity (passive); 10 reps • Integration of patient breathing and / or relaxation into manual treatment (optional)	☐ ☐	
Neurological: • Follow-up neurological examination (optional) (In addition to study baseline testing; not used as outcome)	☐	10 mins
Physical activity: • Revisit experiences with SSS programme • Reinforce importance of exercise (optional) • Explore barriers to adherence and provide advice on overcoming those (optional) • Practice session 6 home-based sensorimotor, stretch, strength (SSS) programme content • Decide on suitable modifications for SSS programme (optional) • Provide session 6 handout / link	☐ ☐ ☐ ☐ ☐ ☐	5 mins
Diabetes management and education: • Conversation about dietary DM management (optional) • Education about DM (optional) • Education about DPN and other complications of DM, including management strategies (optional) • Provision of NeuOst educational materials (optional)	☐ ☐ ☐ ☐	
Home programme content Week 6: *Cardio:* • 3x p.w.: Engage in agreed amount of aerobic activity *SSS (sensorimotor, stretch, strength) exercises:* • 2x p.d.: 10x slump 'tenser' neurodynamic movement • 1x p.d.: 10x sit-to-stand • 30-60 secs: Wobble cushion exercise (modifiable from simple two-legged standing to single-leg reaching) • 1x p.d.: 10x heel-to-toe raises • 1x p.d.: 10x per side standing knee raises • 1x p.d.: 10x supine bridge • 1x p.d.: 10x per side Star-steps (front, side, back large steps)		

PRECIS-2 table for *NeuOst* feasibility trial

Table 43: PRECIS-2 table for NeuOst feasibility trial: Each domain is scored on a scale from 1 (very explanatory) to 5 (very pragmatic), indicating how similar each design feature is compared to usual care. Adapted from Loudon et al. (2015).

	Domain	Score	Rationale
1	Eligibility criteria	4	The exclusion criteria are active foot ulcerations, which would likely be considered a contraindication to local manual therapy by many practitioners. Further, people with amputated limbs and other advanced complications of DM are not eligible (advanced renal failure). Whilst amputations are common in DPN and osteopaths may choose to treat people with (unilateral) amputations for their symptoms on the remaining limb or for phantom limb pain, these patients were excluded due to potentially reduced comparability in terms of disability levels and ability to engage in the home-exercise component of *NeuOst*. We will be clear in the dissemination of findings that this *NeuOst* may require adaptation to this subgroup of diabetic neuropathy patients. Further, treatments during the feasibility trial are markedly subsidised, meaning that trial participants may be financially less well-off than those who can afford *NeuOst* outside a trial. Apart from that, participants are not excluded for any medication, comorbidity, co-diagnosis, or physical characteristics (e.g., BMI). Therefore, only a single point was deduced.
2	Recruitment Path	4	Recruitment to *NeuOst* is through local newspapers, social media, and potentially NHS services. Whilst osteopaths commonly employ social media to advertise their services, newspaper advertising and advertising targeted at a specific patient population may represent a slightly larger effort than usual. At the same time, people with pDPN are not typically seen by osteopaths, so that anyone wanting to build a patient base of people with pDPN in their normal clinical practice would have to employ additional recruitment pathways.
3	Setting	1	The *NeuOst* feasibility trial will be run at a single site only. This site is the teaching clinic at the UK's largest osteopathic academic institution.

			If *NeuOst* is offered as training to qualified osteopaths, most would deliver the intervention in a private practice setting. Practically, however, it is questionable how large the implication of this domain for applicability of trial findings are, since processes at the trial site are comparable to those in routine practice.
4	Organisation	4	The same training as during the trial would be available to any qualified osteopaths wanting to deliver *NeuOst*, as part of the regular Continued Professional Development scheme. Other than two-years' minimum experience since graduation, trial providers do not have to fulfil particular criteria. In fact, those *without* experience with working with people with neuropathies are preferred, making the trial providers more representative of the general population of UK osteopaths. Given that there are only two providers in *NeuOst*, it cannot be guaranteed that these are not more experienced or ambitious than the average osteopath. No additional resources will be available to deliver *NeuOst* during the trial, and the clinic administrative is comparable to routine practice. Potentially, the frequency of care delivery may be more than what some providers offer in clinical practice, but variability of treatment frequency is currently unclear.
5	Flexibility of experimental intervention (delivery)	2	The treatment is semi-protocolised and treatment protocol compliance is monitored during *NeuOst* to increase the confidence in that the trial findings are due to elements of the intervention itself. If protocol compliance is found to be wanting during interim checking, measures will be taken to enhance provider compliance. This explanatory approach is appropriate at this stage because of the newly developed nature of the intervention and the potential for uncontrolled heterogeneity given the breath of potential treatment components. At the same time, the intervention protocol entails many elements that are optional, allowing for significant flexibility and individualisation in the intervention delivery. This is likely, representative of how osteopaths would apply *NeuOst* in practice. Conversely, some elements of routine osteopathic practice are not permissible (taping, dry needling, etc) and the exercise protocol is rather prescriptive.
6	Flexibility of experimental intervention (adherence)	4	Encouragement to attend treatment sessions is likely more than in usual care, with the importance of attendance for research purposes highlighted to participants. Participants may be contacted in case of missed appointments, which does happen in routine care, too. Participants are not

			excluded from the trial if they miss appointments and adherence to the home exercise programme is not encouraged beyond what would be considered best-practice.
7	Follow-up	2	Follow-up contact and data collection is significantly more than what would occur in normal practice, with outcome collection at the end of the treatment period and via a long telephone interview 4 weeks after the treatment period. However, no additional visits to the research site will be required.
8	Outcome	N/A	Not applicable, as primary outcomes of a feasibility trial cannot be expected to be clinically relevant. Referring to the secondary outcomes, their relevance to routine practice is likely high, having been chosen based on existing evidence specifically for their potential importance for people with pDPN. However, a primary clinical outcome has not been chosen at this stage, so that this domain is not rated.
9	Analysis	(5)	For feasibility outcomes, all data will be included in the analysis, irrespective of participants' trial completion or missing data. Whilst compliant with PRECIS-2 wording, this tells us little about the real-world applicability of the intervention and more about the real-world feasibility of a full-scale trial. For clinical outcomes (not rated here), both intention-to-treat and completer analyses will be explored.

Appendix 6: Academic development and awards

The following is a selection of intra- and extracurricular training obtained during the PhD project. Teaching activities are listed separately. Contributions to academic conferences are listed in Table 44, awards and grants in Table 45, and teaching and supervision activities in Table 46.

Grad School Courses:

1. Ensuring Integrity 1: Plagiarism Awareness (Online Course) 01/10/2019
2. Maximising Your Management Skills: Putting Project Management into Action 21/10/2019
3. Impact in Academia: Making an Impact 26/02/2020
4. Global Postgraduate Retreats: Impact and Influence 03/03/2020
5. Data Science: R Programming 18/11/2020
6. Presentation Skills 12/05/2021
7. Introduction to Philosophy – Knowledge, Truth and Science? 10/05/2021
8. Preparing for Thesis Submission 10/10/2022

Extra-curricular courses:

1. Introduction to Good Clinical Practice (GCP) eLearning, NIHR, October 2019
2. Systematic reviews and Meta-Analysis, University of York, September 2019 (3 days)
3. Essentials of Clinical Trials Programme, London School of Hygiene & Tropical Medicine, 5th – 9th July 2021
4. ICTU Contemporary Clinical trials Course, Imperial College London Clinical Trials Unit, Autumn 2021 (10 dates)

Table 44: Conference contributions and presentations. Presentations were the result of a competitive application process, unless indicated as 'invited'. Presentation titles in quotation marks. TBC, to be confirmed.

Conference contribution	Location and date
Invited topical lecture (virtual): **EFIC Virtual Pain Education Summit 2021**, "Application of digital technology to facilitate practice"	Virtual, 5 November 2021
Invited Webinar (virtual), **EuroSpine**, "Remote Consultations"	Virtual, May 2021
Presentation (virtual): **Innovations scientifiques, economiques et reglementaires, iCEPS**, "Méthodes d'utilisation du placebo et de la mise en aveugle des essais cliniques pour les thérapies physiques, psychologiques et d'auto-gestion pour la douleur : une revue systématique avec méta-analyse"	Virtual, 01 April 2021
Presentation (virtual): **Australian Trials Methodology Conference**, "Sham Control Methods in Physical, Psychological, and Self-Management Intervention Trials for Pain: A Systematic Review and Meta-Analysis"	Virtual, 07 December 2021
Presentation (virtual): **Institute of Osteopathy Conference**, "Intervention Development and Feasibility Study of Osteopathic Management for People with Painful Diabetic Polyneuropathy"	Virtual, November 2022
Presentation (virtual): **National Council for Osteopathic Research**, "What Do Clinical Trials Actually Tell Us and What Use Is That? Current Challenges of Clinical Trial Design in Physical Interventions and Possible Ways Forward"	Virtual, January 2021
Presentation (virtual): **Annual Congress of the Osteopathie Schule Deutschland**, "The role of the placebo effect in manual therapy practice"	Virtual, November 2021
Presentation: **The Society for Back Pain Research annual congress**, "Recommendations For The Development, Implementation, And Reporting Of (Sham) Control Interventions In Efficacy Trials Of Physical, Psychological, And Self-Management Therapies - The CoPPS Statement."	Warwick, UK, June 2022

Poster: **World Congress of the International Association for the Study of Pain**, "Developing a Sham Control Intervention for a Complex Multimodal Therapy Following a New Consensus Guideline: A Mixed Methods Process Evaluation"	Toronto, Canada, September 2022
Symposium presentation: **Kongress der Deutschen Schmerzgesellschaft**, "Interventionsentwicklung und Sham-kontrollierte Machbarkeitstudie osteopathischen Managements für Menschen mit schmerzhafter diabetischer peripherer Polyneuropathie"	Mannheim, Germany, October 2022
Invited presentation: **The Royal Society of Medicine**, A clinical masterclass on treatments for painful neuropathies, "An Innovative Drug-Free Treatment for People with Painful Diabetic Neuropathy: Laying the Foundations for Clinical Testing and Future Perspectives for Patients"	London, UK, 20 March 2023
Upcoming	
Invited small-group presentation and discussion: **HealthXL**, "Masterclass: Design and Control Conditions for Clinical Trials of Digital Health Solutions"	Virtual, 29 March 2023
Invited lecture on the course The Science of the Placebo Effect: **Karolinska Institute**, "Clinical trial design and placebo effects"	Stockholm, Sweden, 03 May 2023
Symposium presentation: **Society for Interdisciplinary Placebo Studies, annual conference**, "A new Guideline for Control Interventions in Efficacy and Mechanistic Trials of Physical, Psychological, and Self-Management Therapies - The CoPPS Statement and Its Practical Application."	Duisburg, Germany, May 2023
Invited webinar: **British Pain Society**, title TBC	Late 2023, exact date TBC
Invited keynote presentation (virtual): **Universidad de La Sabana, Bogotá, III International Meeting of Academics in Manual Therapy**, title TBC	Universidad de La Sabana, Bogotá, Colombia
Invited keynote presentation and workshops: **The McKenzie Institute Conference**, title TBC	Ottawa, Canada, October 2024

Table 45 lists research awards and funding support obtained during the PhD project (not considering the PhD studentship itself).

Table 45: Awards and personal grants. Unless otherwise indicated, David Hohenschurz-Schmidt was the sole recipient of an award or the principal investigator on a grant.

Award or grant	**Awarded for**
Community of Practice Award (Institute of Osteopathy, 2020)	Leading a task force that facilitated the transition to video and phone consultations for musculoskeletal practice during the early phase of the COVID-19 pandemic, providing webinar training to over 1,700 clinicians in three languages and publishing dedicated guidance (Hohenschurz-Schmidt et al., 2020b).
Research Grant of £19,863 (The Osteopathic Foundation, 2021)	For the *NeuOst* project.
Travel Fellowship of £2,500 (Society for Back Pain Research, 2021)	To facilitate an international and multidisciplinary research collaboration for the *NeuOst* project
New Investigator Award (Society for Back Pain Research, 2022)	In recognition of the work on control interventions presented at the annual conference in Warwick.
Named Collaborator on Research Grant of £18,600 (The Osteopathic Foundation, 2022)	For a project investigating the trustworthiness of osteopathic research and to build research capacity within the profession by mentoring an aspiring osteopathic academic.

Table 46: Overview of teaching and supervision activities

Activities and setting	Hours per Year
Lectures, osteopathic higher education (undergraduate level)	12 (2019) 46 (2020) 74 (2022) 20 (2023)
Student supervision, higher education (undergraduate level)	20 (2020) 28 (2021) 20 (2022) 15 (2023)

Appendix 7: Outreach activities and additional contributions

Apart from work on clinical trial methods and the *NeuOst* project, I participated in or led various other endeavours in the field of nonpharmacological intervention science and practice and gave numerous presentations. In brief, I strove to encourage attention to methodological rigour in the scientific community, and enable clinicians to appraise and utilise evidence effectively. As an osteopath by training, I also attempted to provide clinicians with an appreciation of the complexity of clinical practice, highlighting the role of contextual factors and the importance of interpersonal skills. These activities are described and reflected on below; Presentations are summarised in Table 44.

Engagement with scientific and clinician communities

Through letters to editors (Hohenschurz-Schmidt et al., 2021, 2022a), my involvement in a designated methodological article (Draper-Rodi et al., 2022a), and conference presentations (e.g., 2021 National Council of Osteopathic Research conference), I sought to educate stakeholders about the link between trial design and the interpretation of results. At all points have I highlighted that the challenges are complex, likely not amenable to one-size-fits-all solutions, and that it is mainly the responsibility of authors to transparently report their methods and offer a truthful interpretation. Attempts to improve reporting across the explanatory-pragmatic spectrum are seen widely elsewhere (Zwarenstein et al., 2008; Howick et al., 2020), also in PPS interventions (Alvarez et al., 2016, 2020). Gerard Alvarez and Sònia Roura Carvajal from Barcelona are contributing important research to pragmatic trial methods and reporting in manual therapy, and we are currently collaborating on a manual therapy-specific systematic review. Most recently, I was invited to serve on the 2023 IASP Global Year Task Force on Integrative Pain Care (https://www.iasp-pain.org/advocacy/global-year/integrative-pain-care/). Here, I led the development of a fact sheet on evaluation methods for integrative models of pain care (publication imminent), highlighting the challenges and opportunities of complex intervention research. This endeavour was also useful to consider how the methods discussed in

the IMMPACT and CoPPS statements relate to other research methodologies, such as preclinical, experimental, and single case designs (See Discussion chapter of this thesis). Similarly, I have been invited to contribute to a position statement on research waste in spinal manipulation research, collaborating with the renown Prof Rachelle Buchbinder.

Next to improved research methods and reporting, evidence-based practice requires active engagement of clinicians (Leach, 2006). Research will only translate into clinical practice if critical appraisal and decision-making capacities of clinicians are promoted, allowing them to identify high-quality research. Therefore, continued engagement at the level of undergraduate and continued professional education is required. My own involvement in the practitioner-oriented paper '4 M's to make sense of evidence – Avoiding the propagation of mistakes, misinterpretation, misrepresentation and misinformation' (Draper-Rodi et al., 2022a) is an example of such efforts. Similarly, the involvement of several osteopathic clinicians in the reviewer crowd of the control intervention review can be understood as a form of engagement. I have also supervised several BSc and MSc projects at University College of Osteopathy, Osteopathie Schule Deutschland, Aarhus University, Imperial College London, and the Haute école de santé Fribourg. All these projects related to my own areas of research interest, and I strove to instil a sense of importance for rigorous methodology. One project resulted in a published systematic review (Bini et al., 2022). Next, I will mentor an aspiring osteopathic academic as part of an 8-month project on research trustworthiness in osteopathy and help the National Council for Osteopathic Research in setting up a support network which will provide osteopathic researchers with feedback on study protocols.

Promoting a reconceptualisation of MSK practice

Beyond promoting an understanding of research methods, I have tried to help practitioners reconceptualise manual and physical therapies: away from a biomechanical-reductionist understanding towards conceptual frameworks that embrace intervention complexity and the multifactorial nature of ill-health. In doing so, I strove to equip clinicians with practical tools to make positive experiences with.

First, I initiated a remote consultation 'task force' during the early phase of the COVID-19 pandemic. Over the course of three weeks, we developed and delivered webinars for musculoskeletal (MSK) clinicians introducing them to the key considerations for seeing patients remotely. Ten webinars were held in three languages. Over 1,700 people attended live webinars, recordings continue to be available for free (https://www.uco.ac.uk/free-remote-consultation-training-osteopaths). We subsequently recorded a podcast episode on the topic (https://www.wordsmatter-education.com/blog/podcast-0x-working-remotely-and-relying-on-words-with-david-hohenshurz-schmidt) and published a well-cited Masterclass article in PAIN Reports (Hohenschurz-Schmidt et al., 2020b). While the primary objective was to facilitate care continuity during the pandemic, our optimism about remote MSK care implied that hands-on techniques were not the only, and potentially not the most important component of manual therapies, highlighting communication and best-practice advice as potentially powerful therapeutic elements. The questions and indeed criticism that this message evoked from clinicians (since reflected by several qualitative studies, e.g., Derbyshire et al., 2021; Malliaras et al., 2021) led us to carefully underscore the benefits of pain self-management, relationship-building, and various communication strategies in our subsequent article (Hohenschurz-Schmidt et al., 2020).

Secondly, to foster such a reconceptualisation of manual therapy practice, I volunteered to organise the programme of the 2021 online conference of the Osteopathie Schule Deutschland. Having worked with the school founder Torsten Liem for years, I knew that the

designated topic “Osteopathy Meets Psychology” would not necessarily result in an evidence-based conference programme. I therefore reached out to leaders in the field who I knew where promoting evidence-based care, scheduling about two-thirds of the speakers, including Professors Frank Keefe and Lisa Roberts (Full programme here: https://osteopathie-schule.de/kongress/online-conference-2021/). While the feedback and discussions at the conference were largely positive, I would have hoped for more engagement with practitioners during a virtual conference that was simultaneously translated into four languages. Apparently, biomechanically oriented conferences still attract larger audiences and more engagement in the manual therapy field.

My third approach to broaden our understanding of manual therapy practice was through evidence-based discussions of the placebo and nocebo effects, highlighting the importance of so-called contextual factors for therapeutic outcomes. This evolved partly through my association with Professor Lene Vase (Aarhus University) and partly through the pertinence of the placebo effect for my blinding-related work. I attempted to positively re-frame the placebo effect, for example during my own lecture at the above osteopathic conference or a German-language podcast (Episode #16: https://physiobib.de/physiobib-podcast/), but also through dedicated papers and webinars (in preparation). Supervising two BSc projects, we explored the use of language in osteopathic consultations, notably regarding the creation of positive and negative expectations. Leading a multi-professional collaboration, I also published an ‘in invitation to reflect’ on the nocebo effect and other undesirable effects in chiropractic, osteopathy, and physiotherapy (Hohenschurz-Schmidt et al., 2022d). In January 2023, this was the most-downloaded article of the journal Musculoskeletal Science & Practice of the past 3 months, and we are currently planning a podcast episode on the topic (https://www.wordsmatter-education.com/blog/episode-78). Currently, I am leading a systematic review with a planned meta-analysis to quantify placebo effects in three-armed trials of PPS interventions (i.e., comparing no treatment against ‘sham’ control interventions). Despite my attempts to frame the increased awareness of communication and contextual factors in clinical practice positively, conflicts with traditional explanatory models arise on occasion, and apprehension due to

impingements on clinicians' sense of professional identity is often palpable during interactions. My own teaching of osteopathic courses on clinical reasoning, pain, or the autonomic nervous system have long since emphasised the relative importance of communication, lifestyle, socioeconomic factors, and psycho-neuro-endocrine mechanisms for health and disease. These courses reach about 120 undergraduate students a year.

In summary, while reconceptualisation towards evidence-based and psychosocially informed frameworks is advancing in academic circles, there are several challenges to this process in manual and physical therapy clinician communities requiring ongoing attention.

Expanding the scope of manual therapy practice

Apart from improving the basis for clinical research in PPS interventions and expanding the conceptual understanding of manual therapists, the project has begun to explore a novel way to conservatively manage a group of people with multiple comorbidities and often severely impacted quality of life. By providing targeted multidisciplinary training to osteopaths, their potential scope of practice is hoped to expand in line with the reconceptualisation towards contemporary practice models. Drawing on multiple existing promising management and treatment options resulted in a treatment worthy of further investigation. Finally, integrating a touch-based component into the management of people with a painful polyneuropathy opens an avenue into the investigation of this component's mechanisms. The *NeuOst* project also challenges the contemporary self-payer models of manual practice. As touched on briefly in Chapter 6 and in an article on undesirable effects of chiropractic, osteopathy, and physiotherapy (Hohenschurz-Schmidt et al., 2022d), alternative financing models and innovative models of care are required if osteopaths want to care for people from disadvantaged and marginalised communities. To do so would mean to play a role in the addressing of important public health needs (Briggs et al., 2020), align with priorities of nonpharmacological intervention research for people with pain (Becker et al., 2018), and potentially equip manual therapists with the much-needed updated concepts to justify their existence (Vogel, 2021).

Exploring digital health technologies for pain management

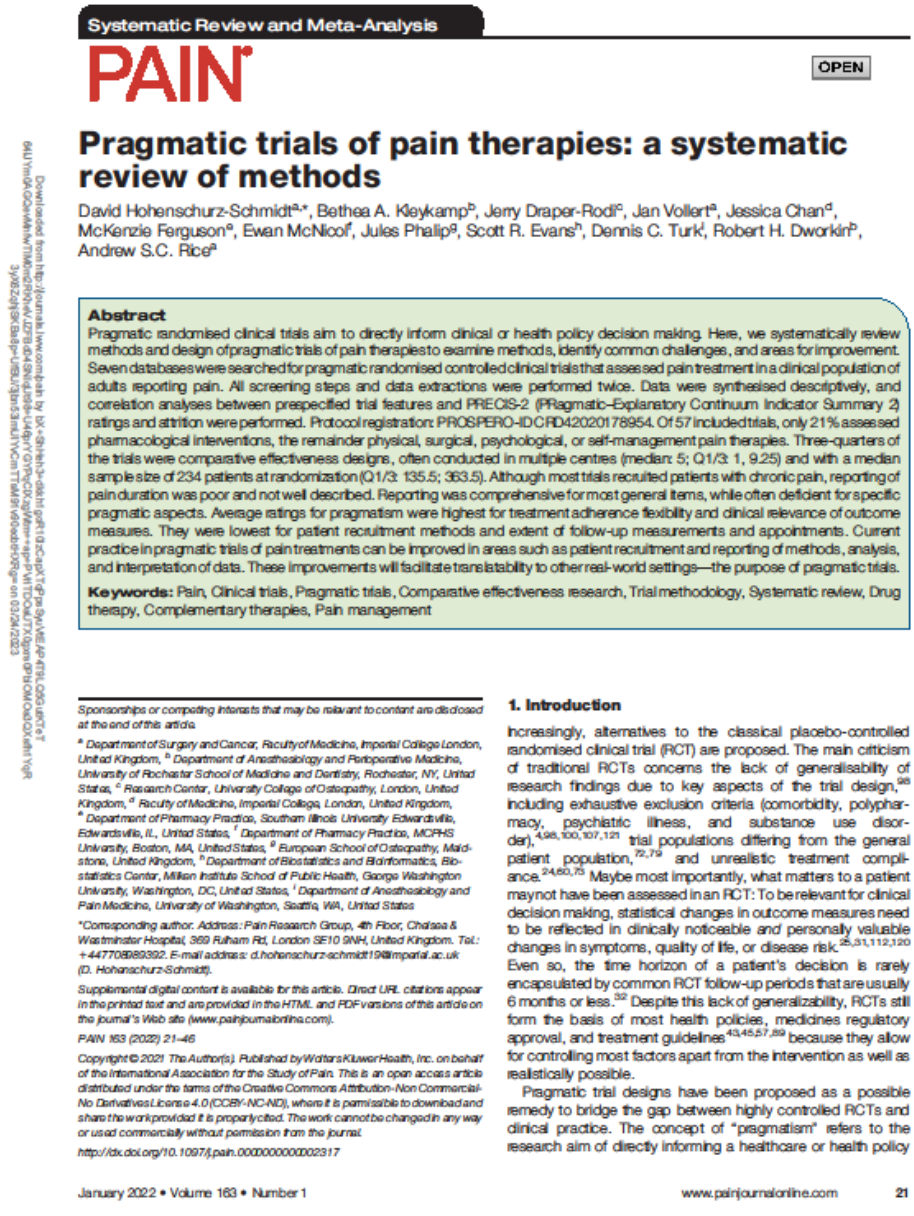
Finally, an innovative approach to pain management is the utilisation of technology to deliver care and engage patients in effective self-care strategies. Through COVID-19 and the remote consultations task force I was first exposed to the topic. My involvement led to me being approached by a start-up company wanting to develop online consultation programme with osteopaths. With my regular input, the endeavour evolved towards a multimodal pain management application that is now going into its third piloting phase. Through this project, I became 'thought leader' on digital therapeutic (DTx) of the online community 'HealthXL', where I am delivering a masterclass on clinical trial design (https://portal-beta.healthxl.com/communities/29dab95f-6ad6-4a8a-b3eb-67ea1a4ea6ec?s=bgmbi&tab=members). Once again, I am combining different research interest with the overarching objective to instil research rigour, an urgent need in the rapidly developing and only partly regulated field of DTx. It must be noted that the digitalisation of healthcare is not without downsides. I am aware of the risk to further-promote the cutting of public services at the favour of private companies, usually bringing with it a deterioration of care quality and working conditions; what has been called the 'uberisation' of healthcare (Al Dahdah, 2022).

Appendix 8: Peer-reviewed publications

For convenience, the main articles published as part of this PhD project are provided as appendices. To view the documents, click the respective icon which will open a Google Drive link to download the pdf file from.

Main publications

1. Hohenschurz-Schmidt, D., Kleykamp, B.A., Draper-Rodi, J., Vollert, J., Chan, J., Ferguson, M., McNicol, E., Phalip, J., Evans, S.R., Turk, D.C., Dworkin, R.H., Rice, A.S.C., 2022. **Pragmatic trials of pain therapies: a systematic review of methods**. PAIN 163, 21. https://doi.org/10.1097/j.pain.0000000000002317



Systematic Review and Meta-Analysis

PAIN

OPEN

Pragmatic trials of pain therapies: a systematic review of methods

David Hohenschurz-Schmidt[a,*], Bethea A. Kleykamp[b], Jerry Draper-Rodi[c], Jan Vollert[a], Jessica Chan[d], McKenzie Ferguson[e], Ewan McNicol[f], Jules Phalip[g], Scott R. Evans[h], Dennis C. Turk[i], Robert H. Dworkin[b], Andrew S.C. Rice[a]

Abstract

Pragmatic randomised clinical trials aim to directly inform clinical or health policy decision making. Here, we systematically review methods and design of pragmatic trials of pain therapies to examine methods, identify common challenges, and areas for improvement. Seven databases were searched for pragmatic randomised controlled clinical trials that assessed pain treatment in a clinical population of adults reporting pain. All screening steps and data extractions were performed twice. Data were synthesised descriptively, and correlation analyses between prespecified trial features and PRECIS-2 (PRagmatic–Explanatory Continuum Indicator Summary 2) ratings and attrition were performed. Protocol registration: PROSPERO-ID CRD42020178954. Of 57 included trials, only 21% assessed pharmacological interventions, the remainder physical, surgical, psychological, or self-management pain therapies. Three-quarters of the trials were comparative effectiveness designs, often conducted in multiple centres (median: 5; Q1/3: 1, 9.25) and with a median sample size of 234 patients at randomization (Q1/3: 135.5; 363.5). Although most trials recruited patients with chronic pain, reporting of pain duration was poor and not well described. Reporting was comprehensive for most general items, while often deficient for specific pragmatic aspects. Average ratings for pragmatism were highest for treatment adherence flexibility and clinical relevance of outcome measures. They were lowest for patient recruitment methods and extent of follow-up measurements and appointments. Current practice in pragmatic trials of pain treatments can be improved in areas such as patient recruitment and reporting of methods, analysis, and interpretation of data. These improvements will facilitate translatability to other real-world settings—the purpose of pragmatic trials.

Keywords: Pain, Clinical trials, Pragmatic trials, Comparative effectiveness research, Trial methodology, Systematic review, Drug therapy, Complementary therapies, Pain management

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, United Kingdom, [b] Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States, [c] Research Center, University College of Osteopathy, London, United Kingdom, [d] Faculty of Medicine, Imperial College, London, United Kingdom, [e] Department of Pharmacy Practice, Southern Illinois University Edwardsville, Edwardsville, IL, United States, [f] Department of Pharmacy Practice, MCPHS University, Boston, MA, United States, [g] European School of Osteopathy, Maidstone, United Kingdom, [h] Department of Biostatistics and Bioinformatics, Biostatistics Center, Milken Institute School of Public Health, George Washington University, Washington, DC, United States, [i] Department of Anesthesiology and Pain Medicine, University of Washington, Seattle, WA, United States

**Corresponding author. Address: Pain Research Group, 4th Floor, Chelsea & Westminster Hospital, 369 Fulham Rd, London SE10 9NH, United Kingdom. Tel.: +447708883392. E-mail address: d.hohenschurz-schmidt19@imperial.ac.uk (D. Hohenschurz-Schmidt).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 163 (2022) 21–46

Copyright © 2021 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Association for the Study of Pain. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

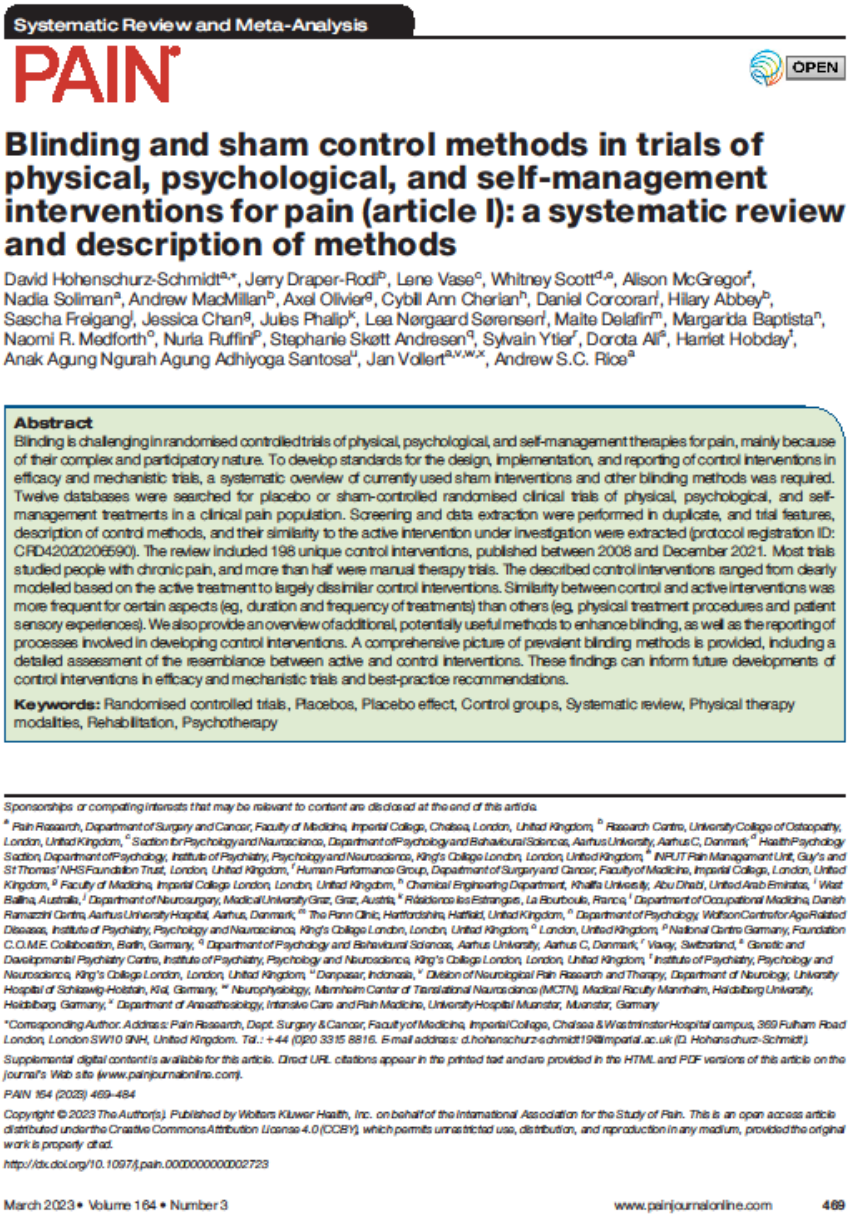
http://dx.doi.org/10.1097/j.pain.0000000000002317

1. Introduction

Increasingly, alternatives to the classical placebo-controlled randomised clinical trial (RCT) are proposed. The main criticism of traditional RCTs concerns the lack of generalisability of research findings due to key aspects of the trial design,[98] including exhaustive exclusion criteria (comorbidity, polypharmacy, psychiatric illness, and substance use disorder),[4,98,100,107,121] trial populations differing from the general patient population,[72,79] and unrealistic treatment compliance.[24,60,75] Maybe most importantly, what matters to a patient may not have been assessed in an RCT: To be relevant for clinical decision making, statistical changes in outcome measures need to be reflected in clinically noticeable *and* personally valuable changes in symptoms, quality of life, or disease risk.[25,31,112,120] Even so, the time horizon of a patient's decision is rarely encapsulated by common RCT follow-up periods that are usually 6 months or less.[32] Despite this lack of generalizability, RCTs still form the basis of most health policies, medicines regulatory approval, and treatment guidelines[43,45,57,89] because they allow for controlling most factors apart from the intervention as well as realistically possible.

Pragmatic trial designs have been proposed as a possible remedy to bridge the gap between highly controlled RCTs and clinical practice. The concept of "pragmatism" refers to the research aim of directly informing a healthcare or health policy

2. Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., Scott, W., McGregor, A., Soliman, N., MacMillan, A., Olivier, A., Cherian, C.A., Corcoran, D., Abbey, H., Freigang, S., Chan, J., Phalip, J., Nørgaard Sørensen, L., Delafin, M., Baptista, M., Medforth, N.R., Ruffini, N., Skøtt Andresen, S., Ytier, S., Ali, D., Hobday, H., Santosa, A.A.N.A.A., Vollert, J., Rice, A.S.C., 2023. **Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article I): a systematic review and description of methods**. PAIN 164, 469. https://doi.org/10.1097/j.pain.0000000000002723



Systematic Review and Meta-Analysis

PAIN

OPEN

Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article I): a systematic review and description of methods

David Hohenschurz-Schmidt[a,*], Jerry Draper-Rodi[b], Lene Vase[c], Whitney Scott[d,e], Alison McGregor[f], Nadia Soliman[a], Andrew MacMillan[b], Axel Olivier[g], Cybill Ann Cherian[h], Daniel Corcoran[i], Hilary Abbey[b], Sascha Freigang[j], Jessica Chan[g], Jules Phalip[k], Lea Nørgaard Sørensen[l], Maite Delafin[m], Margarida Baptista[n], Naomi R. Medforth[o], Nuria Ruffini[p], Stephanie Skøtt Andresen[q], Sylvain Ytier[r], Dorota Ali[s], Harriet Hobday[t], Anak Agung Ngurah Agung Adhiyoga Santosa[u], Jan Vollert[a,v,w,x], Andrew S.C. Rice[a]

Abstract

Blinding is challenging in randomised controlled trials of physical, psychological, and self-management therapies for pain, mainly because of their complex and participatory nature. To develop standards for the design, implementation, and reporting of control interventions in efficacy and mechanistic trials, a systematic overview of currently used sham interventions and other blinding methods was required. Twelve databases were searched for placebo or sham-controlled randomised clinical trials of physical, psychological, and self-management treatments in a clinical pain population. Screening and data extraction were performed in duplicate, and trial features, description of control methods, and their similarity to the active intervention under investigation were extracted (protocol registration ID: CRD42020206590). The review included 198 unique control interventions, published between 2008 and December 2021. Most trials studied people with chronic pain, and more than half were manual therapy trials. The described control interventions ranged from clearly modelled based on the active treatment to largely dissimilar control interventions. Similarity between control and active interventions was more frequent for certain aspects (eg, duration and frequency of treatments) than others (eg, physical treatment procedures and patient sensory experiences). We also provide an overview of additional, potentially useful methods to enhance blinding, as well as the reporting of processes involved in developing control interventions. A comprehensive picture of prevalent blinding methods is provided, including a detailed assessment of the resemblance between active and control interventions. These findings can inform future developments of control interventions in efficacy and mechanistic trials and best-practice recommendations.

Keywords: Randomised controlled trials, Placebos, Placebo effect, Control groups, Systematic review, Physical therapy modalities, Rehabilitation, Psychotherapy

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] Pain Research, Department of Surgery and Cancer, Faculty of Medicine, Imperial College, Chelsea, London, United Kingdom, [b] Research Centre, University College of Osteopathy, London, United Kingdom, [c] Section for Psychology and Neuroscience, Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus C, Denmark, [d] Health Psychology Section, Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, [e] NIHR Pain Management Unit, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom, [f] Human Performance Group, Department of Surgery and Cancer, Faculty of Medicine, Imperial College, London, United Kingdom, [g] Faculty of Medicine, Imperial College London, London, United Kingdom, [h] Chemical Engineering Department, Khalifa University, Abu Dhabi, United Arab Emirates, [i] West Ballina, Australia, [j] Department of Neurosurgery, Medical University Graz, Graz, Austria, [k] Résidence les Estranges, La Bourboule, France, [l] Department of Occupational Medicine, Danish Ramazzini Centre, Aarhus University Hospital, Aarhus, Denmark, [m] The Penn Clinic, Hertfordshire, Hatfield, United Kingdom, [n] Department of Psychology, Wolfson Centre for Age-Related Diseases, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, [o] London, United Kingdom, [p] National Centre Germany, Foundation C.O.M.E. Collaboration, Berlin, Germany, [q] Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus C, Denmark, [r] Vevey, Switzerland, [s] Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, [t] Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom, [u] Denpasar, Indonesia, [v] Division of Neurological Pain Research and Therapy, Department of Neurology, University Hospital of Schleswig-Holstein, Kiel, Germany, [w] Neurophysiology, Mannheim Center of Translational Neuroscience (MCTN), Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany, [x] Department of Anaesthesiology, Intensive Care and Pain Medicine, University Hospital Muenster, Muenster, Germany

*Corresponding Author. Address: Pain Research, Dept. Surgery & Cancer, Faculty of Medicine, Imperial College, Chelsea & Westminster Hospital campus, 369 Fulham Road London, London SW10 9NH, United Kingdom. Tel.: +44 (0)20 3315 8816. E-mail address: d.hohenschurz-schmidt19@imperial.ac.uk (D. Hohenschurz-Schmidt).

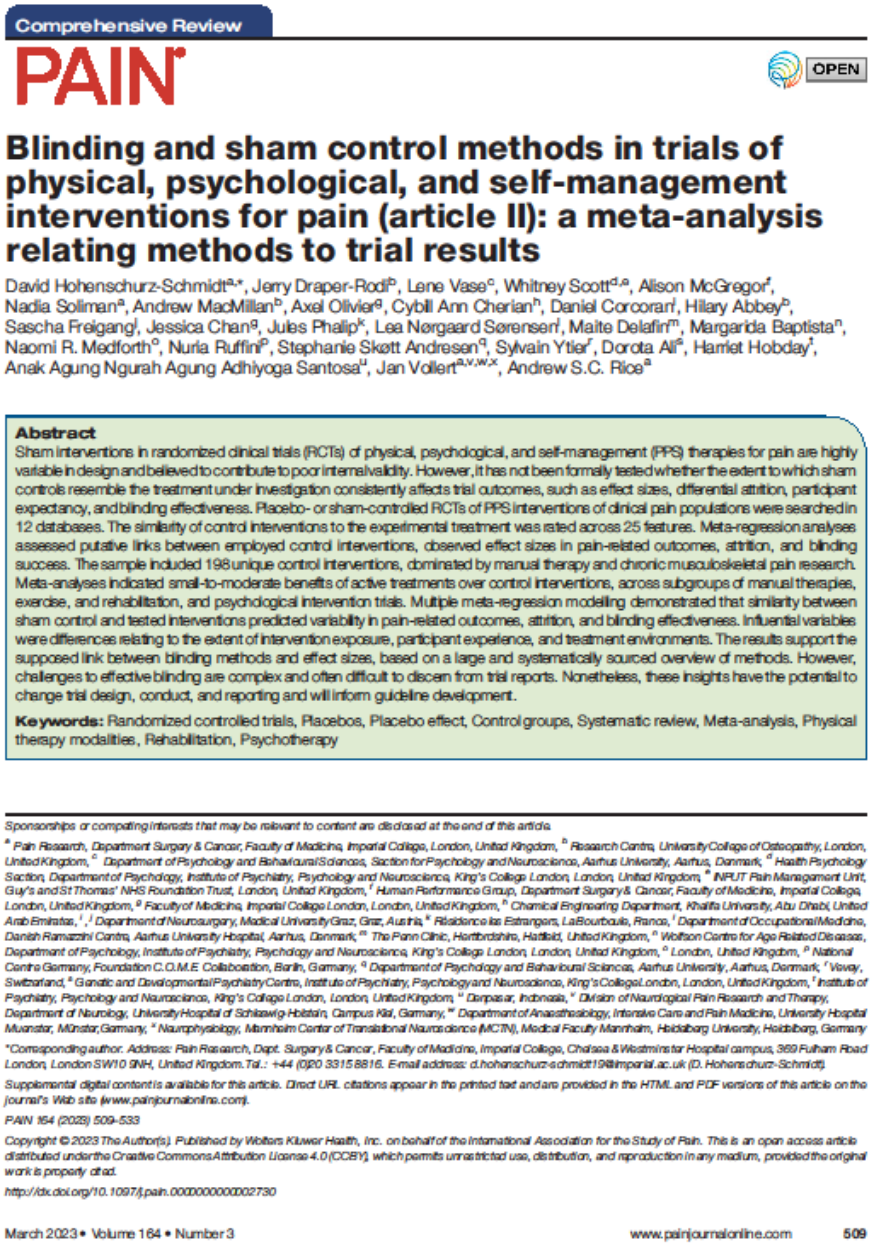
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 164 (2023) 469–484

Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Association for the Study of Pain. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

http://dx.doi.org/10.1097/j.pain.0000000000002723

3. Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., Scott, W., McGregor, A., Soliman, N., MacMillan, A., Olivier, A., Cherian, C.A., Corcoran, D., Abbey, H., Freigang, S., Chan, J., Phalip, J., Sørensen, L.N., Delafin, M., Baptista, M., Medforth, N.R., Ruffini, N., Andresen, S.S., Ytier, S., Ali, D., Hobday, H., Ngurah Agung Adhiyoga Santosa, A.A., Vollert, J., Rice, A.S.C., 2023. **Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article II): a meta-analysis relating methods to trial results.** PAIN 164, 509. https://doi.org/10.1097/j.pain.0000000000002730



Comprehensive Review

PAIN

OPEN

Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article II): a meta-analysis relating methods to trial results

David Hohenschurz-Schmidt[a,*], Jerry Draper-Rodi[b], Lene Vase[c], Whitney Scott[d,e], Alison McGregor[f], Nadia Soliman[a], Andrew MacMillan[b], Axel Olivier[g], Cybill Ann Cherian[h], Daniel Corcoran[i], Hilary Abbey[b], Sascha Freigang[j], Jessica Chan[g], Jules Phalip[k], Lea Nørgaard Sørensen[l], Malte Delafin[m], Margarida Baptista[n], Naomi R. Medforth[o], Nuria Ruffini[p], Stephanie Skøtt Andresen[q], Sylvain Ytier[r], Dorota Ali[s], Harriet Hobday[t], Anak Agung Ngurah Agung Adhiyoga Santosa[u], Jan Vollert[a,v,w,x], Andrew S.C. Rice[a]

Abstract

Sham interventions in randomized clinical trials (RCTs) of physical, psychological, and self-management (PPS) therapies for pain are highly variable in design and believed to contribute to poor internal validity. However, it has not been formally tested whether the extent to which sham controls resemble the treatment under investigation consistently affects trial outcomes, such as effect sizes, differential attrition, participant expectancy, and blinding effectiveness. Placebo- or sham-controlled RCTs of PPS interventions of clinical pain populations were searched in 12 databases. The similarity of control interventions to the experimental treatment was rated across 25 features. Meta-regression analyses assessed putative links between employed control interventions, observed effect sizes in pain-related outcomes, attrition, and blinding success. The sample included 198 unique control interventions, dominated by manual therapy and chronic musculoskeletal pain research. Meta-analyses indicated small-to-moderate benefits of active treatments over control interventions, across subgroups of manual therapies, exercise, and rehabilitation, and psychological intervention trials. Multiple meta-regression modelling demonstrated that similarity between sham control and tested interventions predicted variability in pain-related outcomes, attrition, and blinding effectiveness. Influential variables were differences relating to the extent of intervention exposure, participant experience, and treatment environments. The results support the supposed link between blinding methods and effect sizes, based on a large and systematically sourced overview of methods. However, challenges to effective blinding are complex and often difficult to discern from trial reports. Nonetheless, these insights have the potential to change trial design, conduct, and reporting and will inform guideline development.

Keywords: Randomized controlled trials, Placebos, Placebo effect, Control groups, Systematic review, Meta-analysis, Physical therapy modalities, Rehabilitation, Psychotherapy

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Pain Research, Department Surgery & Cancer, Faculty of Medicine, Imperial College, London, United Kingdom,* [b] *Research Centre, University College of Osteopathy, London, United Kingdom,* [c] *Department of Psychology and Behavioural Sciences, Section for Psychology and Neuroscience, Aarhus University, Aarhus, Denmark,* [d] *Health Psychology Section, Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom,* [e] *INPUT Pain Management Unit, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom,* [f] *Human Performance Group, Department Surgery & Cancer, Faculty of Medicine, Imperial College, London, United Kingdom,* [g] *Faculty of Medicine, Imperial College London, London, United Kingdom,* [h] *Chemical Engineering Department, Khalifa University, Abu Dhabi, United Arab Emirates,* [i,j] *Department of Neurosurgery, Medical University Graz, Graz, Austria,* [k] *Résidence les Estrangers, La Bourboule, France,* [l] *Department of Occupational Medicine, Danish Ramazzini Centre, Aarhus University Hospital, Aarhus, Denmark,* [m] *The Penn Clinic, Hertfordshire, Hatfield, United Kingdom,* [n] *Wolfson Centre for Age Related Diseases, Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom,* [o] *London, United Kingdom,* [p] *National Centre Germany, Foundation C.O.M.E Collaboration, Berlin, Germany,* [q] *Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark,* [r] *Vevey, Switzerland,* [s] *Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom,* [t] *Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom,* [u] *Denpasar, Indonesia,* [v] *Division of Neurological Pain Research and Therapy, Department of Neurology, University Hospital of Schleswig-Holstein, Campus Kiel, Germany,* [w] *Department of Anaesthesiology, Intensive Care and Pain Medicine, University Hospital Muenster, Münster, Germany,* [x] *Neurophysiology, Mannheim Center of Translational Neuroscience (MCTN), Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany*

**Corresponding author. Address: Pain Research, Dept. Surgery & Cancer, Faculty of Medicine, Imperial College, Chelsea & Westminster Hospital campus, 369 Fulham Road London, London SW10 9NH, United Kingdom. Tel.: +44 (0)20 3315 8816. E-mail address: d.hohenschurz-schmidt19@imperial.ac.uk (D. Hohenschurz-Schmidt).*

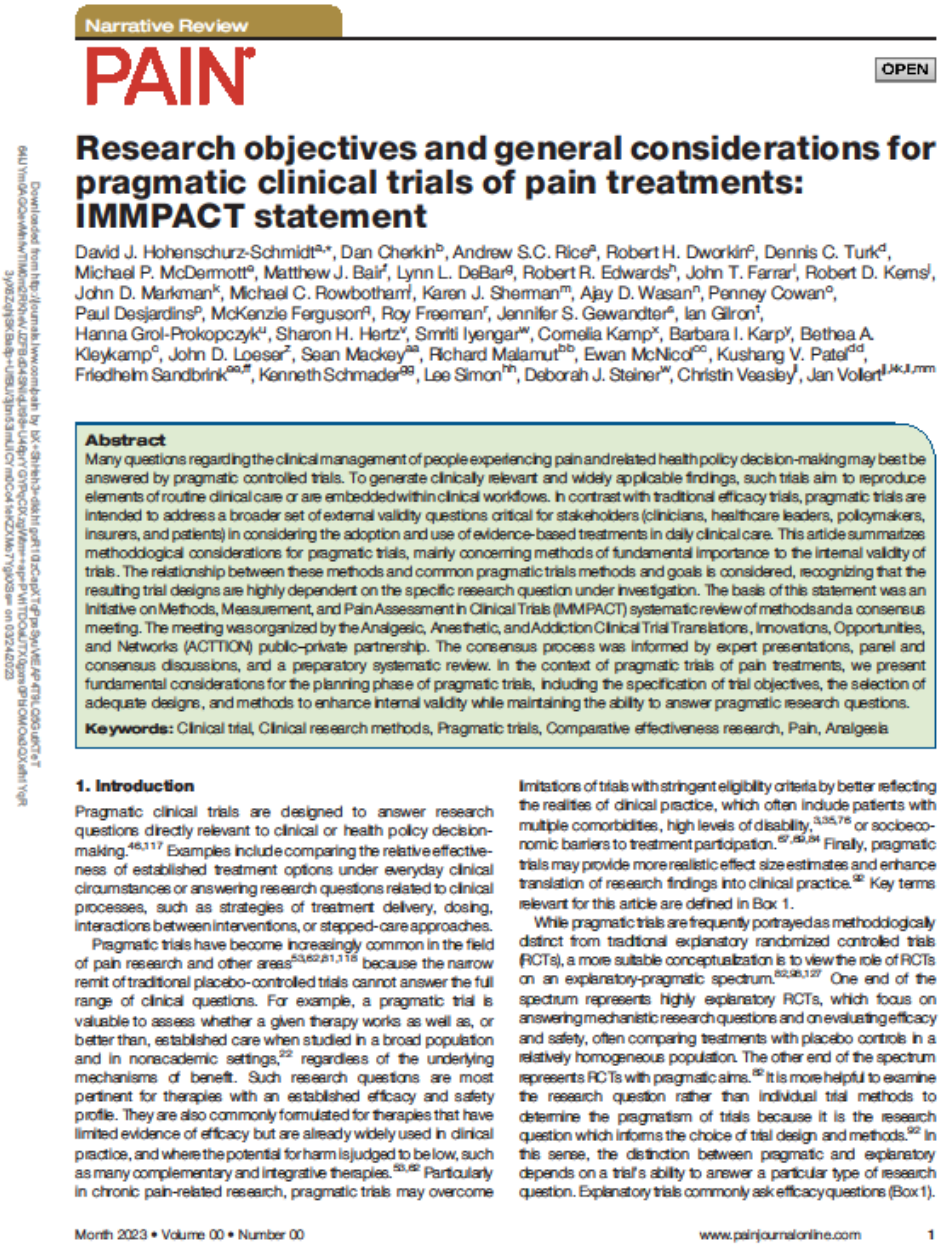
Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 164 (2023) 509–533

Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the International Association for the Study of Pain. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

http://dx.doi.org/10.1097/j.pain.0000000000002730

4. Hohenschurz-Schmidt, D., Cherkin, D., Rice, A.S.C., Dworkin, R.H., Turk, D.C., McDermott, M.P., Bair, M.J., DeBar, L.L., Edwards, R.R., Farrar, J.T., Kerns, R.D., Markman, J.D., Rowbotham, M.C., Sherman, K.J., Wasan, A.D., Cowan, P., Desjardins, P., Ferguson, M., Freeman, R., Gewandter, J.S., Gilron, I., Grol-Prokopczyk, H., Hertz, S.H., Iyengar, S., Kamp, C., Karp, B.I., Kleykamp, B.A., Loeser, J.D., Mackey, S., Malamut, R., McNicol, E., Patel, K.V., Sandbrink, F., Schmader, K., Simon, L., Steiner, D.J., Veasley, C., Vollert, J., 2023. **Research objectives and general considerations for pragmatic clinical trials of pain treatments: IMMPACT statement.** PAIN 10.1097/j.pain.0000000000002888. https://doi.org/10.1097/j.pain.0000000000002888



Narrative Review

PAIN

OPEN

Research objectives and general considerations for pragmatic clinical trials of pain treatments: IMMPACT statement

David J. Hohenschurz-Schmidt[a,*], Dan Cherkin[b], Andrew S.C. Rice[a], Robert H. Dworkin[c], Dennis C. Turk[d], Michael P. McDermott[e], Matthew J. Bair[f], Lynn L. DeBar[g], Robert R. Edwards[h], John T. Farrar[i], Robert D. Kerns[j], John D. Markman[k], Michael C. Rowbotham[l], Karen J. Sherman[m], Ajay D. Wasan[n], Penney Cowan[o], Paul Desjardins[p], McKenzie Ferguson[q], Roy Freeman[r], Jennifer S. Gewandter[s], Ian Gilron[t], Hanna Grol-Prokopczyk[u], Sharon H. Hertz[v], Smriti Iyengar[w], Cornelia Kamp[x], Barbara I. Karp[y], Bethea A. Kleykamp[c], John D. Loeser[z], Sean Mackey[aa], Richard Malamut[bb], Ewan McNicol[cc], Kushang V. Patel[dd], Friedhelm Sandbrink[ee,ff], Kenneth Schmader[gg], Lee Simon[hh], Deborah J. Steiner[w], Christin Veasley[i], Jan Vollert[ij,kk,ll,mm]

Abstract

Many questions regarding the clinical management of people experiencing pain and related health policy decision-making may best be answered by pragmatic controlled trials. To generate clinically relevant and widely applicable findings, such trials aim to reproduce elements of routine clinical care or are embedded within clinical workflows. In contrast with traditional efficacy trials, pragmatic trials are intended to address a broader set of external validity questions critical for stakeholders (clinicians, healthcare leaders, policymakers, insurers, and patients) in considering the adoption and use of evidence-based treatments in daily clinical care. This article summarizes methodological considerations for pragmatic trials, mainly concerning methods of fundamental importance to the internal validity of trials. The relationship between these methods and common pragmatic trials methods and goals is considered, recognizing that the resulting trial designs are highly dependent on the specific research question under investigation. The basis of this statement was an Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) systematic review of methods and a consensus meeting. The meeting was organized by the Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION) public–private partnership. The consensus process was informed by expert presentations, panel and consensus discussions, and a preparatory systematic review. In the context of pragmatic trials of pain treatments, we present fundamental considerations for the planning phase of pragmatic trials, including the specification of trial objectives, the selection of adequate designs, and methods to enhance internal validity while maintaining the ability to answer pragmatic research questions.

Keywords: Clinical trial, Clinical research methods, Pragmatic trials, Comparative effectiveness research, Pain, Analgesia

1. Introduction

Pragmatic clinical trials are designed to answer research questions directly relevant to clinical or health policy decision-making.[46,117] Examples include comparing the relative effectiveness of established treatment options under everyday clinical circumstances or answering research questions related to clinical processes, such as strategies of treatment delivery, dosing, interactions between interventions, or stepped-care approaches.

Pragmatic trials have become increasingly common in the field of pain research and other areas[53,62,81,118] because the narrow remit of traditional placebo-controlled trials cannot answer the full range of clinical questions. For example, a pragmatic trial is valuable to assess whether a given therapy works as well as, or better than, established care when studied in a broad population and in nonacademic settings,[22] regardless of the underlying mechanisms of benefit. Such research questions are most pertinent for therapies with an established efficacy and safety profile. They are also commonly formulated for therapies that have limited evidence of efficacy but are already widely used in clinical practice, and where the potential for harm is judged to be low, such as many complementary and integrative therapies.[53,62] Particularly in chronic pain-related research, pragmatic trials may overcome limitations of trials with stringent eligibility criteria by better reflecting the realities of clinical practice, which often include patients with multiple comorbidities, high levels of disability,[3,35,75] or socioeconomic barriers to treatment participation.[57,60,84] Finally, pragmatic trials may provide more realistic effect size estimates and enhance translation of research findings into clinical practice.[92] Key terms relevant for this article are defined in Box 1.

While pragmatic trials are frequently portrayed as methodologically distinct from traditional explanatory randomized controlled trials (RCTs), a more suitable conceptualization is to view the role of RCTs on an explanatory-pragmatic spectrum.[82,98,127] One end of the spectrum represents highly explanatory RCTs, which focus on answering mechanistic research questions and on evaluating efficacy and safety, often comparing treatments with placebo controls in a relatively homogeneous population. The other end of the spectrum represents RCTs with pragmatic aims.[82] It is more helpful to examine the research question rather than individual trial methods to determine the pragmatism of trials because it is the research question which informs the choice of trial design and methods.[92] In this sense, the distinction between pragmatic and explanatory depends on a trial's ability to answer a particular type of research question. Explanatory trials commonly ask efficacy questions (Box 1).

5. Hohenschurz-Schmidt, D., Vase, L., Scott, W., Annoni, M., Ajayi, O.K., Barth, J., Bennell, K., Berna, C., Bialosky, J., Braithwaite, F., Finnerup, N.B., Williams, A.C. de C., Carlino, E., Cerritelli, F., Chaibi, A., Cherkin, D., Colloca, L., Côté, P., Darnall, B.D., Evans, R., Fabre, L., Faria, V., French, S., Gerger, H., Häuser, W., Hinman, R.S., Ho, D., Janssens, T., Jensen, K., Johnston, C., Lunde, S.J., Keefe, F., Kerns, R.D., Koechlin, H., Kongsted, A., Michener, L.A., Moerman, D.E., Musial, F., Newell, D., Nicholas, M., Palermo, T.M., Palermo, S., Peerdeman, K.J., Pogatzki-Zahn, E.M., Puhl, A.A., Roberts, L., Rossettini, G., Matthiesen, S.T., Underwood, M., Vaucher, P., Vollert, J., Wartolowska, K., Weimer, K., Werner, C.P., Rice, A.S.C., Draper-Rodi, J., 2023. **Recommendations for the development, implementation, and reporting of control interventions in efficacy and mechanistic trials of physical, psychological, and self-management therapies: the CoPPS Statement.** BMJ 381, e072108. https://doi.org/10.1136/bmj-2022-072108



RESEARCH METHODS AND REPORTING

Check for updates

Recommendations for the development, implementation, and reporting of control interventions in efficacy and mechanistic trials of physical, psychological, and self-management therapies: the CoPPS Statement

For numbered affiliations see end of the article

Correspondence to:
D Hohenschurz-Schmidt
d.hohenschurz@imperial.ac.uk
(ORCID 0000-0002-1964-6069)

Additional material is published online only. To view please visit the journal online.

Cite this as: BMJ 2023;381:e072108
http://dx.doi.org/10.1136/bmj-2022-072108

Accepted: 05 April 2023

David Hohenschurz-Schmidt,[1,2] Lene Vase,[3] Whitney Scott,[4] Marco Annoni,[5] Oluwafemi K Ajayi,[6] Jürgen Barth,[7] Kim Bennell,[8] Chantal Berna,[9] Joel Bialosky,[10] Felicity Braithwaite,[11] Nanna B Finnerup,[12] Amanda C de C Williams,[13] Elisa Carlino,[14] Francesco Cerritelli,[15] Aleksander Chaibi,[16] Dan Cherkin,[17] Luana Colloca,[18] Pierre Côté,[19] Beth D Darnall,[20] Roni Evans,[21] Laurent Fabre,[22] Vanda Faria,[23] Simon French,[24] Heike Gerger,[25] Winfried Häuser,[26] Rana S Hinman,[27] Dien Ho,[28] Thomas Janssens,[29] Karin Jensen,[30] Chris Johnston,[31] Sigrid Juhl Lunde,[3] Francis Keefe,[32] Robert D Kerns,[33] Helen Koechlin,[34] Alice Kongsted,[35] Lori A Michener,[36] Daniel E Moerman,[37] Frauke Musial,[38] David Newell,[39] Michael Nicholas,[40] Tonya M Palermo,[41] Sara Palermo,[42] Kaya J Peerdeman,[43] Esther M Pogatzki-Zahn,[44] Aaron A Puhl,[45] Lisa Roberts,[46] Giacomo Rossettini,[47] Susan Tomczak Matthiesen,[3] Martin Underwood,[48] Paul Vaucher,[49] Jan Vollert,[50] Karolina Wartolowska,[51] Katja Weimer,[52] Christoph Patrick Werner,[53] Andrew S C Rice,[1] Jerry Draper-Rodi[2,54]

Control interventions (often called "sham," "placebo," or "attention controls") are essential for studying the efficacy or mechanism of physical, psychological, and self-management interventions in clinical trials. This article presents core recommendations for designing, conducting, and reporting control interventions to establish a quality standard in non-pharmacological intervention research. A framework of additional considerations supports researchers' decision making in this context. We also provide a reporting checklist for control interventions to enhance research transparency, usefulness, and rigour.

SUMMARY POINTS

The design, conduct, and reporting of control interventions are fundamental for interpreting efficacy and mechanistic clinical trials of physical, psychological, and self-management interventions

We developed this guideline using a three-round Delphi study with 64 experts in placebo research and/or clinical trials. This panel was presented with a systematic review of control and blinding methods and a meta-analysis that linked design features to trial endpoints. Consensus items informed interviews with eight patient partners. Finally, 44 Delphi participants and two patient representatives discussed the draft manuscript and results from stakeholder interviews at online consensus meetings

For clinical trials of treatment efficacy or mechanisms, we recommend designing control interventions that are as similar as possible to the tested intervention, apart from the components examined by the study

Structured planning, early stakeholder engagement, feasibility work, and piloting will improve the quality and acceptability of control interventions

When participant blinding is an objective, blinding effectiveness should be routinely assessed and reported

Detailed and transparent reporting will improve the interpretation and repeatability of clinical trials

Randomised controlled trials (RCTs) that test the efficacy of a drug against a "placebo control" are well established in drug research. In principle, placebo controls mimic the treatment under investigation but lack its supposed active ingredient (for example, as pharmacologically inert sugar pills or saline injections).[1] Such specifically designed control interventions control for the natural course of the disease and regression to the mean. Importantly, these interventions also account for context-dependent effects, such as those produced by patient-provider interactions and treatment related rituals. Researchers can use such interventions to blind trial participants, treatment providers, and outcome assessors to trial participants' group allocation, concealing whether the investigational treatment or the control is administered. Through the indistinguishability of drug and control, all trial participants are exposed to similar treatment contexts, which should lead to comparable expectations of treatment benefit.

In non-pharmacological trials, methodological criticism of control interventions (often called "sham" controls, representing the equivalent of placebo controls) contributes to concerns about research quality.[2-7] Ultimately, these concerns impede adequate consideration of many therapies for clinical guidelines or reimbursement. While some of these concerns are

Further publications during the PhD project

6. Wartolowska, K.A., Hohenschurz-Schmidt, D., Vase, L., Aronson, J.K., 2022. The importance of using placebo controls in non-pharmacological randomised trials. PAIN 10.1097/j.pain.0000000000002839. https://doi.org/10.1097/j.pain.0000000000002839
7. Smith, S., Normahani, P., Lane, T., Hohenschurz-Schmidt, D., Oliver, N., Davies, A.H., 2022. Prevention and Management Strategies for Diabetic Neuropathy. Life 12, 1185. https://doi.org/10.3390/life12081185
8. Smith, S., Normahani, P., Lane, T., Hohenschurz-Schmidt, D., Oliver, N., Davies, A.H., 2022. Pathogenesis of Distal Symmetrical Polyneuropathy in Diabetes. Life 12, 1074. https://doi.org/10.3390/life12071074
9. MacMillan, A., Hohenschurz-Schmidt, D., Migliarini, D.V., Draper-Rodi, D.J., 2022. Discrimination, bullying or harassment in undergraduate education in the osteopathic, chiropractic and physiotherapy professions: A systematic review with critical interpretive synthesis. International Journal of Educational Research Open 3, 100105. https://doi.org/10.1016/j.ijedro.2021.100105
10. Hohenschurz-Schmidt, D., Thomson, O.P., Rossettini, G., Miciak, M., Newell, D., Roberts, L., Vase, L., Draper-Rodi, J., 2022. Avoiding nocebo and other undesirable effects in chiropractic, osteopathy and physiotherapy: An invitation to reflect. Musculoskeletal Science & Practice 0. https://doi.org/10.1016/j.msksp.2022.102677
11. Hohenschurz-Schmidt, D., Scott, W., Park, C., Christopoulos, G., Vogel, S., Draper-Rodi, J., 2022. Telemedizinisches Schmerzmanagement von muskuloskeletalem Schmerz. Schmerz. https://doi.org/10.1007/s00482-022-00659-5
12. Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., 2022. Dissimilar Control Interventions in Clinical Trials Undermine Interpretability. JAMA Psychiatry 79, 271–272. https://doi.org/10.1001/jamapsychiatry.2021.3963
13. Draper-Rodi, J., Vaucher, P., Hohenschurz-Schmidt, D., Morin, C., Thomson, O.P., 2022. 4 M's to make sense of evidence – Avoiding the propagation of mistakes, misinterpretation, misrepresentation and misinformation. International Journal of Osteopathic Medicine 44, 29–35. https://doi.org/10.1016/j.ijosm.2022.04.007
14. Bini, P., Hohenschurz-Schmidt, D., Masullo, V., Pitt, D., Draper-Rodi, J., 2022. The effectiveness of manual and exercise therapy on headache intensity and frequency among patients with cervicogenic headache: a systematic review and meta-analysis. Chiropr Man Therap 30, 49. https://doi.org/10.1186/s12998-022-00459-9
15. Thomson, O.P., MacMillan, A., Draper-Rodi, J., Vaucher, P., Ménard, M., Vaughan, B., Morin, C., Alvarez, G., Sampath, K.K., Cerritelli, F., Shaw, R., Cymet, T.C., Bright, P., Hohenschurz-Schmidt, D., Vogel, S., 2021. Opposing vaccine hesitancy during the COVID-19 pandemic - A critical commentary and united statement of an international osteopathic research community. International Journal of Osteopathic Medicine 39, A1–A6. https://doi.org/10.1016/j.ijosm.2021.02.002
16. Hohenschurz-Schmidt, D., Vollert, J., Vogel, S., Rice, A.S.C., Draper-Rodi, J., 2021. Performing and interpreting randomized clinical trials. Journal of Osteopathic Medicine 121, 443–445. https://doi.org/10.1515/jom-2020-0320

17. Evans, M.C., Wade, C., Hohenschurz-Schmidt, D., Lally, P., Ugwudike, A., Shah, K., Bangerter, N., Sharp, D.J., Rice, A.S.C., 2021. Magnetic Resonance Imaging as a Biomarker in Diabetic and HIV-Associated Peripheral Neuropathy: A Systematic Review-Based Narrative. Frontiers in Neuroscience 15, 1169. https://doi.org/10.3389/fnins.2021.727311
18. Hohenschurz-Schmidt, D., Scott, W., Park, C., Christopoulos, G., Vogel, S., Draper-Rodi, J., 2020. Remote management of musculoskeletal pain: a pragmatic approach to the implementation of video and phone consultations in musculoskeletal practice. PAIN Reports 5, e878. https://doi.org/10.1097/PR9.0000000000000878

References

Abraha, I., Montedori, A., 2010. Modified intention to treat reporting in randomised controlled trials: systematic review. BMJ 340, c2697. https://doi.org/10.1136/bmj.c2697

Adams, A.S., Schmittdiel, J.A., Altschuler, A., Bayliss, E.A., Neugebauer, R., Ma, L., Dyer, W., Clark, J., Cook, B., Willyoung, D., Jaffe, M., Young, J.D., Kim, E., Boggs, J.M., Prosser, L.A., Wittenberg, E., Callaghan, B., Shainline, M., Hippler, R.M., Grant, R.W., 2019. Automated symptom and treatment side effect monitoring for improved quality of life among adults with diabetic peripheral neuropathy in primary care: a pragmatic, cluster, randomized, controlled trial. Diabetic Medicine 36, 52–61. https://doi.org/10.1111/dme.13840

Aghabati, N., Mohammadi, E., Pour Esmaiel, Z., 2010. The effect of therapeutic touch on pain and fatigue of cancer patients undergoing chemotherapy. Evidence-based complementary and alternative medicine : eCAM 7, 375–81. https://doi.org/10.1093/ecam/nen006

Ahmad, I., Verma, S., Noohu, M.M., Shareef, Mohd.Y., Hussain, M.E., 2020. Sensorimotor and gait training improves proprioception, nerve function, and muscular activation in patients with diabetic peripheral neuropathy: a randomized control trial. J Musculoskelet Neuronal Interact 20, 234–248.

Ajimsha, M.S., Binsu, D., Chithra, S., 2014. Effectiveness of myofascial release in the management of plantar heel pain: a randomized controlled trial. Foot (Edinburgh, Scotland) 24, 66–71. https://doi.org/10.1016/j.foot.2014.03.005

Akers, J., Aguiar-Ibáñez, R., Baba-Akbari, A., 2009. Systematic reviews: CRD's guidance for undertaking reviews in health care. York, UK: Centre for Reviews and Dissemination, University of York.

Al Dahdah, M., 2022. Digital markets and the commercialization of healthcare in Africa: the case of Kenya. Globalizations 0, 1–13. https://doi.org/10.1080/14747731.2022.2135422

Albert, H.B., Manniche, C., 2012. The efficacy of systematic active conservative treatment for patients with severe sciatica: a single-blind, randomized, clinical, controlled trial. Spine 37, 531–42. https://doi.org/10.1097/BRS.0b013e31821ace7f

Ali, J., Antonelli, M., Bastian, L., Becker, W., Brandt, C.A., Burgess, D.J., Burns, A., Cohen, S.P., Davis, A.F., Dearth, C.L., Dziura, J., Edwards, R., Erdos, J., Farrokhi, S., Fritz, J., Geda, M., George, S.Z., Goertz, C., Goodie, J., Hastings, S.N., Heapy, A., Ilfeld, B.M., Katsovich, L., Kerns, R.D., Kyriakides, T.C., Lee, A., Long, C.R., Luther, S.L., Martino, S., Matheny, M.E., McGeary, D., Midboe, A., Pasquina, P., Peduzzi, P., Raffanello, M., Rhon, D., Rosen, M., Esposito, E.R., Scarton, D., Hastings, S.N., Seal, K., Silliker, N., Taylor, S., Taylor, S.L., Tsui, M., Wright, F.S., Zeliadt, S., 2022a. Optimizing the Impact of Pragmatic Clinical Trials for Veteran and Military Populations: Lessons From the Pain Management Collaboratory. Military Medicine 187, 179–185. https://doi.org/10.1093/milmed/usab458

Ali, J., Davis, A.F., Burgess, D.J., Rhon, D.I., Vining, R., Young-McCaughan, S., Green, S., Kerns, R.D., 2022b. Justice and equity in pragmatic clinical trials: Considerations for pain research within integrated health systems. Learning Health Systems 6, e10291. https://doi.org/10.1002/lrh2.10291

Allen, K.D., Oddone, E.Z., Coffman, C.J., Datta, S.K., Juntilla, K.A., Lindquist, J.H., Walker, T.A., Weinberger, M., Bosworth, H.B., 2010. Telephone-Based Self-management of Osteoarthritis. Annals of Internal Medicine 153, 570–579. https://doi.org/10.7326/0003-4819-153-9-201011020-00006

Alvarez, G., Cerritelli, F., Urrutia, G., 2016. Using the template for intervention description and replication (TIDieR) as a tool for improving the design and reporting of manual therapy interventions. Manual Therapy 24, 85–89. https://doi.org/10.1016/j.math.2016.03.004

Alvarez, G., Núñez-Cortés, R., Solà, I., Sitjà-Rabert, M., Fort-Vanmeerhaeghe, A., Fernández, C., Bonfill, X., Urrútia, G., 2021. Sample size, study length, and inadequate controls were the most common self-acknowledged limitations in manual therapy trials: A methodological review. Journal of Clinical Epidemiology 130, 96–106. https://doi.org/10.1016/j.jclinepi.2020.10.018

Alvarez, G., Solà, I., Sitjà-Rabert, M., Fort-Vanmeerhaeghe, A., Gich, I., Fernández, C., Bonfill, X., Urrútia, G., 2020. A methodological review revealed that reporting of trials in manual therapy has not improved over time. Journal of Clinical Epidemiology 121, 32–44. https://doi.org/10.1016/j.jclinepi.2020.01.006

Alvesson, M., Sköldberg, K., 2017. Reflexive Methodology: New Vistas for Qualitative Research. SAGE.

Amato Nesbit, S., Sharma, R., Waldfogel, J.M., Zhang, A., Bennett, W.L., Yeh, H.-C., Chelladurai, Y., Feldman, D., Robinson, K.A., Dy, S.M., 2019. Non-pharmacologic treatments for symptoms of diabetic peripheral neuropathy: a systematic review. Curr Med Res Opin 35, 15–25. https://doi.org/10.1080/03007995.2018.1497958

Amatya, B., Young, J., Khan, F., 2018. Non-pharmacological interventions for chronic pain in multiple sclerosis. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.CD012622.pub2

Anand, R., Norrie, J., Bradley, J.M., McAuley, D.F., Clarke, M., 2020. Fool's gold? Why blinded trials are not always best. BMJ 368, l6228. https://doi.org/10.1136/bmj.l6228

Andersen, F., Anjum, R.L., Rocca, E., 2019. Philosophical bias is the one bias that science cannot avoid. eLife 8, e44929. https://doi.org/10.7554/eLife.44929

Andersen Hammond, E., Pitz, M., Steinfeld, K., Lambert, P., Shay, B., 2020. An Exploratory Randomized Trial of Physical Therapy for the Treatment of Chemotherapy-Induced Peripheral Neuropathy. Neurorehabil Neural Repair 34, 235–246. https://doi.org/10.1177/1545968319899918

Andreae, S.J., Andreae, L.J., Richman, J.S., Cherrington, A.L., Safford, M.M., 2020. Peer-Delivered Cognitive Behavioral Training to Improve Functioning in Patients With Diabetes: A Cluster-Randomized Trial. The Annals of Family Medicine 18, 15–23. https://doi.org/10.1370/afm.2469

Annoni, M., 2018. The Ethics of Placebo Effects in Clinical Practice and Research. Int Rev Neurobiol 139, 463–484. https://doi.org/10.1016/bs.irn.2018.07.031

Antunes, R.P., Sales, C.M.D., Elliott, R., 2020. The clinical utility of the Personal Questionnaire (PQ): a mixed methods study. Counselling Psychology Quarterly 33, 25–45. https://doi.org/10.1080/09515070.2018.1439451

Arcos-Carmona, I.M., Castro-Sanchez, A.M., Mataran-Penarrocha, G.A., Gutierrez-Rubio, A.B., Ramos-Gonzalez, E., Moreno-Lorenzo, C., 2011. [Effects of aerobic exercise program and

relaxation techniques on anxiety, quality of sleep, depression, and quality of life in patients with fibromyalgia: a randomized controlled trial]. Efectos de un programa de ejercicios aerobicos y tecnicas de relajacion sobre el estado de ansiedad, calidad del sueno, depresion y calidad de vida en pacientes con fibromialgia: ensayo clinico aleatorizado. 137, 398–401. https://doi.org/10.1016/j.medcli.2010.09.045

Armijo-Olivo, S., Dennett, L., Arienti, C., Dahchi, M., Arokoski, J., Heinemann, A.W., Malmivaara, A., 2020. Blinding in Rehabilitation Research: Empirical Evidence on the Association Between Blinding and Treatment Effect Estimates. American Journal of Physical Medicine & Rehabilitation 99, 198–209. https://doi.org/10.1097/PHM.0000000000001377

Armijo-Olivo, S., Fuentes, J., da Costa, B.R., Saltaji, H., Ha, C., Cummings, G.G., 2017. Blinding in Physical Therapy Trials and Its Association with Treatment Effects: A Meta-epidemiological Study. American Journal of Physical Medicine & Rehabilitation 96, 34–44. https://doi.org/10.1097/PHM.0000000000000521

Ashar, Y.K., Gordon, A., Schubiner, H., Uipi, C., Knight, K., Anderson, Z., Carlisle, J., Polisky, L., Geuter, S., Flood, T.F., Kragel, P.A., Dimidjian, S., Lumley, M.A., Wager, T.D., 2021. Effect of Pain Reprocessing Therapy vs Placebo and Usual Care for Patients With Chronic Back Pain: A Randomized Clinical Trial. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2021.2669

Ashar, Y.K., Wager, T.D., 2022. Dissimilar Control Interventions in Clinical Trials Undermine Interpretability—Reply. JAMA Psychiatry. https://doi.org/10.1001/jamapsychiatry.2021.3966

Australian Government Department of Health, 2021. The National Strategic Action Plan for Pain Management [WWW Document]. Australian Government Department of Health and Aged Care. URL https://www.health.gov.au/resources/publications/the-national-strategic-action-plan-for-pain-management?language=en (accessed 12.15.22).

Aycock, D.M., Hayat, M.J., Helvig, A., Dunbar, S.B., Clark, P.C., 2018. Essential considerations in developing attention control groups in behavioral research. Res Nurs Health 41, 320–328. https://doi.org/10.1002/nur.21870

Bagley, H.J., Short, H., Harman, N.L., Hickey, H.R., Gamble, C.L., Woolfall, K., Young, B., Williamson, P.R., 2016. A patient and public involvement (PPI) toolkit for meaningful and flexible involvement in clinical trials – a work in progress. Research Involvement and Engagement 2, 15. https://doi.org/10.1186/s40900-016-0029-8

Bai, Y., Ma, J.-H., Yu, Y., Wang, Z.-W., 2022. Effect of Cognitive-Behavioral Therapy or Mindfulness Therapy on Pain and Quality of Life in Patients with Diabetic Neuropathy: A Systematic Review and Meta-Analysis. Pain Management Nursing 23, 861–870. https://doi.org/10.1016/j.pmn.2022.05.005

Balk, E.M., Earley, A., Raman, G., Avendano, E.A., Pittas, A.G., Remington, P.L., 2015. Combined Diet and Physical Activity Promotion Programs to Prevent Type 2 Diabetes Among Persons at Increased Risk: A Systematic Review for the Community Preventive Services Task Force. Ann Intern Med 163, 437–451. https://doi.org/10.7326/M15-0452

Ball, E.F., Nur Shafina Muhammad Sharizan, E., Franklin, G., Rogozińska, E., 2017. Does mindfulness meditation improve chronic pain? A systematic review. Current Opinion in Obstetrics and Gynecology 29, 359–366. https://doi.org/10.1097/GCO.0000000000000417

Bang, H., Flaherty, S.P., Kolahi, J., Park, J., 2010. Blinding assessment in clinical trials: A review of statistical methods and a proposal of blinding assessment protocol. Clinical Research and Regulatory Affairs 27, 42–51. https://doi.org/10.3109/10601331003777444

Bang, H., Ni, L., Davis, C.E., 2004. Assessment of blinding in clinical trials. Controlled Clinical Trials 25, 143–156. https://doi.org/10.1016/j.cct.2003.10.016

Barnett, K., Mercer, S.W., Norbury, M., Watt, G., Wyke, S., Guthrie, B., 2012. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 380, 37–43. https://doi.org/10.1016/S0140-6736(12)60240-2

Barth, J., Kern, A., Lüthi, S., Witt, C.M., 2019. Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS). BMJ Open 9, e026712. https://doi.org/10.1136/bmjopen-2018-026712

Baskerville, R., Ricci-Cabello, I., Roberts, N., Farmer, A., 2017. Impact of accelerometer and pedometer use on physical activity and glycaemic control in people with Type 2 diabetes: a systematic review and meta-analysis. Diabet Med 34, 612–620. https://doi.org/10.1111/dme.13331

Baskin, T.W., Tierney, S.C., Minami, T., Wampold, B.E., 2003. Establishing Specificity in Psychotherapy: A Meta-Analysis of Structural Equivalence of Placebo Controls. Journal of Consulting 71, 973–979.

Bastian, L.A., Cohen, S.P., Katsovich, L., Becker, W.C., Brummett, B.R., Burgess, D.J., Crunkhorn, A.E., Denneson, L.M., Frank, J.W., Goertz, C., Ilfeld, B., Kanzler, K.E., Krishnaswamy, A., LaChappelle, K., Martino, S., Mattocks, K., McGeary, C.A., Reznik, T.E., Rhon, D.I., Salsbury, S.A., Seal, K.H., Semiatin, A.M., Shin, M.H., Simon, C.B., Teyhen, D.S., Zamora, K., Kerns, Robert D., the NIH-DOD-VA Pain Management Collaboratory, 2020. Stakeholder Engagement in Pragmatic Clinical Trials: Emphasizing Relationships to Improve Pain Management Delivery and Outcomes. Pain Medicine 21, S13–S20. https://doi.org/10.1093/pm/pnaa333

Bautista, R., 2012. An Overlooked Approach in Survey Research: Total Survey Error, in: Gideon, L. (Ed.), Handbook of Survey Methodology for the Social Sciences. Springer, New York, NY, pp. 37–49. https://doi.org/10.1007/978-1-4614-3876-2_4

Beall, J., 2016. Dangerous Predatory Publishers Threaten Medical Research. J Korean Med Sci 31, 1511–1513. https://doi.org/10.3346/jkms.2016.31.10.1511

Beard, D.J., Campbell, M.K., Blazeby, J.M., Carr, A.J., Weijer, C., Cuthbertson, B.H., Buchbinder, R., Pinkney, T., Bishop, F.L., Pugh, J., Cousins, S., Harris, I.A., Lohmander, L.S., Blencowe, N., Gillies, K., Probst, P., Brennan, C., Cook, A., Farrar-Hockley, D., Savulescu, J., Huxtable, R., Rangan, A., Tracey, I., Brocklehurst, P., Ferreira, M.L., Nicholl, J., Reeves, B.C., Hamdy, F., Rowley, S.C., Cook, J.A., 2020. Considerations and methods for placebo controls in surgical trials (ASPIRE guidelines). The Lancet 395, 828–838. https://doi.org/10.1016/S0140-6736(19)33137-X

Beard, D.J., Rees, J.L., Cook, J.A., Rombach, I., Cooper, C., Merritt, N., Shirkey, B.A., Donovan, J.L., Gwilym, S., Savulescu, J., Moser, J., Gray, A., Jepson, M., Tracey, I., Judge, A., Wartolowska, K., Carr, A.J., CSAW Study Group, 2018. Arthroscopic subacromial decompression for subacromial shoulder pain (CSAW): a multicentre, pragmatic, parallel group, placebo-controlled, three-group, randomised surgical trial. Lancet 391, 329–338. https://doi.org/10.1016/S0140-6736(17)32457-1

Becker, W.C., DeBar, L.L., Heapy, A.A., Higgins, D., Krein, S.L., Lisi, A., Makris, U.E., Allen, K.D., 2018. A Research Agenda for Advancing Non-pharmacological Management of Chronic Musculoskeletal Pain: Findings from a VHA State-of-the-art Conference. J Gen Intern Med 33, 11–15. https://doi.org/10.1007/s11606-018-4345-6

Beedie, C., Benedetti, F., Barbiani, D., Camerone, E., Cohen, E., Coleman, D., Davis, A., Elsworth-Edelsten, C., Flowers, E., Foad, A., Harvey, S., Hettinga, F., Hurst, P., Lane, A., Lindheimer, J., Raglin, J., Roelands, B., Schiphof-Godart, L., Szabo, A., 2018. Consensus statement on placebo effects in sports and exercise: The need for conceptual clarity, methodological rigour, and the elucidation of neurobiological mechanisms. European Journal of Sport Science 18, 1383–1389. https://doi.org/10.1080/17461391.2018.1496144

Beedie, C., Foad, A., Hurst, P., 2015. Capitalizing on the Placebo Component of Treatments. Current Sports Medicine Reports 14, 284–287. https://doi.org/10.1249/JSR.0000000000000172

Beever, D., Swaby, L., 2019. An evaluation of risk-based monitoring in pragmatic trials in UK Clinical Trials Units. Trials 20, 556. https://doi.org/10.1186/s13063-019-3619-6

Benedetti, F., 2020. Placebo Effects: Understanding the Mechanisms in Health and Disease, 3rd ed. Oxford University Press, Oxford, England.

Benedetti, F., Frisaldi, E., Shaibani, A., 2022. Thirty Years of Neuroscientific Investigation of Placebo and Nocebo: The Interesting, the Good, and the Bad. Annual Review of Pharmacology and Toxicology 62, 323–340. https://doi.org/10.1146/annurev-pharmtox-052120-104536

Berdal, G., Bø, I., Dager, T.N., Dingsør, A., Eppeland, S.G., Hagfors, J., Hamnes, B., Mowinckel, P., Nielsen, M., Sand-Svartrud, A.-L., Slungaard, B., Wigers, S.H., Hagen, K.B., Dagfinrud, H.S., Kjeken, I., 2018. Structured Goal Planning and Supportive Telephone Follow-up in Rheumatology Care: Results From a Pragmatic, Stepped-Wedge, Cluster-Randomized Trial. Arthritis Care Res (Hoboken) 70, 1576–1586. https://doi.org/10.1002/acr.23520

Bernhardsson, S., Larsson, M.E.H., Johansson, K., Öberg, B., 2017. "In the physio we trust": A qualitative study on patients' preferences for physiotherapy. Physiotherapy Theory and Practice 33, 535–549. https://doi.org/10.1080/09593985.2017.1328720

Bernstein, I.A., Malik, Q., Carville, S., Ward, S., 2017. Low back pain and sciatica: summary of NICE guidance. BMJ 356, i6748. https://doi.org/10.1136/bmj.i6748

Bialosky, J.E., Beneciuk, J.M., Bishop, M.D., Coronado, R.A., Penza, C.W., Simon, C.B., George, S.Z., 2017a. Unraveling the Mechanisms of Manual Therapy: Modeling an Approach. J Orthop Sports Phys Ther 48, 8–18. https://doi.org/10.2519/jospt.2018.7476

Bialosky, J.E., Bishop, M.D., Penza, C.W., 2017b. Placebo Mechanisms of Manual Therapy: A Sheep in Wolf 's Clothing? J Orthop Sports Phys Ther 47, 301–304. https://doi.org/10.2519/jospt.2017.0604

Bialosky, J.E., Bishop, M.D., Price, D.D., Robinson, M.E., Vincent, K.R., George, S.Z., 2009. A randomized sham-controlled trial of a neurodynamic technique in the treatment of carpal tunnel syndrome. The Journal of orthopaedic and sports physical therapy 39, 709–23. https://doi.org/10.2519/jospt.2009.3117

Bialosky, J.E., George, S.Z., Horn, M.E., Price, D.D., Staud, R., Robinson, M.E., 2014. Spinal manipulative therapy-specific changes in pain sensitivity in individuals with low back pain

(NCT01168999). The journal of pain : official journal of the American Pain Society 15, 136–48. https://doi.org/10.1016/j.jpain.2013.10.005

Bini, P., Hohenschurz-Schmidt, D., Masullo, V., Pitt, D., Draper-Rodi, J., 2022. The effectiveness of manual and exercise therapy on headache intensity and frequency among patients with cervicogenic headache: a systematic review and meta-analysis. Chiropractic & Manual Therapies 30, 49. https://doi.org/10.1186/s12998-022-00459-9

Birch, S., Alraek, T., Kim, K.H., Lee, M.S., 2016. Placebo-Controlled Trials in Acupuncture: Problems and Solutions, in: Leung, S., Hu, H. (Eds.), Evidence-Based Research Methods for Chinese Medicine. Springer, Singapore, pp. 55–64. https://doi.org/10.1007/978-981-10-2290-6_4

Birch, S., Bovey, M., Robinson, N., 2020. Acupuncture for chronic primary pain – are UK guidelines now consistent with other countries? European Journal of Integrative Medicine 41, 101257. https://doi.org/10.1016/j.eujim.2020.101257

Bishop, F.L., Coghlan, B., Geraghty, A.W., Everitt, H., Little, P., Holmes, M.M., Seretis, D., Lewith, G., 2017. What techniques might be used to harness placebo effects in non-malignant pain? A literature review and survey to develop a taxonomy. BMJ Open 7, e015516. https://doi.org/10.1136/bmjopen-2016-015516

Bishop, M.D., Torres-Cueco, R., Gay, C.W., Lluch-Girbés, E., Beneciuk, J.M., Bialosky, J.E., 2015. What effect can manual therapy have on a patient's pain experience? Pain Management 5, 455–464. https://doi.org/10.2217/pmt.15.39

Blatch-Jones, A.J., Pek, W., Kirkpatrick, E., Ashton-Key, M., 2018. Role of feasibility and pilot studies in randomised controlled trials: a cross-sectional study. BMJ Open 8, e022233. https://doi.org/10.1136/bmjopen-2018-022233

Blazey, P., Crossley, K.M., Ardern, C.L., Middelkoop, M. van, Scott, A., Khan, K.M., 2022. It is time for consensus on 'consensus statements.' Br J Sports Med 56, 306–307. https://doi.org/10.1136/bjsports-2021-104578

Boddy, C.R., 2016. Sample size for qualitative research. Qualitative Market Research: An International Journal 19, 426–432. https://doi.org/10.1108/QMR-06-2016-0053

Bonney, R., Phillips, T.B., Ballard, H.L., Enck, J.W., 2016. Can citizen science enhance public understanding of science? Public Underst Sci 25, 2–16. https://doi.org/10.1177/0963662515607406

Bono Mira, M., 2008. EFFECTS OF THE MANIPULATION OF D6 IN THE GLUCEMIA OF DIABETIC PATIENTS [WWW Document].

Booth, A., Clarke, M., Dooley, G., Ghersi, D., Moher, D., Petticrew, M., Stewart, L., 2012. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Systematic Reviews 1, 2. https://doi.org/10.1186/2046-4053-1-2

Bordoni, B., 2019. The Benefits and Limitations of Evidence-based Practice in Osteopathy. Cureus 11. https://doi.org/10.7759/cureus.6093

Bordoni, B., Escher, A.R., 2023. The Osteopath's Imprint: Osteopathic Medicine Under the Nanoscopic Lens. Cureus 15, e33914. https://doi.org/10.7759/cureus.33914

Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R., 2021. Introduction to Meta-Analysis. John Wiley & Sons.

Bornhöft, G., Maxion-Bergemann, S., Wolf, U., Kienle, G.S., Michalsen, A., Vollmar, H.C., Gilbertson, S., Matthiessen, P.F., 2006. Checklist for the qualitative evaluation of clinical

studies with particular focus on external validity and model validity. BMC Medical Research Methodology 6, 56. https://doi.org/10.1186/1471-2288-6-56

Boutron, I., Altman, D.G., Moher, D., Schulz, K.F., Ravaud, P., for the CONSORT NPT Group, 2017. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. Ann Intern Med 167, 40. https://doi.org/10.7326/M17-0046

Boutron, I., Guittet, L., Estellat, C., Moher, D., Hróbjartsson, A., Ravaud, P., 2007. Reporting Methods of Blinding in Randomized Trials Assessing Nonpharmacological Treatments. PLOS Medicine 4, e61. https://doi.org/10.1371/journal.pmed.0040061

Boutron, I., Haneef, R., Yavchitz, A., Baron, G., Novack, J., Oransky, I., Schwitzer, G., Ravaud, P., 2019. Three randomized controlled trials evaluating the impact of "spin" in health news stories reporting studies of pharmacologic treatments on patients'/caregivers' interpretation of treatment benefit. BMC Medicine 17, 105. https://doi.org/10.1186/s12916-019-1330-9

Boutron, I., Moher, D., Altman, D.G., Schulz, K.F., Ravaud, P., for the CONSORT Group, 2008. Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. Ann Intern Med 148, 295. https://doi.org/10.7326/0003-4819-148-4-200802190-00008

Boutron, I., Ravaud, P., 2011. Blinding in Nonpharmacological Randomized Controlled Trials, in: Randomized Clinical Trials of Nonpharmacologic Treatments. CRC Press LLC, London, UNITED KINGDOM, pp. 3–8.

Boutron, I., Tubach, F., Giraudeau, B., Ravaud, P., 2004. Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. Journal of Clinical Epidemiology 57, 543–550. https://doi.org/10.1016/j.jclinepi.2003.12.010

Braithwaite, F.A., Walters, J.L., Li, L.S.K., Moseley, G.L., Williams, M.T., McEvoy, M.P., 2019. Blinding Strategies in Dry Needling Trials: Systematic Review and Meta-Analysis. Phys Ther 99, 1461–1480. https://doi.org/10.1093/ptj/pzz111

Braithwaite, F.A., Walters, J.L., Li, L.S.K., Moseley, G.L., Williams, M.T., McEvoy, M.P., 2018. Effectiveness and adequacy of blinding in the moderation of pain outcomes: Systematic review and meta-analyses of dry needling trials. Peerj 6, e5318. https://doi.org/10.7717/peerj.5318

Braithwaite, F.A., Walters, J.L., Moseley, G.L., Williams, M.T., McEvoy, M.P., 2020a. Towards more homogenous and rigorous methods in sham-controlled dry needling trials: two Delphi surveys. Physiotherapy 106, 12–23. https://doi.org/10.1016/j.physio.2019.11.004

Braithwaite, F.A., Walters, J.L., Moseley, G.L., Williams, M.T., McEvoy, M.P., 2020b. Towards more credible shams for physical interventions: A Delphi survey. Clinical Trials 17, 295–305. https://doi.org/10.1177/1740774520910365

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative Research in Psychology 3, 77–101. https://doi.org/10.1191/1478088706qp063oa

Brett, J., Staniszewska, S., Mockford, C., Herron-Marx, S., Hughes, J., Tysall, C., Suleman, R., 2014. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expectations 17, 637–650. https://doi.org/10.1111/j.1369-7625.2012.00795.x

Brienza, C., 2016. Degrees of (Self-)Exploitation: Learning to Labour in the Neoliberal University. Journal of Historical Sociology 29, 92–111. https://doi.org/10.1111/johs.12119

Briggs, A.M., Shiffman, J., Shawar, Y.R., Åkesson, K., Ali, N., Woolf, A.D., 2020. Global health policy in the 21st century: Challenges and opportunities to arrest the global disability burden from musculoskeletal health conditions. Best Practice & Research Clinical Rheumatology, Bone and Joint Health - the Future 34, 101549. https://doi.org/10.1016/j.berh.2020.101549

Brose, S.W., Jennings, D.C., Kwok, J., Stuart, C.L., O'Connell, S.M., Pauli, H.A., Liu, B., 2013. Sham Manual Medicine Protocol for Cervical Strain-Counterstrain Research. PM&R 5, 400–407. https://doi.org/10.1016/j.pmrj.2013.01.005

Brouwers, M.C., Kho, M.E., Browman, G.P., Burgers, J.S., Cluzeau, F., Feder, G., Fervers, B., Graham, I.D., Grimshaw, J., Hanna, S.E., Littlejohns, P., Makarski, J., Zitzelsberger, L., 2010. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ 182, E839–E842. https://doi.org/10.1503/cmaj.090449

Bruehl, S., Apkarian, A.V., Ballantyne, J.C., Berger, A., Borsook, D., Chen, W.G., Farrar, J.T., Haythornthwaite, J.A., Horn, S.D., Iadarola, M.J., Inturrisi, C.E., Lao, L., Mackey, S., Mao, J., Sawczuk, A., Uhl, G.R., Witter, J., Woolf, C.J., Zubieta, J.-K., Lin, Y., 2013. Personalized Medicine and Opioid Analgesic Prescribing for Chronic Pain: Opportunities and Challenges. The Journal of Pain 14, 103–113. https://doi.org/10.1016/j.jpain.2012.10.016

Buchbinder, R., Underwood, M., Hartvigsen, J., Maher, C.G., 2020. The Lancet Series call to action to reduce low value care for low back pain: an update. PAIN 161, S57. https://doi.org/10.1097/j.pain.0000000000001869

Bueno, M., Harrison, D., 2019. What happened to the principle of equipoise in the planning, designing and conducting of placebo-controlled trials for neonatal procedural pain? Pediatrics & Neonatology 60, 479–480. https://doi.org/10.1016/j.pedneo.2019.05.003

Bushnell, M.C., Frangos, E., Madian, N., 2021. Non-pharmacological Treatment of Pain: Grand Challenge and Future Opportunities. Frontiers in Pain Research 2.

Buttagat, V., Narktro, T., Onsrira, K., Pobsamai, C., 2016. Short-term effects of traditional Thai massage on electromyogram, muscle tension and pain among patients with upper back pain associated with myofascial trigger points. Complementary Therapies in Medicine 28, 8–12. https://doi.org/10.1016/j.ctim.2016.07.004

Cai, X., Gewandter, J.S., He, H., Turk, D.C., Dworkin, R.H., McDermott, M.P., 2020. Estimands and missing data in clinical trials of chronic pain treatments: advances in design and analysis. PAIN 161, 2308–2320. https://doi.org/10.1097/j.pain.0000000000001937

Çakici, N., Fakkel, T.M., van Neck, J.W., Verhagen, A.P., Coert, J.H., 2016. Systematic review of treatments for diabetic peripheral neuropathy. Diabet Med 33, 1466–1476. https://doi.org/10.1111/dme.13083

Califf, R.M., Sugarman, J., 2015. Exploring the ethical and regulatory issues in pragmatic clinical trials. Clinical Trials 12, 436–441. https://doi.org/10.1177/1740774515598334

Calner, T., Isaksson, G., Michaelson, P., 2021. Physiotherapy treatment experiences of persons with persistent musculoskeletal pain: A qualitative study. null 37, 28–37. https://doi.org/10.1080/09593985.2019.1622162

Calster, B.V., Wynants, L., Riley, R.D., Smeden, M. van, Collins, G.S., 2021. Methodology over metrics: current scientific standards are a disservice to patients and society. Journal of Clinical Epidemiology 138, 219–226. https://doi.org/10.1016/j.jclinepi.2021.05.018

Campbell, C.M., Edwards, R.R., 2012. Ethnic differences in pain and pain management. Pain Management 2, 219–230. https://doi.org/10.2217/pmt.12.7

Campbell, M.K., Entwistle, V.A., Cuthbertson, B.H., Skea, Z.C., Sutherland, A.G., McDonald, A.M., Norrie, J.D., Carlson, R.V., Bridgman, S., KORAL study group, 2011. Developing a placebo-controlled trial in surgery: issues of design, acceptability and feasibility. Trials 12, 50. https://doi.org/10.1186/1745-6215-12-50

Carnes, D., Mars, T., Plunkett, A., Nanke, L., Abbey, H., 2017. A mixed methods evaluation of a third wave cognitive behavioural therapy and osteopathic treatment programme for chronic pain in primary care (OsteoMAP). International Journal of Osteopathic Medicine 24, 12–17. https://doi.org/10.1016/j.ijosm.2017.03.005

Carnes, D., Mars, T.S., Mullinger, B., Froud, R., Underwood, M., 2010. Adverse events and manual therapy: A systematic review. Manual Therapy 15, 355–363. https://doi.org/10.1016/j.math.2009.12.006

Carpenter, J.R., Roger, J.H., Kenward, M.G., 2013. Analysis of Longitudinal Trials with Protocol Deviation: A Framework for Relevant, Accessible Assumptions, and Inference via Multiple Imputation. Journal of Biopharmaceutical Statistics 23, 1352–1371. https://doi.org/10.1080/10543406.2013.834911

Carpenter, M., 2016. can an osteopathic rib-raising technique influence blood glucose levels in type 2 diabetic patients? An ABAB single case research design. [WWW Document].

Carvalho, C., Caetano, J.M., Cunha, L., Rebouta, P., Kaptchuk, T.J., Kirsch, I., 2016. Open-label placebo treatment in chronic low back pain: a randomized controlled trial. Pain 157, 2766–2772. https://doi.org/10.1097/j.pain.0000000000000700

Cashin, A.G., Lee, H., Bagg, M.K., O'Hagan, E., Traeger, A.C., Kamper, S.J., Folly, T., Jones, M.D., Booth, J., McAuley, J.H., 2020. A systematic review highlights the need to improve the quality and applicability of trials of physical therapy interventions for low back pain. Journal of Clinical Epidemiology 126, 106–115. https://doi.org/10.1016/j.jclinepi.2020.06.025

Castelnuovo, G., Pietrabissa, G., Manzoni, G.M., Cattivelli, R., Rossi, A., Novelli, M., Varallo, G., Molinari, E., 2017. Cognitive behavioral therapy to aid weight loss in obese patients: Current perspectives. Psychology Research and Behavior Management 10. https://doi.org/10.2147/PRBM.S113278

Castro-Sanchez, A.M., Mataran-Penarrocha, G.A., Arroyo-Morales, M., Saavedra-Hernandez, M., Fernandez-Sola, C., Moreno-Lorenzo, C., 2011a. Effects of myofascial release techniques on pain, physical function, and postural stability in patients with fibromyalgia: a randomized controlled trial. Clinical Rehabilitation 25, 800–13. https://doi.org/10.1177/0269215511399476

Castro-Sanchez, A.M., Mataran-Penarrocha, G.A., Sanchez-Labraca, N., Quesada-Rubio, J.M., Granero-Molina, J., Moreno-Lorenzo, C., 2011b. A randomized controlled trial investigating the effects of craniosacral therapy on pain and heart rate variability in fibromyalgia patients. Clinical Rehabilitation 25, 25–35. https://doi.org/10.1177/0269215510375909

Cerritelli, F., Verzella, M., Cicchitti, L., D'Alessandro, G., Vanacore, N., 2016. The paradox of sham therapy and placebo effect in osteopathy. Medicine (Baltimore) 95, e4728. https://doi.org/10.1097/MD.0000000000004728

Chaibi, A., Benth, J.S., Tuchin, P.J., Russell, M.B., 2017. Chiropractic spinal manipulative therapy for migraine: a three-armed, single-blinded, placebo, randomized controlled trial. Eur J Neurol 24, 143–153. https://doi.org/10.1111/ene.13166

Chaibi, A., Šaltytė Benth, J., Bjørn Russell, M., 2015. Validation of Placebo in a Manual Therapy Randomized Controlled Trial. Sci Rep 5, 11774. https://doi.org/10.1038/srep11774

Chalkidou, K., Tunis, S., Whicher, D., Fowler, R., Zwarenstein, M., 2012. The role for pragmatic randomized controlled trials (pRCTs) in comparative effectiveness research. Clinical Trials 9, 436–446. https://doi.org/10.1177/1740774512450097

Chan, A.-W., Tetzlaff, J.M., Altman, D.G., Laupacis, A., Gøtzsche, P.C., Krleža-Jerić, K., Hróbjartsson, A., Mann, H., Dickersin, K., Berlin, J.A., Doré, C.J., Parulekar, W.R., Summerskill, W.S.M., Groves, T., Schulz, K.F., Sox, H.C., Rockhold, F.W., Rennie, D., Moher, D., 2013. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. Annals of Internal Medicine 158, 200–207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Chatchawan, U., Eungpinichpong, W., Plandee, P., Yamauchi, J., 2015. Effects of Thai Foot Massage on Balance Performance in Diabetic Patients with Peripheral Neuropathy: A Randomized Parallel-Controlled Trial. Med Sci Monit Basic Res 21, 68–75. https://doi.org/10.12659/MSMBR.894163

Chen, K.W., Perlman, A., Liao, J.G., Lam, A., Staller, J., Sigal, L.H., 2008. Effects of external qigong therapy on osteoarthritis of the knee. A randomized controlled trial. Clinical Rheumatology 27, 1497–505. https://doi.org/10.1007/s10067-008-0955-4

Chen, P.-H.A., Cheong, J.H., Jolly, E., Elhence, H., Wager, T.D., Chang, L.J., 2019. Socially transmitted placebo effects. Nat Hum Behav 3, 1295–1305. https://doi.org/10.1038/s41562-019-0749-5

Chen, W., Yang, G., Liu, B., Manheimer, E., Liu, J.-P., 2013. Manual Acupuncture for Treatment of Diabetic Peripheral Neuropathy: A Systematic Review of Randomized Controlled Trials. PLOS ONE 8, e73764. https://doi.org/10.1371/journal.pone.0073764

Cherkin, D.C., 2021. Are Methods for Evaluating Medications Appropriate for Evaluating Nonpharmacological Treatments for Pain?—Challenges for an Emerging Field of Research. JAMA Internal Medicine 181, 328–329. https://doi.org/10.1001/jamainternmed.2020.7081

Cherkin, D.C., Sherman, K.J., Balderson, B.H., Cook, A.J., Anderson, M.L., Hawkes, R.J., Hansen, K.E., Turner, J.A., 2016. Effect of Mindfulness-Based Stress Reduction vs Cognitive Behavioral Therapy or Usual Care on Back Pain and Functional Limitations in Adults With Chronic Low Back Pain: A Randomized Clinical Trial. JAMA 315, 1240–1249. https://doi.org/10.1001/jama.2016.2323

Chesterton, L.S., Blagojevic-Bucknall, M., Burton, C., Dziedzic, K.S., Davenport, G., Jowett, S.M., Myers, H.L., Oppong, R., Rathod-Mistry, T., van der Windt, D.A., Hay, E.M., Roddy, E., 2018. The clinical and cost-effectiveness of corticosteroid injection versus night splints for carpal tunnel syndrome (INSTINCTS trial): an open-label, parallel group, randomised

controlled trial. Lancet 392, 1423–1433. https://doi.org/10.1016/S0140-6736(18)31572-1

Christian, J.B., Brouwer, E.S., Girman, C.J., Bennett, D., Davis, K.J., Dreyer, N.A., 2020. Masking in Pragmatic Trials: Who, What, and When to Blind. Ther Innov Regul Sci 54, 431–436. https://doi.org/10.1007/s43441-019-00073-7

Ciechanowski, P., Sullivan, M., Jensen, M., Romano, J., Summers, H., 2003. The relationship of attachment style to depression, catastrophizing and health care utilization in patients with chronic pain. Pain 104, 627–637. https://doi.org/10.1016/S0304-3959(03)00120-9

Clark, L., Fairhurst, C., Hewitt, C.E., Birks, Y., Brabyn, S., Cockayne, S., Rodgers, S., Hicks, K., Hodgson, R., Littlewood, E., Torgerson, D.J., 2014. A methodological review of recent meta-analyses has found significant heterogeneity in age between randomized groups. Journal of Clinical Epidemiology 67, 1016–1024. https://doi.org/10.1016/j.jclinepi.2014.04.007

Colagiuri, B., 2010. Participant expectancies in double-blind randomized placebo-controlled trials: potential limitations to trial validity. Clin Trials 7, 246–255. https://doi.org/10.1177/1740774510367916

Colagiuri, B., Sharpe, L., Scott, A., 2019. The Blind Leading the Not-So-Blind: A Meta-Analysis of Blinding in Pharmacological Trials for Chronic Pain. The Journal of Pain 20, 489–500. https://doi.org/10.1016/j.jpain.2018.09.002

Colberg, S.R., Sigal, R.J., Yardley, J.E., Riddell, M.C., Dunstan, D.W., Dempsey, P.C., Horton, E.S., Castorino, K., Tate, D.F., 2016. Physical Activity/Exercise and Diabetes: A Position Statement of the American Diabetes Association. Diabetes Care 39, 2065–2079. https://doi.org/10.2337/dc16-1728

Coleman, E.A., Chugh, A., Williams, M.V., Grigsby, J., Glasheen, J.J., McKenzie, M., Min, S.-J., 2013. Understanding and Execution of Discharge Instructions. Am J Med Qual 28, 383–391. https://doi.org/10.1177/1062860612472931

Collinge, W., Kahn, J., Walton, T., Kozak, L., Bauer-Wu, S., Fletcher, K., Yarnold, P., Soltysik, R., 2013. Touch, Caring, and Cancer: randomized controlled trial of a multimedia caregiver education program. Support Care Cancer 21, 1405–1414. https://doi.org/10.1007/s00520-012-1682-6

Colloca, L., 2019a. The Placebo Effect in Pain Therapies. Annu Rev Pharmacol Toxicol 59, 191–211. https://doi.org/10.1146/annurev-pharmtox-010818-021542

Colloca, L., 2019b. The Placebo Effect in Pain Therapies. Annu Rev Pharmacol Toxicol 59, 191–211. https://doi.org/10.1146/annurev-pharmtox-010818-021542

Colloca, L., Barsky, A.J., 2020. Placebo and Nocebo Effects. New England Journal of Medicine 382, 554–561. https://doi.org/10.1056/NEJMra1907805

Consorti, G., Castagna, C., Tramontano, M., Longobardi, M., Castagna, P., Di Lernia, D., Lunghi, C., 2023. Reconceptualizing Somatic Dysfunction in the Light of a Neuroaesthetic Enactive Paradigm. Healthcare 11, 479. https://doi.org/10.3390/healthcare11040479

Constantino, M.J., Vîslă, A., Coyne, A.E., Boswell, J.F., 2018. A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. Psychotherapy 55, 473–485. https://doi.org/10.1037/pst0000169

Cook, T.D., Campbell, D.T., Shadish, W., 2002. Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin Boston, MA.

Coon, C.D., Cook, K.F., 2018. Moving from significance to real-world meaning: methods for interpreting change in clinical outcome assessment scores. Qual Life Res 27, 33–40. https://doi.org/10.1007/s11136-017-1616-3

Corry, M., Porter, S., McKenna, H., 2019. The redundancy of positivism as a paradigm for nursing research. Nursing Philosophy 20, e12230. https://doi.org/10.1111/nup.12230

Costa, N., Olson, R., Mescouto, K., Hodges, P.W., Dillon, M., Evans, K., Walsh, K., Jensen, N., Setchell, J., 2022. Uncertainty in low back pain care – insights from an ethnographic study. Disability and Rehabilitation 0, 1–12. https://doi.org/10.1080/09638288.2022.2040615

Coste, J., Medkour, T., Maigne, J.-Y., Pérez, M., Laroche, F., Perrot, S., 2021. Osteopathic medicine for fibromyalgia: a sham-controlled randomized clinical trial. Therapeutic Advances in Musculoskeletal 13, 1759720X211009017. https://doi.org/10.1177/1759720X211009017

Council for International Organizations of Medical Sciences, 2016. 2016 International ethical guidelines for health-related research involving humans [WWW Document]. Council for International Organizations of Medical Sciences. URL https://cioms.ch/publications/product/international-ethical-guidelines-for-health-related-research-involving-humans/ (accessed 1.18.23).

Cowie, M.R., Blomster, J.I., Curtis, L.H., Duclaux, S., Ford, I., Fritz, F., Goldman, S., Janmohamed, S., Kreuzer, J., Leenay, M., Michel, A., Ong, S., Pell, J.P., Southworth, M.R., Stough, W.G., Thoenes, M., Zannad, F., Zalewski, A., 2017. Electronic health records to facilitate clinical research. Clin Res Cardiol 106, 1–9. https://doi.org/10.1007/s00392-016-1025-6

Cragg, W.J., Hurley, C., Yorke-Edwards, V., Stenning, S.P., 2021. Dynamic methods for ongoing assessment of site-level risk in risk-based monitoring of clinical trials: A scoping review. Clinical Trials 1740774520976561. https://doi.org/10.1177/1740774520976561

Craig, P., Campbell, M., 2015. Evaluability Assessment: a systematic approach to deciding whether and how to evaluate programmes and policies: a What Works Scotland Working paper. https://doi.org/10.13140/RG.2.1.2007.4725

Cramer, H., 2022. In the Same Boat: The Evidence Base of (the Whole of) Medicine. Journal of Integrative and Complementary Medicine 28, 697–698. https://doi.org/10.1089/jicm.2022.29109.editorial

Cronbach, L.J., Shapiro, K., 1982. Designing evaluations of educational and social programs. The Jossey-Bass series in social and behavioral science and in higher education.

Croskerry, P., 2009. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Adv Health Sci Educ Theory Pract 14 Suppl 1, 27–35. https://doi.org/10.1007/s10459-009-9182-2

Curtis, P., Gaylord, S.A., Park, J., Faurot, K.R., Coble, R., Suchindran, C., Coeytaux, R.R., Wilkinson, L., Mann, J.D., 2011. Credibility of low-strength static magnet therapy as an attention control intervention for a randomized controlled study of CranioSacral therapy for migraine headaches. Journal of alternative and complementary medicine (New York, N.Y.) 17, 711–21. https://doi.org/10.1089/acm.2010.0277

D'Alessandro, G., Ruffini, N., Aquino, A., Galli, M., Innocenti, M., Tramontano, M., Cerritelli, F., 2022. Differences between experimental and placebo arms in manual therapy trials: a methodological review. BMC Medical Research Methodology 22, 219. https://doi.org/10.1186/s12874-022-01704-8

D'Alessandro, G., Ruffini, N., Iacopini, A., Annoni, M., Kossowsky, J., Cerritelli, F., 2021. Overcoming placebo-related challenges in manual therapy trials: The 'whats and hows' and the 'touch equality assumption' proposals. International Journal of Osteopathic Medicine 42, 5–10. https://doi.org/10.1016/j.ijosm.2021.10.003

Dal-Ré, R., Boer, A. de, James, S.K., 2020. The design can limit PRECIS-2 retrospective assessment of the clinical trial explanatory/pragmatic features. Journal of Clinical Epidemiology 126, 193–201. https://doi.org/10.1016/j.jclinepi.2020.03.027

Dal-Ré, R., Janiaud, P., Ioannidis, J.P.A., 2018. Real-world evidence: How pragmatic are randomized controlled trials labeled as pragmatic? BMC Med 16, 49. https://doi.org/10.1186/s12916-018-1038-2

Daniali, H., Flaten, M.A., 2019. A Qualitative Systematic Review of Effects of Provider Characteristics and Nonverbal Behavior on Pain, and Placebo and Nocebo Effects. Frontiers in Psychiatry 10.

Dave, A.D., Ruaño, G., Kost, J., Wang, X., 2022. Automated Extraction of Pain Symptoms: A Natural Language Approach using Electronic Health Records. Pain Physician.

Davies, B., Cramp, F., Gauntlett-Gilbert, J., Wynick, D., McCabe, C.S., 2015. The role of physical activity and psychological coping strategies in the management of painful diabetic neuropathy – A systematic review of the literature. Physiotherapy 101, 319–326. https://doi.org/10.1016/j.physio.2015.04.003

Davies, M., Brophy, S., Williams, R., Taylor, A., 2006. The Prevalence, Severity, and Impact of Painful Diabetic Peripheral Neuropathy in Type 2 Diabetes. Diabetes Care 29, 1518–1522. https://doi.org/10.2337/dc05-2228

Davis, A.J., Hettinga, F., Beedie, C., 2020. You don't need to administer a placebo to elicit a placebo effect: Social factors trigger neurobiological pathways to enhance sports performance. European Journal of Sport Science 20, 302–312. https://doi.org/10.1080/17461391.2019.1635212

Davis, M.C., Zautra, A.J., 2013. An online mindfulness intervention targeting socioemotional regulation in fibromyalgia: results of a randomized controlled trial. Annals of behavioral medicine : a publication of the Society of Behavioral Medicine 46, 273–84. https://doi.org/10.1007/s12160-013-9513-7

Dean, B.J.F., Collins, J., Thurley, N., Rombach, I., Bennell, K., 2021. Exercise therapy with or without other physical therapy interventions versus placebo interventions for osteoarthritis –Systematic review. Osteoarthritis and Cartilage Open 3, 100195. https://doi.org/10.1016/j.ocarto.2021.100195

Dechartres, A., Trinquart, L., Boutron, I., Ravaud, P., 2013. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. BMJ 346, f2304. https://doi.org/10.1136/bmj.f2304

Dekkers, O.M., Elm, E. von, Algra, A., Romijn, J.A., Vandenbroucke, J.P., 2010. How to assess the external validity of therapeutic trials: a conceptual approach. Int J Epidemiol 39, 89–94. https://doi.org/10.1093/ije/dyp174

Delitto, A., Patterson, C.G., Stevans, J.M., Freburger, J.K., Khoja, S.S., Schneider, M.J., Greco, C.M., Freel, J.A., Sowa, G.A., Wasan, A.D., Brennan, G.P., Hunter, S.J., Minick, K.I., Wegener, S.T., Ephraim, P.L., Beneciuk, J.M., George, S.Z., Saper, R.B., 2021. Stratified care to prevent chronic low back pain in high-risk patients: The TARGET trial. A multi-site pragmatic

cluster randomized trial. EClinicalMedicine 34, 100795. https://doi.org/10.1016/j.eclinm.2021.100795

Derbyshire, S., Field, J., Vennik, J., Sanders, M., Newell, D., 2021. “Chiropractic is manual therapy, not talk therapy”: a qualitative analysis exploring perceived barriers to remote consultations by chiropractors. Chiropractic & Manual Therapies 29, 47. https://doi.org/10.1186/s12998-021-00404-2

Devilly, G.J., Borkovec, T.D., 2000. Psychometric properties of the credibility/expectancy questionnaire. Journal of Behavior Therapy and Experimental Psychiatry 31, 73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

Di Franco, M., Iannuccelli, C., Bazzichi, L., Atzeni, F., Consensi, A., Salaffi, F., Pietropaolo, M., Alessandri, C., Basili, S., Olivieri, M., Bombardieri, S., Valesini, G., Sarzi-Puttini, P., 2011. Misdiagnosis in fibromyalgia: a multicentre study. Clin Exp Rheumatol 29, S104-108.

Díaz Cerrato, I., 2008. MODIFICACIONES EN LA PIO Y EN LA TA EN DIABÉTICOS TIPO I TRAS LA MANIPULACIÓN GLOBAL OAA SEGÚN FRYETTE.(ECA) [WWW Document].

Dixit, S., Asiri, F., 2014. Pharmacological and nonpharmacological therapies in the management of diabetic peripheral neuropathy in type 2 diabetes: A comprehensive review. Journal of Cardiovascular Disease Research 5.

Djulbegovic, B., Guyatt, G.H., 2017. Progress in evidence-based medicine: a quarter century on. The Lancet 390, 415–423. https://doi.org/10.1016/S0140-6736(16)31592-6

Draper-Rodi, J., Vaucher, P., Hohenschurz-Schmidt, D., Morin, C., Thomson, O.P., 2022a. 4 M’s to make sense of evidence – Avoiding the propagation of mistakes, misinterpretation, misrepresentation and misinformation. International Journal of Osteopathic Medicine. https://doi.org/10.1016/j.ijosm.2022.04.007

Draper-Rodi, J., Vaucher, P., Hohenschurz-Schmidt, D., Morin, C., Thomson, O.P., 2022b. 4 M’s to make sense of evidence – Avoiding the propagation of mistakes, misinterpretation, misrepresentation and misinformation. International Journal of Osteopathic Medicine 44, 29–35. https://doi.org/10.1016/j.ijosm.2022.04.007

Drucker, A.M., Chan, A.-W., 2020. Blindsided: challenging the dogma of masking in clinical trials. BMJ 368, m229. https://doi.org/10.1136/bmj.m229

Duncan, E.M., Bennett, T., Gillies, K., 2020. Assessing effective interventions to improve trial retention: do they contain behaviour change techniques? Trials 21, 213. https://doi.org/10.1186/s13063-020-4151-4

Dunlop, B.W., Thase, M.E., Wun, C.-C., Fayyad, R., Guico-Pabia, C.J., Musgnung, J., Ninan, P.T., 2012. A Meta-analysis of Factors Impacting Detection of Antidepressant Efficacy in Clinical Trials: The Importance of Academic Sites. Neuropsychopharmacol 37, 2830–2836. https://doi.org/10.1038/npp.2012.153

Dworkin, J.D., McKeown, A., Farrar, J.T., Gilron, I., Hunsinger, M., Kerns, R.D., McDermott, M.P., Rappaport, B.A., Turk, D.C., Dworkin, R.H., Gewandter, J.S., 2016. Deficiencies in reporting of statistical methodology in recent randomized trials of nonpharmacologic pain treatments: ACTTION systematic review. J Clin Epidemiol 72, 56–65. https://doi.org/10.1016/j.jclinepi.2015.10.019

Dworkin, R.H., Evans, S.R., Mbowe, O., McDermott, M.P., 2021. Essential statistical principles of clinical trials of pain treatments. PAIN Reports 6, e863. https://doi.org/10.1097/PR9.0000000000000863

Dworkin, R.H., Turk, D.C., Farrar, J.T., Haythornthwaite, J.A., Jensen, M.P., Katz, N.P., Kerns, R.D., Stucki, G., Allen, R.R., Bellamy, N., Carr, D.B., Chandler, J., Cowan, P., Dionne, R., Galer, B.S., Hertz, S., Jadad, A.R., Kramer, L.D., Manning, D.C., Martin, S., McCormick, C.G., McDermott, M.P., McGrath, P., Quessy, S., Rappaport, B.A., Robbins, W., Robinson, J.P., Rothman, M., Royal, M.A., Simon, L., Stauffer, J.W., Stein, W., Tollett, J., Wernicke, J., Witter, J., 2005. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. PAIN 113, 9. https://doi.org/10.1016/j.pain.2004.09.012

Dworkin, R.H., Turk, D.C., Peirce-Sandner, S., Baron, R., Bellamy, N., Burke, L.B., Chappell, A., Chartier, K., Cleeland, C.S., Costello, A., Cowan, P., Dimitrova, R., Ellenberg, S., Farrar, J.T., French, J.A., Gilron, I., Hertz, S., Jadad, A.R., Jay, G.W., Kalliomäki, J., Katz, N.P., Kerns, R.D., Manning, D.C., McDermott, M.P., McGrath, P.J., Narayana, A., Porter, L., Quessy, S., Rappaport, B.A., Rauschkolb, C., Reeve, B.B., Rhodes, T., Sampaio, C., Simpson, D.M., Stauffer, J.W., Stucki, G., Tobias, J., White, R.E., Witter, J., 2010. Research design considerations for confirmatory chronic pain clinical trials: IMMPACT recommendations. PAIN® 149, 177–193. https://doi.org/10.1016/j.pain.2010.02.018

Dworkin, R.H., Turk, D.C., Peirce-Sandner, S., Burke, L.B., Farrar, J.T., Gilron, I., Jensen, M.P., Katz, N.P., Raja, S.N., Rappaport, B.A., Rowbotham, M.C., Backonja, M.-M., Baron, R., Bellamy, N., Bhagwagar, Z., Costello, A., Cowan, P., Fang, W.C., Hertz, S., Jay, G.W., Junor, R., Kerns, R.D., Kerwin, R., Kopecky, E.A., Lissin, D., Malamut, R., Markman, J.D., McDermott, M.P., Munera, C., Porter, L., Rauschkolb, C., Rice, A.S.C., Sampaio, C., Skljarevski, V., Sommerville, K., Stacey, B.R., Steigerwald, I., Tobias, J., Trentacosti, A.M., Wasan, A.D., Wells, G.A., Williams, J., Witter, J., Ziegler, D., 2012. Considerations for improving assay sensitivity in chronic pain clinical trials: IMMPACT recommendations. PAIN 153, 1148–1158. https://doi.org/10.1016/j.pain.2012.03.003

Dworkin, R.H., Turk, D.C., Wyrwich, K.W., Beaton, D., Cleeland, C.S., Farrar, J.T., Haythornthwaite, J.A., Jensen, M.P., Kerns, R.D., Ader, D.N., Brandenburg, N., Burke, L.B., Cella, D., Chandler, J., Cowan, P., Dimitrova, R., Dionne, R., Hertz, S., Jadad, A.R., Katz, N.P., Kehlet, H., Kramer, L.D., Manning, D.C., McCormick, C., McDermott, M.P., McQuay, H.J., Patel, S., Porter, L., Quessy, S., Rappaport, B.A., Rauschkolb, C., Revicki, D.A., Rothman, M., Schmader, K.E., Stacey, B.R., Stauffer, J.W., von Stein, T., White, R.E., Witter, J., Zavisic, S., 2008. Interpreting the Clinical Importance of Treatment Outcomes in Chronic Pain Clinical Trials: IMMPACT Recommendations. The Journal of Pain 9, 105–121. https://doi.org/10.1016/j.jpain.2007.09.005

Eapen, Z.J., Lauer, M.S., Temple, R.J., 2014. The Imperative of Overcoming Barriers to the Conduct of Large, Simple Trials. JAMA 311, 1397–1398. https://doi.org/10.1001/jama.2014.1030

Eguaras, N., Rodriguez-Lopez, E.S., Lopez-Dicastillo, O., Franco-Sierra, M.A., Ricard, F., Oliva-Pascual-Vaca, A., 2019. Effects of Osteopathic Visceral Treatment in Patients with Gastroesophageal Reflux: A Randomized Controlled Trial. Journal of Clinical Medicine 8. https://doi.org/10.3390/jcm8101738

Eklund, A., Jensen, I., Lohela-Karlsson, M., Hagberg, J., Leboeuf-Yde, C., Kongsted, A., Bodin, L., Axén, I., 2018. The Nordic Maintenance Care program: Effectiveness of chiropractic maintenance care versus symptom-guided treatment for recurrent and persistent low back pain—A pragmatic randomized controlled trial. PLOS ONE 13, e0203029. https://doi.org/10.1371/journal.pone.0203029

Eldridge, S.M., Chan, C.L., Campbell, M.J., Bond, C.M., Hopewell, S., Thabane, L., Lancaster, G.A., 2016a. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 355, i5239. https://doi.org/10.1136/bmj.i5239

Eldridge, S.M., Lancaster, G.A., Campbell, M.J., Thabane, L., Hopewell, S., Coleman, C.L., Bond, C.M., 2016b. Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. PLoS One 11. https://doi.org/10.1371/journal.pone.0150205

Emanuel, E.J., Miller, F.G., 2001. The Ethics of Placebo-Controlled Trials — A Middle Ground. N Engl J Med 345, 915–919. https://doi.org/10.1056/NEJM200109203451211

Emery, P., Horton, S., Dumitru, R.B., Naraghi, K., van der Heijde, D., Wakefield, R.J., Hensor, E.M.A., Buch, M.H., 2020. Pragmatic randomised controlled trial of very early etanercept and MTX versus MTX with delayed etanercept in RA: the VEDERA trial. Ann Rheum Dis 79, 464–471. https://doi.org/10.1136/annrheumdis-2019-216539

Estabrooks, P.A., Glasgow, R.E., Dzewaltowski, D.A., 2003. Physical activity promotion through primary care. JAMA 289, 2913–2916. https://doi.org/10.1001/jama.289.22.2913

Evans, S.R., Bigelow, R., Chuang-Stein, C., Ellenberg, S.S., Gallo, P., He, W., Jiang, Q., Rockhold, F., 2019. Presenting Risks and Benefits: Helping the Data Monitoring Committee Do Its Job. Ann Intern Med 172, 119–125. https://doi.org/10.7326/M19-1491

Evans, S.R., Follmann, D., 2016. Using Outcomes to Analyze Patients Rather than Patients to Analyze Outcomes: A Step Toward Pragmatism in Benefit:Risk Evaluation. Statistics in Biopharmaceutical Research 8, 386–393. https://doi.org/10.1080/19466315.2016.1207561

Evans, S.R., Rubin, D., Follmann, D., Pennello, G., Huskins, W.C., Powers, J.H., Schoenfeld, D., Chuang-Stein, C., Cosgrove, S.E., Fowler, V.G., Jr, Lautenbach, E., Chambers, H.F., 2015. Desirability of Outcome Ranking (DOOR) and Response Adjusted for Duration of Antibiotic Risk (RADAR). Clinical Infectious Diseases 61, 800–806. https://doi.org/10.1093/cid/civ495

Evers, A.W.M., Colloca, L., Blease, C., Annoni, M., Atlas, L.Y., Benedetti, F., Bingel, U., Büchel, C., Carvalho, C., Colagiuri, B., Crum, A.J., Enck, P., Gaab, J., Geers, A.L., Howick, J., Jensen, K.B., Kirsch, I., Meissner, K., Napadow, V., Peerdeman, K.J., Raz, A., Rief, W., Vase, L., Wager, T.D., Wampold, B.E., Weimer, K., Wiech, K., Kaptchuk, T.J., Klinger, R., Kelley, J.M., 2018. Implications of Placebo and Nocebo Effects for Clinical Practice: Expert Consensus. PPS 87, 204–210. https://doi.org/10.1159/000490354

Faden, R.R., Kass, N.E., Goodman, S.N., Pronovost, P., Tunis, S., Beauchamp, T.L., 2013. An Ethics Framework for a Learning Health Care System: A Departure from Traditional Research Ethics and Clinical Ethics. Hastings Center Report 43, S16–S27. https://doi.org/10.1002/hast.134

Faltinsen, E., Todorovac, A., Bruun, L.S., Hróbjartsson, A., Gluud, C., Kongerslev, M.T., Simonsen, E., Storebø, O.J., 2022. Control interventions in randomised trials among people with mental health disorders. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.MR000050.pub2

Fawkes, C., Carnes, D., 2021. Patient reported outcomes in a large cohort of patients receiving osteopathic care in the United Kingdom. PLOS ONE 16, e0249719. https://doi.org/10.1371/journal.pone.0249719

Felson, D.T., Redmond, A.C., Chapman, G.J., Smith, T.O., Hamilton, D.F., Jones, R.K., Holt, C.A., Callaghan, M.J., Mason, D.J., Conaghan, P.G., 2016. Recommendations for the conduct of efficacy trials of treatment devices for osteoarthritis: a report from a working group of the Arthritis Research UK Osteoarthritis and Crystal Diseases Clinical Studies Group. Rheumatology (Oxford) 55, 320–326. https://doi.org/10.1093/rheumatology/kev328

Ferguson, M.C., McNicol, E., Kleykamp, B.A., Sandoval, K., Haroutounian, S., Holzer, K.J., Kerns, R.D., Veasley, C., Turk, D.C., Dworkin, R.H., 2022. Perspectives on participation in clinical trials among individuals with pain, depression, and/or anxiety: an ACTTION scoping review. The Journal of Pain. https://doi.org/10.1016/j.jpain.2022.09.001

Ferini-Strambi, L., 2017. Neuropathic Pain and Sleep: A Review. Pain Ther 6, 19–23. https://doi.org/10.1007/s40122-017-0089-y

Fernandez-Carnero, J., Sierra-Silvestre, E., Beltran-Alacreu, H., Gil-Martinez, A., La Touche, R., 2019. Neural Tension Technique Improves Immediate Conditioned Pain Modulation in Patients with Chronic Neck Pain: A Randomized Clinical Trial. Pain medicine (Malden, Mass.) 20, 1227–1235. https://doi.org/10.1093/pm/pny115

Finnegan-John, J., Molassiotis, A., Richardson, A., Ream, E., 2013. A Systematic Review of Complementary and Alternative Medicine Interventions for the Management of Cancer-Related Fatigue. Integr Cancer Ther 12, 276–290. https://doi.org/10.1177/1534735413485816

Finnerup, N.B., Attal, N., Haroutounian, S., McNicol, E., Baron, R., Dworkin, R.H., Gilron, I., Haanpää, M., Hansson, P., Jensen, T.S., Kamerman, P.R., Lund, K., Moore, A., Raja, S.N., Rice, A.S.C., Rowbotham, M., Sena, E., Siddall, P., Smith, B.H., Wallace, M., 2015. Pharmacotherapy for neuropathic pain in adults: a systematic review and meta-analysis. The Lancet Neurology 14, 162–173. https://doi.org/10.1016/S1474-4422(14)70251-0

Finucane, L.M., Stokes, E., Briggs, A.M., 2023. Its everyone's responsibility: Responding to the global burden of musculoskeletal health impairment. Musculoskeletal Science and Practice 64, 102743. https://doi.org/10.1016/j.msksp.2023.102743

Fiore, L.D., Lavori, P.W., 2016. Integrating Randomized Comparative Effectiveness Research with Patient Care. New England Journal of Medicine 374, 2152–2158. https://doi.org/10.1056/NEJMra1510057

Fischer, F., Lewith, G., Witt, C.M., Linde, K., von Ammon, K., Cardini, F., Falkenberg, T., Fønnebø, V., Johannessen, H., Reiter, B., Uehleke, B., Weidenhammer, W., Brinkhaus, B., 2014. A research roadmap for complementary and alternative medicine - what we need to know by 2020. Forsch Komplementmed 21, e1-16. https://doi.org/10.1159/000360744

Fischer, H.F., Junne, F., Witt, C., Ammon, K. von, Cardini, F., Fønnebø, V., Johannessen, H., Lewith, G., Uehleke, B., Weidenhammer, W., Brinkhaus, B., 2012. Key Issues in Clinical and Epidemiological Research in Complementary and Alternative Medicine – a Systematic Literature Review. CMR 19, 51–60. https://doi.org/10.1159/000343126

Fitch, K., Bernstein, S.J., Aguilar, M.D., Burnand, B., LaCalle, J.R., Lazaro, P., van het Loo, M., McDonnell, J., Vader, J., Kahan, J.P., 2001. The RAND/UCLA Appropriateness Method User's Manual. RAND Corporation.

Fjeldsoe, B., Neuhaus, M., Winkler, E., Eakin, E., 2011. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychology 30, 99–109. https://doi.org/10.1037/a0021974

Fønnebø, V., Grimsgaard, S., Walach, H., Ritenbaugh, C., Norheim, A.J., MacPherson, H., Lewith, G., Launsø, L., Koithan, M., Falkenberg, T., Boon, H., Aickin, M., 2007. Researching complementary and alternative treatments – the gatekeepers are not at home. BMC Medical Research Methodology 7, 7. https://doi.org/10.1186/1471-2288-7-7

Food and Drug Administration, 2016. Non-Inferiority Clinical Trials to Establish Effectiveness Guidance for Industry.

Ford, I., Norrie, J., 2016. Pragmatic Trials. New England Journal of Medicine 375, 454–463. https://doi.org/10.1056/NEJMra1510059

Franssen, W.M.A., Franssen, G.H.L.M., Spaas, J., Solmi, F., Eijnde, B.O., 2020. Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. Int J Behav Nutr Phys Act 17, 57. https://doi.org/10.1186/s12966-020-00955-2

Fraser, J.J., Saliba, S.A., Hart, J.M., Park, J.S., Hertel, J., 2020. Effects of midfoot joint mobilization on ankle-foot morphology and function following acute ankle sprain. A crossover clinical trial. Musculoskeletal science & practice 46, 102130. https://doi.org/10.1016/j.msksp.2020.102130

Freed, B., Williams, B., Situ, X., Landsman, V., Kim, J., Moroz, A., Bang, H., Park, J.J., 2021. Blinding, sham, and treatment effects in randomized controlled trials for back pain in 2000–2019: A review and meta-analytic approach. Clinical Trials 18, 361–370. https://doi.org/10.1177/1740774520984870

Freedland, K.E., King, A.C., Ambrosius, W.T., Mayo-Wilson, E., Mohr, D.C., Czajkowski, S.M., Thabane, L., Collins, L.M., Rebok, G.W., Treweek, S.P., Cook, T.D., Edinger, J.D., Stoney, C.M., Campo, R.A., Young-Hyman, D., Riley, W.T., 2019. The selection of comparators for randomized controlled trials of health-related behavioral interventions: recommendations of an NIH expert panel. Journal of Clinical Epidemiology 110, 74–81. https://doi.org/10.1016/j.jclinepi.2019.02.011

Fregni, F., Imamura, M., Chien, H.F., Lew, H.L., Boggio, P., Kaptchuk, T.J., Riberto, M., Hsing, W.T., Battistella, L.R., Furlan, A., 2010. Challenges and Recommendations for Placebo Controls in Randomized Trials in Physical and Rehabilitation Medicine. Am J Phys Med Rehabil 89, 160–172. https://doi.org/10.1097/PHM.0b013e3181bc0bbd

Frieden, T.R., 2017. Evidence for Health Decision Making — Beyond Randomized, Controlled Trials. New England Journal of Medicine 377, 465–475. https://doi.org/10.1056/NEJMra1614394

Froment, F.P., Olson, K.A., Hooper, T.L., Shaffer, S.M., Sizer, P.S., Woodhouse, L.J., Brismée, J.-M., 2019. Large variability found in musculoskeletal physiotherapy scope of practice throughout WCPT and IFOMPT affiliated countries: An international survey. Musculoskeletal Science and Practice 42, 104–119. https://doi.org/10.1016/j.msksp.2019.04.012

Fu, Q., Yang, H., Zhang, L., Liu, Yang, Li, X., Dai, M., Yang, Y., Yang, S., Xie, Y., Liu, Ying, Fu, L., Liu, Z., Zhang, Q., 2020. Traditional Chinese medicine foot bath combined with acupoint massage for the treatment of diabetic peripheral neuropathy: A systematic review and meta-analysis of 31 RCTs. Diabetes/Metabolism Research and Reviews 36, e3218. https://doi.org/10.1002/dmrr.3218

Furukawa, T.A., Noma, H., Caldwell, D.M., Honyashiki, M., Shinohara, K., Imai, H., Chen, P., Hunot, V., Churchill, R., 2014. Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. Acta Psychiatrica Scandinavica 130, 181–192. https://doi.org/10.1111/acps.12275

Gardner, T., Refshauge, K., McAuley, J., Goodall, S., Hübscher, M., Smith, L., 2015. Patient led goal setting in chronic low back pain—What goals are important to the patient and are they aligned to what we measure? Patient Education and Counseling 98, 1035–1038. https://doi.org/10.1016/j.pec.2015.04.012

Garland, E.L., Manusov, E.G., Froeliger, B., Kelly, A., Williams, J.M., Howard, M.O., 2014. Mindfulness-oriented recovery enhancement for chronic pain and prescription opioid misuse: Results from an early-stage randomized controlled trial. Journal of Consulting and Clinical Psychology 82, 448–459. https://doi.org/10.1037/a0035798

Gartlehner, G., Affengruber, L., Titscher, V., Noel-Storr, A., Dooley, G., Ballarini, N., König, F., 2020. Single-reviewer abstract screening missed 13 percent of relevant studies: a crowd-based, randomized controlled trial. Journal of Clinical Epidemiology 121, 20–28. https://doi.org/10.1016/j.jclinepi.2020.01.005

Gastaldon, C., Mosler, F., Toner, S., Tedeschi, F., Bird, V.J., Barbui, C., Priebe, S., 2019. Are trials of psychological and psychosocial interventions for schizophrenia and psychosis included in the NICE guidelines pragmatic? A systematic review. PLoS One 14, e0222891. https://doi.org/10.1371/journal.pone.0222891

General Osteopathic Council, 2022. Graduate Outcomes and Standards for Education and Training.

Gerdesmeyer, L., Klueter, T., Rahlfs, V.W., Muderis, M.A., Saxena, A., Gollwitzer, H., Harrasser, N., Stukenberg, M., Prehn-Kristensen, A., 2017. Randomized Placebo-Controlled Placebo Trial to Determine the Placebo Effect Size. Pain physician 20, 387–396.

Gewandter, J.S., Dworkin, R.H., Turk, D.C., Devine, E.G., Hewitt, D., Jensen, M.P., Katz, N.P., Kirkwood, A.A., Malamut, R., Markman, J.D., Vrijens, B., Burke, L., Campbell, J.N., Carr, D.B., Conaghan, P.G., Cowan, P., Doyle, M.K., Edwards, R.R., Evans, S.R., Farrar, J.T., Freeman, R., Gilron, I., Juge, D., Kerns, R.D., Kopecky, E.A., McDermott, M.P., Niebler, G., Patel, K.V., Rauck, R., Rice, A.S.C., Rowbotham, M., Sessler, N.E., Simon, L.S., Singla, N., Skljarevski, V., Tockarshewsky, T., Vanhove, G.F., Wasan, A.D., Witter, J., 2019. Improving study conduct and data quality in clinical trials of chronic pain treatments: IMMPACT recommendations. The Journal of Pain. https://doi.org/10.1016/j.jpain.2019.12.003

Gewandter, J.S., McDermott, M.P., McKeown, A., Smith, S.M., Pawlowski, J.R., Poli, J.J., Rothstein, D., Williams, M.R., Bujanover, S., Farrar, J.T., Gilron, I., Katz, N.P., Rowbotham, M.C., Turk, D.C., Dworkin, R.H., 2014. Reporting of intention-to-treat analyses in recent analgesic clinical trials: ACTTION systematic review and recommendations. PAIN® 155, 2714–2719. https://doi.org/10.1016/j.pain.2014.09.039

Gewandter, J.S., McKeown, A., McDermott, M.P., Dworkin, J.D., Smith, S.M., Gross, R.A., Hunsinger, M., Lin, A.H., Rappaport, B.A., Rice, A.S.C., Rowbotham, M.C., Williams, M.R., Turk, D.C., Dworkin, R.H., 2015. Data Interpretation in Analgesic Clinical Trials With Statistically Nonsignificant Primary Analyses: An ACTTION Systematic Review. The Journal of Pain 16, 3–10. https://doi.org/10.1016/j.jpain.2014.10.003

Gilbert, A.W., Billany, J.C.T., Adam, R., Martin, L., Tobin, R., Bagdai, S., Galvin, N., Farr, I., Allain, A., Davies, L., Bateson, J., 2020. Rapid implementation of virtual clinics due to COVID-19: report and early evaluation of a quality improvement initiative. BMJ Open Qual 9, e000985. https://doi.org/10.1136/bmjoq-2020-000985

Gilron, I., Blyth, F., Smith, B.H., 2020. Translating clinical trials into improved real-world management of pain: convergence of translational, population-based, and primary care research. PAIN 161, 36–42. https://doi.org/10.1097/j.pain.0000000000001684

Glasgow, R.E., Gaglio, B., Bennett, G., Jerome, G.J., Yeh, H.-C., Sarwer, D.B., Appel, L., Colditz, G., Wadden, T.A., Wells, B., 2012. Applying the PRECIS Criteria to Describe Three Effectiveness Trials of Weight Loss in Obese Patients with Comorbid Conditions. Health Services Research 47, 1051–1067. https://doi.org/10.1111/j.1475-6773.2011.01347.x

Glasgow, R.E., McKay, H.G., Piette, J.D., Reynolds, K.D., 2001. The RE-AIM framework for evaluating interventions: what can it tell us about approaches to chronic illness management? Patient Education and Counseling 44, 119–127. https://doi.org/10.1016/S0738-3991(00)00186-5

Gluud, L.L., 2006. Bias in Clinical Intervention Research. Am J Epidemiol 163, 493–501. https://doi.org/10.1093/aje/kwj069

Godwin, M., Ruhland, L., Casson, I., MacDonald, S., Delva, D., Birtwhistle, R., Lam, M., Seguin, R., 2003. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC Medical Research Methodology 3, 28. https://doi.org/10.1186/1471-2288-3-28

Goertz, C.M., Long, C.R., Hondras, M.A., Petri, R., Delgado, R., Lawrence, D.J., Owens, E.F., Meeker, W.C., 2013. Adding Chiropractic Manipulative Therapy to Standard Medical Care for Patients With Acute Low Back Pain: Results of a Pragmatic Randomized Comparative Effectiveness Study. Spine 38, 627. https://doi.org/10.1097/BRS.0b013e31827733e7

Goldstein, C.E., Weijer, C., Brehaut, J.C., Fergusson, D.A., Grimshaw, J.M., Horn, A.R., Taljaard, M., 2018. Ethical issues in pragmatic randomized controlled trials: a review of the recent literature identifies gaps in ethical argumentation. BMC Medical Ethics 19, 14. https://doi.org/10.1186/s12910-018-0253-x

González, Á.C., Berenguer, S.B., Luque Mañas, J.M., Martin-Pintado-Zugasti, A., 2021. Validation of a sham novel neural mobilization technique in patients with non-specific low back pain: A randomized, placebo-controlled trial. Musculoskeletal Science and Practice 53, 102378. https://doi.org/10.1016/j.msksp.2021.102378

González-Urzelai, V., Palacio-Elua, L., López-de-Munain, J., 2003. Routine primary care management of acute low back pain: adherence to clinical guidelines. Eur Spine J 12, 589–594. https://doi.org/10.1007/s00586-003-0567-2

Gooberman-Hill, R., Jinks, C., Bouças, S.B., Hislop, K., Dziedzic, K.S., Rhodes, C., Burston, A., Adams, J., 2013. Designing a placebo device: involving service users in clinical trial design. Health Expect 16, e100-110. https://doi.org/10.1111/hex.12043

Gordon, K.S., Peduzzi, P., Kerns, R.D., 2020a. Designing Trials with Purpose: Pragmatic Clinical Trials of Nonpharmacological Approaches for Pain Management. Pain Medicine 21, S7–S12. https://doi.org/10.1093/pm/pnaa347

Gordon, K.S., Peduzzi, P., Kerns, R.D., 2020b. Designing Trials with Purpose: Pragmatic Clinical Trials of Nonpharmacological Approaches for Pain Management. Pain Medicine 21, S7–S12. https://doi.org/10.1093/pm/pnaa347

Gough, D., Oliver, S., Thomas, J., 2017. An Introduction to Systematic Reviews. SAGE.

Greenhalgh, T., Hinton, L., Finlay, T., Macfarlane, A., Fahy, N., Clyde, B., Chant, A., 2019. Frameworks for supporting patient and public involvement in research: Systematic review and co-design pilot. Health Expectations 22, 785–801. https://doi.org/10.1111/hex.12888

Greenhouse, J.B., Kaizar, E.E., Kelleher, K., Seltman, H., Gardner, W., 2008. Generalizing from clinical trial data: A case study. The risk of suicidality among pediatric antidepressant users. Statistics in Medicine 27, 1801–1813. https://doi.org/10.1002/sim.3218

Grewal, G.S., Schwenk, M., Lee-Eng, J., Parvaneh, S., Bharara, M., Menzies, R.A., Talal, T.K., Armstrong, D.G., Najafi, B., 2015. Sensor-Based Interactive Balance Training with Visual Joint Movement Feedback for Improving Postural Stability in Diabetics with Peripheral Neuropathy: A Randomized Controlled Trial. GER 61, 567–574. https://doi.org/10.1159/000371846

Griswold, D., Learman, K., Kolber, M.J., O'Halloran, B., Cleland, J.A., 2018. Pragmatically Applied Cervical and Thoracic Nonthrust Manipulation Versus Thrust Manipulation for Patients With Mechanical Neck Pain: A Multicenter Randomized Clinical Trial. J Orthop Sports Phys Ther 48, 137–145. https://doi.org/10.2519/jospt.2018.7738

Grøvle, L., Hasvik, E., Haugen, A.J., 2020. Rescue and concomitant analgesics in placebo-controlled trials of pharmacotherapy for neuropathic pain and low back pain. PAIN 161, 3–10. https://doi.org/10.1097/j.pain.0000000000001690

Grøvle, L., Haugen, A.J., Keller, A., Natvig, B., Brox, J.I., Grotle, M., 2010. The bothersomeness of sciatica: patients' self-report of paresthesia, weakness and leg pain. Eur Spine J 19, 263–269. https://doi.org/10.1007/s00586-009-1042-5

Gu, A.P., Gu, C.N., Ahmed, A.T., Murad, M.H., Wang, Z., Kallmes, D.F., Brinjikji, W., 2017. Sham surgical procedures for pain intervention result in significant improvements in pain: systematic review and meta-analysis. Journal of Clinical Epidemiology 83, 18–23. https://doi.org/10.1016/j.jclinepi.2016.12.010

Guba, E.G., Lincoln, Y.S., 1994. Competing paradigms in qualitative research. Handbook of qualitative research 2, 105.

Guyatt, G., Oxman, A.D., Akl, E.A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., deBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., Schünemann, H.J., 2011. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology 64, 383–394. https://doi.org/10.1016/j.jclinepi.2010.04.026

Hall, S., Lewith, G., Brien, S., Little, P., 2008. An exploratory pilot study to design and assess the credibility of a sham kinesiology treatment. Forsch Komplementmed 15, 321–6. https://doi.org/10.1159/000170401

Hamid, N., 2011. Effects of stress management training on glycemic control in women with type 2 diabetes. Iranian Journal of Endocrinology and Metabolism 13, 346–353.

Harris, B., Duggan, M., Batterham, P., Bartlem, K., Clinton-McHarg, T., Dunbar, J., Fehily, C., Lawrence, D., Morgan, M., Rosenbaum, S., 2018. Australia's Mental and Physical Health Tracker: Background Paper. Australian Health Policy Collaboration, Melbourne, Australia.

Harris, P.A., Taylor, R., Minor, B.L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S.N., REDCap Consortium, 2019. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 95, 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P.A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., Conde, J.G., 2009. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 42, 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hart, T., Bagiella, E., 2012. Design and Implementation of Clinical Trials in Rehabilitation Research. Archives of Physical Medicine and Rehabilitation 93, S117–S126. https://doi.org/10.1016/j.apmr.2011.11.039

Harvey, E., Burton, A.K., Moffett, J.K., Breen, A., 2003. Spinal manipulation for low-back pain: a treatment package agreed by the UK chiropractic, osteopathy and physiotherapy professional associations. Manual Therapy 8, 46–51. https://doi.org/10.1054/math.2002.0472

Hasson, F., Keeney, S., McKenna, H., 2000. Research guidelines for the Delphi survey technique. Journal of Advanced Nursing 32, 1008–1015.

Hawe, P., Shiell, A., Riley, T., 2004. Complex interventions: how "out of control" can a randomised controlled trial be? BMJ 328, 1561–1563. https://doi.org/10.1136/bmj.328.7455.1561

Hawk, C., Minkalis, A.L., Khorsan, R., Daniels, C.J., Homack, D., Gliedt, J.A., Hartman, J.A., Bhalerao, S., 2017. Systematic Review of Nondrug, Nonsurgical Treatment of Shoulder Conditions. J Manipulative Physiol Ther 40, 293–319. https://doi.org/10.1016/j.jmpt.2017.04.001

Haxby Abbott [@Haxby_Abbott], 2023. interesting editorial but incorrect on one key point: there are at least 13 placebo/sham controlled trials of ex in #osteoarthritis although the findings aren't very convincing. Did the field move from efficacy to effectiveness too early? Maybe...½ https://doi.org/10.1016/j.ocarto.2021.100195. Twitter.

Hayden, J.A., 2020. Predatory publishing dilutes and distorts evidence in systematic reviews. Journal of Clinical Epidemiology 121, 117–119. https://doi.org/10.1016/j.jclinepi.2020.01.013

Haynes, B., 1999. Can it work? Does it work? Is it worth it?: The testing of healthcare interventions is evolving. BMJ 319, 652–653. https://doi.org/10.1136/bmj.319.7211.652

He, X., Li, J., Wang, B., Yao, Q., Li, L., Song, R., Shi, X., Zhang, J., 2017. Diabetes self-management education reduces risk of all-cause mortality in type 2 diabetes patients: a systematic review and meta-analysis. Endocrine 55, 712–731. https://doi.org/10.1007/s12020-016-1168-2

Hebert, J.J., Stomski, N.J., French, S.D., Rubinstein, S.M., 2015. Serious Adverse Events and Spinal Manipulative Therapy of the Low Back Region: A Systematic Review of Cases. Journal of Manipulative and Physiological Therapeutics, Special Issue: Adverse Events 38, 677–691. https://doi.org/10.1016/j.jmpt.2013.05.009

Hemming, K., Haines, T.P., Chilton, P.J., Girling, A.J., Lilford, R.J., 2015. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 350. https://doi.org/10.1136/bmj.h391

Hempel, S., Suttorp, M.J., Miles, J.N., Wang, Z., Maglione, M., Morton, S., Johnsen, B., Valentine, D., Shekelle, P.G., 2011. Empirical Evidence of Associations Between Trial Quality and Effect Size, AHRQ Methods for Effective Health Care. Agency for Healthcare Research and Quality (US), Rockville (MD).

Herman, W.H., Pop-Busui, R., Braffett, B.H., Martin, C.L., Cleary, P.A., Albers, J.W., Feldman, E.L., 2012. Use of the Michigan Neuropathy Screening Instrument as a measure of distal symmetrical peripheral neuropathy in Type 1 diabetes: results from the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications. Diabet Med 29, 937–944. https://doi.org/10.1111/j.1464-5491.2012.03644.x

Hernán, M.A., Robins, J.M., 2017. Per-Protocol Analyses of Pragmatic Trials. N Engl J Med 377, 1391–1398. https://doi.org/10.1056/NEJMsm1605385

Hernández-Secorún, M., Vidal-Peracho, C., Márquez-Gonzalvo, S., Corral-de-Toro, J., Müller-Thyssen-Uriarte, J., Rodríguez-Sanz, J., Lucha-López, M.O., Tricás-Moreno, J.M., Hidalgo-García, C., 2021. Exercise and Manual Therapy for Diabetic Peripheral Neuropathy: A Systematic Review. Applied Sciences 11, 5665. https://doi.org/10.3390/app11125665

Hicks, A., Fairhurst, C., Torgerson, D.J., 2018. A simple technique investigating baseline heterogeneity helped to eliminate potential bias in meta-analyses. Journal of Clinical Epidemiology 95, 55–62. https://doi.org/10.1016/j.jclinepi.2017.10.001

Higgins, J.P.T., Altman, D.G., Gotzsche, P.C., Juni, P., Moher, D., Oxman, A.D., Savovic, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., Cochrane Bias Methods Group, Cochrane Statistical Methods Group, 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 343, d5928–d5928. https://doi.org/10.1136/bmj.d5928

Higgins, J.P.T., Thomas, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (Eds.), 2019. Cochrane Handbook for Systematic Reviews of Interventions, Version 6.0, 2nd ed. John Wiley & Sons, Chichester, UK.

Hill-Briggs, F., Adler, N.E., Berkowitz, S.A., Chin, M.H., Gary-Webb, T.L., Navas-Acien, A., Thornton, P.L., Haire-Joshu, D., 2021. Social Determinants of Health and Diabetes: A Scientific Review. Diabetes Care 44, 258–279. https://doi.org/10.2337/dci20-0053

Hodkinson, A., Kontopantelis, E., Adeniji, C., van Marwijk, H., McMillian, B., Bower, P., Panagioti, M., 2021. Interventions Using Wearable Physical Activity Trackers Among Adults With Cardiometabolic Conditions: A Systematic Review and Meta-analysis. JAMA Netw Open 4, e2116382. https://doi.org/10.1001/jamanetworkopen.2021.16382

Hoffmann, T.C., Erueti, C., Glasziou, P.P., 2013. Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. BMJ 347, f3755. https://doi.org/10.1136/bmj.f3755

Hoffmann, T.C., Glasziou, P.P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D.G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S.E., Dixon-Woods, M., McCulloch, P., Wyatt, J.C., Chan, A.-W., Michie, S., 2014. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 348. https://doi.org/10.1136/bmj.g1687

Hohenschurz-Schmidt, D., Calcagnini, G., Dipasquale, O., Jackson, J.B., Medina, S., O'Daly, O., O'Muircheartaigh, J., de Lara Rubio, A., Williams, S.C.R., McMahon, S.B., Makovac, E., Howard, M.A., 2020a. Linking Pain Sensation to the Autonomic Nervous System: The Role

of the Anterior Cingulate and Periaqueductal Gray Resting-State Networks. Front. Neurosci. 14. https://doi.org/10.3389/fnins.2020.00147

Hohenschurz-Schmidt, D., Cherkin, D., Rice, A.S.C., Dworkin, R.H., Turk, D.C., McDermott, M.P., Bair, M.J., DeBar, L.L., Edwards, R.R., Farrar, J.T., Kerns, R.D., Markman, J.D., Rowbotham, M.C., Sherman, K.J., Wasan, A.D., Cowan, P., Desjardins, P., Ferguson, M., Freeman, R., Gewandter, J.S., Gilron, I., Grol-Prokopczyk, H., Hertz, S.H., Iyengar, S., Kamp, C., Karp, B.I., Kleykamp, B.A., Loeser, J.D., Mackey, S., Malamut, R., McNicol, E., Patel, K.V., Sandbrink, F., Schmader, K., Simon, L., Steiner, D.J., Veasley, C., Vollert, J., 2023a. Research objectives and general considerations for pragmatic clinical trials of pain treatments: IMMPACT statement. PAIN 10.1097/j.pain.0000000000002888. https://doi.org/10.1097/j.pain.0000000000002888

Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., 2022a. Dissimilar Control Interventions in Clinical Trials Undermine Interpretability. JAMA Psychiatry 79, 271–272. https://doi.org/10.1001/jamapsychiatry.2021.3963

Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., Scott, W., McGregor, A., Soliman, N., MacMillan, A., Olivier, A., Cherian, C.A., Corcoran, D., Abbey, H., Freigang, S., Chan, J., Phalip, J., Nørgaard Sørensen, L., Delafin, M., Baptista, M., Medforth, N.R., Ruffini, N., Skøtt Andresen, S., Ytier, S., Ali, D., Hobday, H., Santosa, A.A.N.A.A., Vollert, J., Rice, A.S.C., 2023b. Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article I): a systematic review and description of methods. PAIN 164, 469. https://doi.org/10.1097/j.pain.0000000000002723

Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., Scott, W., McGregor, A., Soliman, N., MacMillan, A., Olivier, A., Cherian, C.A., Corcoran, D., Abbey, H., Freigang, S., Chan, J., Phalip, J., Nørgaard Sørensen, L., Delafin, M., Baptista, M., Medforth, N.R., Ruffini, N., Skøtt Andresen, S., Ytier, S., Ali, D., Hobday, H., Santosa, A.A.N.A.A., Vollert, J., Rice, A.S.C., 2022b. Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article I): a systematic review and description of methods. PAIN 10.1097/j.pain.0000000000002723. https://doi.org/10.1097/j.pain.0000000000002723

Hohenschurz-Schmidt, D., Draper-Rodi, J., Vase, L., Scott, W., McGregor, A., Soliman, N., MacMillan, A., Olivier, A., Cherian, C.A., Corcoran, D., Abbey, H., Freigang, S., Chan, J., Phalip, J., Sørensen, L.N., Delafin, M., Baptista, M., Medforth, N.R., Ruffini, N., Andresen, S.S., Ytier, S., Ali, D., Hobday, H., Ngurah Agung Adhiyoga Santosa, A.A., Vollert, J., Rice, A.S.C., 2023c. Blinding and sham control methods in trials of physical, psychological, and self-management interventions for pain (article II): a meta-analysis relating methods to trial results. PAIN 164, 509. https://doi.org/10.1097/j.pain.0000000000002730

Hohenschurz-Schmidt, D., Kleykamp, B.A., Draper-Rodi, J., Vollert, J., Chan, J., Ferguson, M., McNicol, E., Phalip, J., Evans, S.R., Turk, D.C., Dworkin, R.H., Rice, A.S.C., 2022c. Pragmatic trials of pain therapies: a systematic review of methods. PAIN 163, 21. https://doi.org/10.1097/j.pain.0000000000002317

Hohenschurz-Schmidt, D., Scott, W., Park, C., Christopoulos, G., Vogel, S., Draper-Rodi, J., 2020b. Remote management of musculoskeletal pain: a pragmatic approach to the implementation of video and phone consultations in musculoskeletal practice. PAIN Reports 5, e878. https://doi.org/10.1097/PR9.0000000000000878

Hohenschurz-Schmidt, D., Thomson, O.P., Rossettini, G., Miciak, M., Newell, D., Roberts, L., Vase, L., Draper-Rodi, J., 2022d. Avoiding nocebo and other undesirable effects in chiropractic, osteopathy and physiotherapy: An invitation to reflect. Musculoskeletal Science & Practice 0. https://doi.org/10.1016/j.msksp.2022.102677

Hohenschurz-Schmidt, D., Vase, L., Scott, W., Annoni, M., Ajayi, O.K., Barth, J., Bennell, K., Berna, C., Bialosky, J., Braithwaite, F., Finnerup, N.B., Williams, A.C. de C., Carlino, E., Cerritelli, F., Chaibi, A., Cherkin, D., Colloca, L., Côté, P., Darnall, B.D., Evans, R., Fabre, L., Faria, V., French, S., Gerger, H., Häuser, W., Hinman, R.S., Ho, D., Janssens, T., Jensen, K., Johnston, C., Lunde, S.J., Keefe, F., Kerns, R.D., Koechlin, H., Kongsted, A., Michener, L.A., Moerman, D.E., Musial, F., Newell, D., Nicholas, M., Palermo, T.M., Palermo, S., Peerdeman, K.J., Pogatzki-Zahn, E.M., Puhl, A.A., Roberts, L., Rossettini, G., Matthiesen, S.T., Underwood, M., Vaucher, P., Vollert, J., Wartolowska, K., Weimer, K., Werner, C.P., Rice, A.S.C., Draper-Rodi, J., 2023d. Recommendations for the development, implementation, and reporting of control interventions in efficacy and mechanistic trials of physical, psychological, and self-management therapies: the CoPPS Statement. BMJ 381, e072108. https://doi.org/10.1136/bmj-2022-072108

Hohenschurz-Schmidt, D., Vollert, J., Vogel, S., Rice, A.S.C., Draper-Rodi, J., 2021. Performing and interpreting randomized clinical trials. Journal of Osteopathic Medicine 121, 443–445. https://doi.org/10.1515/jom-2020-0320

Holloway, I., Galvin, K., 2016. Qualitative Research in Nursing and Healthcare. John Wiley & Sons.

Holzer, K.J., Veasley, C., Kerns, R.D., Edwards, R.R., Gewandter, J.S., Langford, D.J., Yaeger, L.H., McNicol, E., Ferguson, M., Turk, D.C., Dworkin, R.H., Haroutounian, S., 2022. Partnering with patients in clinical trials of pain treatments: a narrative review. PAIN 163, 1862–1873. https://doi.org/10.1097/j.pain.0000000000002600

Howick, J., Webster, R.K., Rees, J.L., Turner, R., Macdonald, H., Price, A., Evers, A.W.M., Bishop, F., Collins, G.S., Bokelmann, K., Hopewell, S., Knottnerus, A., Lamb, S., Madigan, C., Napadow, V., Papanikitas, A.N., Hoffmann, T., 2020. TIDieR-Placebo: A guide and checklist for reporting placebo and sham controls. PLOS Medicine 17, e1003294. https://doi.org/10.1371/journal.pmed.1003294

Hróbjartsson, A., Emanuelsson, F., Skou Thomsen, A.S., Hilden, J., Brorson, S., 2014. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. Int J Epidemiol 43, 1272–1283. https://doi.org/10.1093/ije/dyu115

Hróbjartsson, A., Forfang, E., Haahr, M.T., Als-Nielsen, B., Brorson, S., 2007. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. Int J Epidemiol 36, 654–663. https://doi.org/10.1093/ije/dym020

Hróbjartsson, A., Gøtzsche, P.C., 2010. Placebo interventions for all clinical conditions. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.CD003974.pub3

Hróbjartsson, A., Miller, F., 2011. Chapter 2: Blinding in Nonpharmacological Randomized Controlled Trials, in: Boutron, I., Ravaud, P., Moher, D. (Eds.), Randomized Clinical Trials of Nonpharmacologic Treatments. Chapman and Hall/CRC, Boca Raton, FL, USA.

Hudson, R., Richmond, A., Sanchez, B., Stevenson, V., Baker, R.T., May, J., Nasypany, A., Reordan, D., 2018. Innovative treatment of clinically diagnosed meniscal tears: a randomized sham-

controlled trial of the Mulligan concept “squeeze” technique. Journal of Manual and Manipulative Therapy 26, 254–263. https://doi.org/10.1080/10669817.2018.1456614
Huebschmann, A.G., Leavitt, I.M., Glasgow, R.E., 2019. Making Health Research Matter: A Call to Increase Attention to External Validity. Annual Review of Public Health 40, 45–63. https://doi.org/10.1146/annurev-publhealth-040218-043945
Hung, A., Baas, C., Bekelman, J., Fitz-Randolph, M., Daniel Mullins, C., 2017. Patient and Stakeholder Engagement in Designing Pragmatic Clinical Trials, in: Birnbaum, H.G., Greenberg, P.E. (Eds.), Decision Making in a World of Comparative Effectiveness Research: A Practical Guide. Springer, Singapore, pp. 137–152. https://doi.org/10.1007/978-981-10-3262-2_11
Hutchinson, P., Moerman, D.E., 2018. The Meaning Response, “Placebo,” and Methods. Perspectives in Biology and Medicine 61, 361–378. https://doi.org/10.1353/pbm.2018.0049
Hutting, N., Caneiro, J.P., Ong’wen, O.M., Miciak, M., Roberts, L., 2022. Patient-centered care in musculoskeletal practice: Key elements to support clinicians to focus on the person. Musculoskeletal Science and Practice 57, 102434. https://doi.org/10.1016/j.msksp.2021.102434
Ilgen, M.A., Bohnert, A.S., Chermack, S., Conran, C., Jannausch, M., Trafton, J., Blow, F.C., 2016. A randomized trial of a pain management intervention for adults receiving substance use disorder treatment. Addiction (abingdon, england) 111, 1385-1393. https://doi.org/10.1111/add.13349
IMMPACT [WWW Document], n.d. URL http://www.immpact.org/index.html (accessed 4.16.20).
Institute of Medicine, 2009. Initial National Priorities for Comparative Effectiveness Research. National Academies Press, Washington, D.C. https://doi.org/10.17226/12648
International Committee of Medical Journal Editors, n.d. ICMJE Recommendations: Defining the Role of Authors and Contributors [WWW Document]. URL https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html (accessed 10.29.22).
International Diabetes Federation, 2019. IDF Diabetes Atlas 9th edition [WWW Document]. URL https://diabetesatlas.org/en/ (accessed 9.8.21).
Iqbal, S., Pipon-Young, L., 2009. Methods-The Delphi method–A guide from Susanne Iqbal and Laura Pipon-Young. Psychologist 22, 598.
Irving, E., van den Bor, R., Welsing, P., Walsh, V., Alfonso-Cristancho, R., Harvey, C., Garman, N., Grobbee, D.E., 2017. Series: Pragmatic trials and real world evidence: Paper 7. Safety, quality and monitoring. Journal of Clinical Epidemiology 91, 6–12. https://doi.org/10.1016/j.jclinepi.2017.05.004
Irving, M., Eramudugolla, R., Cherbuin, N., Anstey, K.J., 2017. A Critical Review of Grading Systems: Implications for Public Health Policy. Eval Health Prof 40, 244–262. https://doi.org/10.1177/0163278716645161
Izgu, N., Gok Metin, Z., Karadas, C., Ozdemir, L., Metinarikan, N., Corapcıoglu, D., 2020. Progressive Muscle Relaxation and Mindfulness Meditation on Neuropathic Pain, Fatigue, and Quality of Life in Patients With Type 2 Diabetes: A Randomized Clinical Trial. Journal of Nursing Scholarship 52, 476–487. https://doi.org/10.1111/jnu.12580

Jackson, W., Zale, E.L., Berman, S.J., Malacarne, A., Lapidow, A., Schatman, M.E., Kulich, R., Vranceanu, A.-M., 2019. Physical functioning and mindfulness skills training in chronic pain: a systematic review. J Pain Res 12, 179–189. https://doi.org/10.2147/JPR.S172733

Jahantigh Akbari, N., Hosseinifar, M., Naimi, S.S., Mikaili, S., Rahbar, S., 2020. The efficacy of physiotherapy interventions in mitigating the symptoms and complications of diabetic peripheral neuropathy: A systematic review. J Diabetes Metab Disord 19, 1995–2004. https://doi.org/10.1007/s40200-020-00652-8

James, K.E., Bloch, D.A., Lee, K.K., Kraemer, H.C., Fuller, R.K., 1996. An Index for Assessing Blindness in a Multi-Centre Clinical Trial: Disulfiram for Alcohol Cessation—a Va Cooperative Study. Statistics in Medicine 15, 1421–1434. https://doi.org/10.1002/(SICI)1097-0258(19960715)15:13<1421::AID-SIM266>3.0.CO;2-H

Javed, S., Petropoulos, I.N., Alam, U., Malik, R.A., 2015. Treatment of painful diabetic neuropathy. Therapeutic Advances in Chronic Disease 6, 15–28. https://doi.org/10.1177/2040622314552071

Johnson, S.B., Park, H.S., Gross, C.P., Yu, J.B., 2018. Complementary Medicine, Refusal of Conventional Cancer Therapy, and Survival Among Patients With Curable Cancers. JAMA Oncology 4, 1375–1381. https://doi.org/10.1001/jamaoncol.2018.2487

Jonas, W.B., 2005. Building an evidence house: challenges and solutions to research in complementary and alternative medicine. Forsch Komplementarmed Klass Naturheilkd 12, 159–167. https://doi.org/10.1159/000085412

Jones, K.F., Wechsler, S., Zulewski, D., Wood, L., 2022. Pharmacological and Nonpharmacological Management of Chemotherapy-Induced Peripheral Neuropathy: A Scoping Review of Randomized Controlled Trials. Journal of Palliative Medicine 25, 964–995. https://doi.org/10.1089/jpm.2021.0512

Jonsdottir, T., Jonsdottir, H., Lindal, E., Oskarsson, G.K., Gunnarsdottir, S., 2015. Predictors for chronic pain-related health care utilization: a cross-sectional nationwide study in Iceland. Health Expectations 18, 2704–2719. https://doi.org/10.1111/hex.12245

Jünger, S., Payne, S.A., Brine, J., Radbruch, L., Brearley, S.G., 2017. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. Palliat Med 31, 684–706. https://doi.org/10.1177/0269216317690685

Kalkman, S., van Thiel, G.J.M.W., Grobbee, D.E., van Delden, J.J.M., 2018. Pragmatic clinical trials: ethical imperatives and opportunities. Drug Discovery Today 23, 1919–1921. https://doi.org/10.1016/j.drudis.2018.06.007

Kalkman, S., van Thiel, G.J.M.W., Grobbee, D.E., van Delden, J.J.M., 2015. Pragmatic randomized trials in drug development pose new ethical questions: a systematic review. Drug Discovery Today 20, 856–862. https://doi.org/10.1016/j.drudis.2015.03.005

Kane, R.L., Wang, J., Garrard, J., 2007. Reporting in randomized clinical trials improved after adoption of the CONSORT statement. Journal of Clinical Epidemiology 60, 241–249. https://doi.org/10.1016/j.jclinepi.2006.06.016

Kaptchuk, T.J., Stason, W.B., Davis, R.B., Legedza, A.R.T., Schnyer, R.N., Kerr, C.E., Stone, D.A., Nam, B.H., Kirsch, I., Goldman, R.H., 2006. Sham device v inert pill: randomised controlled

trial of two placebo treatments. BMJ 332, 391–397. https://doi.org/10.1136/bmj.38726.603310.55

Karjalainen, T., Heikkinen, J., Busija, L., Jokihaara, J., Lewin, A.M., Naylor, J.M., Harris, L., Harris, I.A., Buchbinder, R., Adie, S., 2022. Use of Placebo and Nonoperative Control Groups in Surgical Trials: A Systematic Review and Meta-analysis. JAMA Network Open 5, e2223903. https://doi.org/10.1001/jamanetworkopen.2022.23903

Karran, E.L., Grant, A.R., Moseley, G.L., 2020. Low back pain and the social determinants of health: a systematic review and narrative synthesis. PAIN 161, 2476–2493. https://doi.org/10.1097/j.pain.0000000000001944

Kass, N., Faden, R., Tunis, S., 2012. Addressing Low-Risk Comparative Effectiveness Research in Proposed Changes to US Federal Regulations Governing Research. JAMA 307, 1589–1590. https://doi.org/10.1001/jama.2012.491

Kawchuk, G.N., Haugen, R., Fritz, J., 2009. A true blind for subjects who receive spinal manipulation therapy. Archives of Physical Medicine and Rehabilitation 90, 366–8. https://doi.org/10.1016/j.apmr.2008.08.213

Kec, D., Rajdova, A., Raputova, J., Adamova, B., Srotova, I., Nekvapilova, E.K., Michalcakova, R.N., Horakova, M., Belobradkova, J., Olsovsky, J., Weber, P., Hajas, G., Kaiserova, M., Mazanec, R., Potockova, V., Ehler, E., Forgac, M., Birklein, F., Üçeyler, N., Sommer, C., Bednarik, J., Vlckova, E., 2021. Risk factors for depression and anxiety in painful and painless diabetic polyneuropathy: A multicentre observational cross-sectional study. European Journal of Pain n/a. https://doi.org/10.1002/ejp.1865

Keefe, F.J., Jensen, M.P., Williams, A.C. de C., George, S.Z., 2022. The yin and yang of pragmatic clinical trials of behavioral interventions for chronic pain: balancing design features to maximize impact. PAIN 163, 1215–1219. https://doi.org/10.1097/j.pain.0000000000002546

Keefe, F.J., Main, C.J., George, S.Z., 2018. Advancing Psychologically Informed Practice for Patients With Persistent Musculoskeletal Pain: Promise, Pitfalls, and Solutions. Phys Ther 98, 398–407. https://doi.org/10.1093/ptj/pzy024

Kerns, R.D., Brandt, C.A., Peduzzi, P., for the NIH-DoD-VA Pain Management Collaboratory, 2019. NIH-DoD-VA Pain Management Collaboratory. Pain Medicine 20, 2336–2345. https://doi.org/10.1093/pm/pnz186

Kerns, R.D., Davis, A.F., Fritz, J.M., Keefe, F.J., Peduzzi, P., Rhon, D.I., Taylor, S.L., Vining, R., Yu, Q., Zeliadt, S.B., George, S.Z., 2022. Intervention Fidelity in Pain Pragmatic Trials for Nonpharmacologic Pain Management: Nuanced Considerations for Determining PRECIS-2 Flexibility in Delivery and Adherence. The Journal of Pain 0. https://doi.org/10.1016/j.jpain.2022.12.008

Kerr, F., Wiechula, R., Feo, R., Schultz, T., Kitson, A., 2019. Neurophysiology of human touch and eye gaze in therapeutic relationships and healing: a scoping review. JBI Database System Rev Implement Rep 17, 209–247. https://doi.org/10.11124/JBISRIR-2017-003549

Keshet, Y., 2009. The untenable boundaries of biomedical knowledge: epistemologies and rhetoric strategies in the debate over evaluating complementary and alternative medicine. Health: 13, 131–155.

Khodneva, Y., Richman, J., Andreae, S., Cherrington, A., Safford, M.M., 2020. Peer Support Intervention Improves Pain-Related Outcomes Among Rural Adults With Diabetes and

Chronic Pain at 12-Month Follow-Up. The Journal of rural health : official journal of the American Rural Health Association and the National Rural Health Care Association. https://doi.org/10.1111/jrh.12422
Kiegerl, G., 2007. Does Osteopathy Influence Diabetes Mellitus Type II? [WWW Document].
King, H.H., 2016. Manual Therapy Shown To Improve Diabetic Foot Ulcer Healing. Journal of Osteopathic Medicine 116, 685–686. https://doi.org/10.7556/jaoa.2016.135
Kioskli, K., Scott, W., Winkley, K., Godfrey, E., McCracken, L., 2020. Online Acceptance and Commitment Therapy for people with painful diabetic neuropathy in the United Kingdom: A single-arm feasibility trial. Pain Med 1–35.
Kioskli, K., Scott, W., Winkley, K., Kylakos, S., McCracken, L.M., 2019. Psychosocial Factors in Painful Diabetic Neuropathy: A Systematic Review of Treatment Trials and Survey Studies. Pain Medicine 20, 1756–1773. https://doi.org/10.1093/pm/pnz071
Kipping, K., Maier, C., Bussemas, H.H., Schwarzer, A., 2014. Medication compliance in patients with chronic pain. Pain Physician 17, 81–94.
Kleijnen, J., 2020. Re: Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. BMJ 368. https://doi.org/10.1136/bmj.l6802
Kligler, B., Bair, M.J., Banerjea, R., DeBar, L., Ezeji-Okoye, S., Lisi, A., Murphy, J.L., Sandbrink, F., Cherkin, D.C., 2018. Clinical Policy Recommendations from the VHA State-of-the-Art Conference on Non-Pharmacological Approaches to Chronic Musculoskeletal Pain. J GEN INTERN MED 33, 16–23. https://doi.org/10.1007/s11606-018-4323-z
Kluzek, S., Dean, B., Wartolowska, K.A., 2022. Patient-reported outcome measures (PROMs) as proof of treatment efficacy. BMJ Evidence-Based Medicine 27, 153–155. https://doi.org/10.1136/bmjebm-2020-111573
Koleck, T.A., Dreisbach, C., Bourne, P.E., Bakken, S., 2019. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. Journal of the American Medical Informatics Association 26, 364–379. https://doi.org/10.1093/jamia/ocy173
Koog, Y.H., Jung, W.Y., 2013. Time Course of Placebo Effect of Acupuncture on Pain: A Systematic Review. ISRN Pain 2013, 204108. https://doi.org/10.1155/2013/204108
Krashin, D., Murinova, N., Sullivan, M., 2016. Challenges to Treatment of Chronic Pain and Addiction During the "Opioid Crisis." Curr Pain Headache Rep 20, 65. https://doi.org/10.1007/s11916-016-0596-2
Kratochwill, T.R., Hitchcock, J., Horner, R.H., Levin, J.R., Odom, S.L., Rindskopf, D.M., Shadish, W.R., 2010. Single-case designs technical documentation. What works clearinghouse.
Kroenke, K., Krebs, E.E., Turk, D., Von Korff, M., Bair, M.J., Allen, K.D., Sandbrink, F., Cheville, A.L., DeBar, L., Lorenz, K.A., Kerns, R.D., 2019. Core Outcome Measures for Chronic Musculoskeletal Pain Research: Recommendations from a Veterans Health Administration Work Group. Pain Medicine 20, 1500–1508. https://doi.org/10.1093/pm/pny279
Kruse, C.S., Kristof, C., Jones, B., Mitchell, E., Martinez, A., 2016. Barriers to Electronic Health Record Adoption: a Systematic Literature Review. J Med Syst 40, 252. https://doi.org/10.1007/s10916-016-0628-9
Kuhmann, O., 2008. The impact of osteopathic treatment on intraocular hypertension – an experimental study [WWW Document].

Kumar, S., Beaton, K., Hughes, T., 2013. The effectiveness of massage therapy for the treatment of nonspecific low back pain: a systematic review of systematic reviews. Int J Gen Med 6, 733–741. https://doi.org/10.2147/IJGM.S50243

Kumar, S.P., 2011. Efficacy of segmental stabilization exercise for lumbar segmental instability in patients with mechanical low back pain: A randomized placebo controlled crossover study. North American Journal of Medical Sciences 3, 456–461. https://doi.org/10.4297/najms.2011.3456

Kutner, J.S., Smith, M.C., Corbin, L., Hemphill, L., Benton, K., Mellis, B.K., Beaty, B., Felton, S., Yamashita, T.E., Bryant, L.L., Fairclough, D.L., 2008. Massage Therapy versus Simple Touch to Improve Pain and Mood in Patients with Advanced Cancer. Annals of Internal Medicine 149, 369–379. https://doi.org/10.7326/0003-4819-149-6-200809160-00003

Ladd, B.O., McCrady, B.S., Manuel, J.K., Campbell, W., 2010. Improving the quality of reporting alcohol outcome studies: Effects of the CONSORT statement. Addictive Behaviors 35, 660–666. https://doi.org/10.1016/j.addbeh.2010.02.009

Ladeira, C.E., Cheng, M.S., da Silva, R.A., 2017. Clinical Specialization and Adherence to Evidence-Based Practice Guidelines for Low Back Pain Management: A Survey of US Physical Therapists. J Orthop Sports Phys Ther 47, 347–358. https://doi.org/10.2519/jospt.2017.6561

Lähdeoja, T., Karjalainen, T., Jokihaara, J., Salamh, P., Kavaja, L., Agarwal, A., Winters, M., Buchbinder, R., Guyatt, G., Vandvik, P.O., Ardern, C.L., 2020. Subacromial decompression surgery for adults with shoulder pain: a systematic review with meta-analysis. Br J Sports Med 54, 665–673. https://doi.org/10.1136/bjsports-2018-100486

Lantos, J.D., Wendler, D., Septimus, E., Wahba, S., Madigan, R., Bliss, G., 2015. Considerations in the evaluation and determination of minimal risk in pragmatic clinical trials. Clinical Trials 12, 485–493. https://doi.org/10.1177/1740774515597687

Lau, M.A., McMain, S.F., 2005. Integrating Mindfulness Meditation with Cognitive and Behavioural Therapies: The Challenge of Combining Acceptance- and Change-Based Strategies. Can J Psychiatry 50, 863–869. https://doi.org/10.1177/070674370505001310

Lavazza, C., Galli, M., Abenavoli, A., Maggiani, A., 2021. Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pairwise meta-analysis. BMJ Open 11, e045106. https://doi.org/10.1136/bmjopen-2020-045106

Leach, M.J., 2006. Evidence-based practice: A framework for clinical practice and research design. International Journal of Nursing Practice 12, 248–251. https://doi.org/10.1111/j.1440-172X.2006.00587.x

Lee, C.-J.G., 2012. Reconsidering Constructivism in Qualitative Research. Educational Philosophy and Theory 44, 403–412. https://doi.org/10.1111/j.1469-5812.2010.00720.x

Lee, K., Lewis, G.N., 2020. Short term relief of multisite chronicpain with Bowen Therapy: A double-blind, randomized controlled trial. Journal of Bodywork and Movement Therapies 24, 271–279. https://doi.org/10.1016/j.jbmt.2020.06.025

Lehtola, V., Korhonen, I., Airaksinen, O., 2010. A randomised, placebo-controlled, clinical trial for the short-term effectiveness of manipulative therapy and acupuncture on pain caused by mechanical thoracic spine dysfunction. International Musculoskeletal Medicine 32, 25–32. https://doi.org/10.1179/175361410X12652805807558

Levy, C., Zimmerman, S., Mor, V., Gifford, D., Greenberg, S.A., Klinger, J.H., Lieblich, C., Linnebur, S., McAllister, A., Nazir, A., Pace, D., Stone, R., Resnick, B., Sloane, P.D., Ouslander, J., Gaugler, J.E., 2022. Pragmatic Trials in Long-Term Care: Implementation and Dissemination Challenges and Opportunities. J Am Med Dir Assoc 23, 339–344. https://doi.org/10.1016/j.jamda.2021.11.024

Lewis, C., Khan, A., Souvlis, T., Sterling, M., 2010. A randomised controlled study examining the short-term effects of Strain-Counterstrain treatment on quantitative sensory measures at digitally tender points in the low back. Manual Therapy 15, 536–41. https://doi.org/10.1016/j.math.2010.05.011

Lewis, R.J., 2016. The pragmatic clinical trial in a learning health care system. Clinical Trials 13, 484–492. https://doi.org/10.1177/1740774516655097

Liampas, A., Rekatsina, M., Vadalouca, A., Paladini, A., Varrassi, G., Zis, P., 2020. Non-Pharmacological Management of Painful Peripheral Neuropathies: A Systematic Review. Adv Ther 37, 4096–4106. https://doi.org/10.1007/s12325-020-01462-3

Licciardone, J.C., Kearns, C.M., Hodge, L.M., Minotti, D.E., 2013. Osteopathic Manual Treatment in Patients With Diabetes Mellitus and Comorbid Chronic Low Back Pain: Subgroup Results From the OSTEOPATHIC Trial. Journal of Osteopathic Medicine 113, 468–478. https://doi.org/10.7556/jaoa.2013.113.6.468

Liem, T., 2018. Von A.T. Stills Theorie der osteopathischen Läsion zur somatischen Dysfunktion: Geschichtliche Entwicklung und evidenzbasierte Modelle, die das aktuelle Konzept der somatischen Dysfunktion unterstützen. Osteopathische Medizin 19, 21–26. https://doi.org/10.1016/S1615-9071(18)30083-2

Little, R.J.A., D'Agostino, R., Dickersin, K., Emerson, S.S., Farrar, J.T., Frangakis, C., Hogan, J.W., Molenberghs, G., Murphy, S.A., Neaton, J.D., Rotnitzky, A., Scharfstein, D., Shih, W., Siegel, J.P., Stern, H., 2010. The Prevention and Treatment of Missing Data in Clinical Trials. Pannel on Handling Missing Data in Clinical Trials. The National Academies Press; Washington, D.C.

Locock, L., Boaz, A., 2019. Drawing straight lines along blurred boundaries: Qualitative research, patient and public involvement in medical research, coproduction and co-design. Evidence & Policy 15, 409–421. https://doi.org/10.1332/174426419X15552999451313

Long, Y., Chen, R., Guo, Q., Luo, S., Huang, J., Du, L., 2020. Do acupuncture trials have lower risk of bias over the last five decades? A methodological study of 4 715 randomized controlled trials. PLOS ONE 15, e0234491. https://doi.org/10.1371/journal.pone.0234491

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K.E., Zwarenstein, M., 2015a. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 350, h2147. https://doi.org/10.1136/bmj.h2147

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K.E., Zwarenstein, M., 2015b. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 350. https://doi.org/10.1136/bmj.h2147

Luce, B.R., Kramer, J.M., Goodman, S.N., Connor, J.T., Tunis, S., Whicher, D., Schwartz, J.S., 2009. Rethinking Randomized Clinical Trials for Comparative Effectiveness Research: The Need for Transformational Change. Ann Intern Med 151, 206–209. https://doi.org/10.7326/0003-4819-151-3-200908040-00126

Lynch, M.E., Campbell, F., Clark, A.J., Dunbar, M.J., Goldstein, D., Peng, P., Stinson, J., Tupper, H., 2008. A systematic review of the effect of waiting for treatment for chronic pain. PAIN 136, 97–116. https://doi.org/10.1016/j.pain.2007.06.018

Ma, B., Chen, Z., Xu, J., Wang, Ya-nan, Chen, K., Ke, F., Niu, J., Li, L., Huang, C., Zheng, J., Yang, J., Zhu, Q., Wang, Ya-ping, 2016. Do the CONSORT and STRICTA Checklists Improve the Reporting Quality of Acupuncture and Moxibustion Randomized Controlled Trials Published in Chinese Journals? A Systematic Review and Analysis of Trends. PLOS ONE 11, e0147244. https://doi.org/10.1371/journal.pone.0147244

Machado, L.A.C., Kamper, S.J., Herbert, R.D., Maher, C.G., McAuley, J.H., 2008. Imperfect placebos are common in low back pain trials: a systematic review of the literature. Eur Spine J 17, 889–904. https://doi.org/10.1007/s00586-008-0664-3

MacPherson, H., 2004. Pragmatic clinical trials. Complementary Therapies in Medicine 12, 136–140. https://doi.org/10.1016/j.ctim.2004.07.043

Maddox, C.D., Subialka, J.A., Young, J.L., Rhon, D.I., 2022. Over Half of Clinical Trials of Mobilization and Manipulation for Patients with Low Back Pain May Have Limited Real-World Applicability. A Systematic Review pf 132 Clinical Trials. J Orthop Sports Phys Ther 1–72. https://doi.org/10.2519/jospt.2022.10962

Maggard-Gibbons, M., Maglione, M., Livhits, M., Ewing, B., Maher, A.R., Hu, J., Li, Z., Shekelle, P.G., 2013. Bariatric Surgery for Weight Loss and Glycemic Control in Nonmorbidly Obese Adults With Diabetes: A Systematic Review. JAMA 309, 2250–2261. https://doi.org/10.1001/jama.2013.4851

Magnusson, K., 2021. Interpreting Cohen's d effect size: An interactive viusualization (Version 2.5.0) [Web App] [WWW Document]. R Psychologist. URL https://rpsychologist.com/cohend/ (accessed 6.30.21).

Malliaras, P., Merolli, M., Williams, C.M., Caneiro, J.P., Haines, T., Barton, C., 2021. 'It's not hands-on therapy, so it's very limited': Telehealth use and views among allied health clinicians during the coronavirus pandemic. Musculoskeletal Science and Practice 52, 102340. https://doi.org/10.1016/j.msksp.2021.102340

Mallinckrodt, C., Molenberghs, G., Rathmann, S., 2017. Choosing estimands in clinical trials with missing data. Pharmaceutical Statistics 16, 29–36. https://doi.org/10.1002/pst.1765

Markman, J.D., Gewandter, J.S., Frazer, M.E., 2020. Comparison of a Pain Tolerability Question With the Numeric Rating Scale for Assessment of Self-reported Chronic Pain. JAMA Netw Open 3, e203155. https://doi.org/10.1001/jamanetworkopen.2020.3155

Marya, R., Patel, R., 2021. Inflamed: Deep Medicine and the Anatomy of Injustice, 1st edition. ed. Allen Lane, London, UK.

Mason, S., Tovey, P., Long, A.F., 2002. Evaluating complementary medicine: methodological challenges of randomised controlled trials. BMJ 325, 832–834. https://doi.org/10.1136/bmj.325.7368.832

Masoudi, F.A., Havranek, E.P., Wolfe, P., Gross, C.P., Rathore, S.S., Steiner, J.F., Ordin, D.L., Krumholz, H.M., 2003. Most hospitalized older persons do not meet the enrollment criteria for clinical trials in heart failure. American Heart Journal 146, 250–257. https://doi.org/10.1016/S0002-8703(03)00189-3

Mataix-Cols, D., Andersson, E., 2021. Ten Practical Recommendations for Improving Blinding Integrity and Reporting in Psychotherapy Trials. JAMA Psychiatry 78, 943–944. https://doi.org/10.1001/jamapsychiatry.2021.1419

Mathias, S.D., Crosby, R.D., Qian, Y., Jiang, Q., Dansey, R., Chung, K., 2011. Estimating Minimally Important Differences for the Worst Pain Rating of the Brief Pain Inventory–Short Form. The Journal of Supportive Oncology 9, 72–78. https://doi.org/10.1016/j.suponc.2010.12.004

May, M., 2019. Twenty-five ways clinical trials have changed in the last 25 years. Nature Medicine 25, 2–5. https://doi.org/10.1038/s41591-018-0314-1

Mbuagbaw, L., Aves, T., 2022. Meta-analysis of Pragmatic and Explanatory TrialsExplanatory trials, in: Evangelou, E., Veroniki, A.A. (Eds.), Meta-Research: Methods and Protocols, Methods in Molecular Biology. Springer US, New York, NY, pp. 147–158. https://doi.org/10.1007/978-1-0716-1566-9_9

McCarthy, D.M., Waite, K.R., Curtis, L.M., Engel, K.G., Baker, D.W., Wolf, M.S., 2012. What Did the Doctor Say? Health Literacy and Recall of Medical Instructions. Med Care 50, 277–282. https://doi.org/10.1097/MLR.0b013e318241e8e1

McClintock, A.S., McCarrick, S.M., Garland, E.L., Zeidan, F., Zgierska, A.E., 2019. Brief Mindfulness-Based Interventions for Acute and Chronic Pain: A Systematic Review. The Journal of Alternative and Complementary Medicine 25, 265–278. https://doi.org/10.1089/acm.2018.0351

McGlone, F., Cerritelli, F., Walker, S., Esteves, J., 2017. The role of gentle touch in perinatal osteopathic manual therapy. Neuroscience & Biobehavioral Reviews 72, 1–9. https://doi.org/10.1016/j.neubiorev.2016.11.009

McGuire, H., Longson, D., Adler, A., Farmer, A., Lewin, I., 2016. Management of type 2 diabetes in adults: summary of updated NICE guidance. BMJ 353, i1575. https://doi.org/10.1136/bmj.i1575

McGuirk, J., 2012. The place of touch in Counselling and Psychotherapy and the potential for healing within the therapeutic relationship. Inside Out 68.

McHugh, G.A., Luker, K.A., 2012. Individuals' expectations and challenges following total hip replacement: a qualitative study. Disability and Rehabilitation 34, 1351–1357. https://doi.org/10.3109/09638288.2011.644022

McPherson, S., Reese, C., Wendler, M.C., 2018. Methodology Update: Delphi Studies. Nursing Research 67, 404. https://doi.org/10.1097/NNR.0000000000000297

Meisinger, C., Bongaerts, B.W.C., Heier, M., Amann, U., Kowall, B., Herder, C., Rückert-Eheberg, I.-M., Rathmann, W., Ziegler, D., 2018. Neuropathic pain is not adequately treated in the older general population: Results from the KORA F4 survey. Pharmacoepidemiol Drug Saf 27, 806–814. https://doi.org/10.1002/pds.4559

Meissner, K., Fässler, M., Rücker, G., Kleijnen, J., Hróbjartsson, A., Schneider, A., Antes, G., Linde, K., 2013. Differential Effectiveness of Placebo Treatments: A Systematic Review of Migraine Prophylaxis. JAMA Intern Med 173, 1941–1951. https://doi.org/10.1001/jamainternmed.2013.10391

Meissner, K., Linde, K., 2018. Chapter Fourteen - Are Blue Pills Better Than Green? How Treatment Features Modulate Placebo Effects, in: Colloca, L. (Ed.), International Review

of Neurobiology, Neurobiology of the Placebo Effect Part II. Academic Press, pp. 357–378. https://doi.org/10.1016/bs.irn.2018.07.014

Meredith, S., Ward, M., Booth, G., Fisher, A., Gamble, C., House, H., Landray, M., 2011. Risk-adapted approaches to the management of clinical trials: Guidance from the department of health (DH)/medical research council (MRC)/medicines and healthcare products regulatory agency (MHRA) clinical trials working group. Trials 12, A39.

Messier, S.P., Mihalko, S.L., Beavers, D.P., Nicklas, B.J., DeVita, P., Carr, J.J., Hunter, D.J., Lyles, M., Guermazi, A., Bennell, K.L., Loeser, R.F., 2021. Effect of High-Intensity Strength Training on Knee Pain and Knee Joint Compressive Forces Among Adults With Knee Osteoarthritis: The START Randomized Clinical Trial. JAMA 325, 646–657. https://doi.org/10.1001/jama.2021.0411

Michener, L.A., Kardouni, J.R., Sousa, C.O., Ely, J.M., 2015. Validation of a sham comparator for thoracic spinal manipulation in patients with shoulder pain. Manual Therapy 20, 171–5. https://doi.org/10.1016/j.math.2014.08.008

Miciak, M., Gross, D.P., Joyce, A., 2012. A review of the psychotherapeutic 'common factors' model and its application in physical therapy: the need to consider general effects in physical therapy practice. Scandinavian Journal of Caring Sciences 26, 394–403. https://doi.org/10.1111/j.1471-6712.2011.00923.x

Miller, F.G., 2004. Sham surgery: An ethical analysis. SCI ENG ETHICS 10, 157–166. https://doi.org/10.1007/s11948-004-0073-x

Miller, F.G., Crouch, R.A., 2008. The ethics of placebo-controlled trials, in: Emanuel, E.J., Grady, C.C., Crouch, R.A., Lie, R.K., Miller, F.G., Wendler, D.D. (Eds.), The Oxford Textbook of Clinical Research Ethics. Oxford University Press, pp. 261–272.

Mills, D., Inouye, K., 2021. Problematizing 'predatory publishing': A systematic review of factors shaping publishing motives, decisions, and experiences. Learned Publishing 34, 89–104. https://doi.org/10.1002/leap.1325

Millum, J., Grady, C., 2013. The ethics of placebo-controlled trials: Methodological justifications. Contemporary Clinical Trials 36, 510–514. https://doi.org/10.1016/j.cct.2013.09.003

Moerman, D.E., 2013. Against the "placebo effect": A personal point of view. Complementary Therapies in Medicine, The Placebo Effect and its Ramifications for Clinical Practice and Research: An Expert Workshop, Lake Como, Italy, 4-6 May 2012 21, 125–130. https://doi.org/10.1016/j.ctim.2013.01.005

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., Altman, D.G., 2010a. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. Journal of Clinical Epidemiology 63, e1–e37. https://doi.org/10.1016/j.jclinepi.2010.03.004

Moher, D., Schulz, K.F., Simera, I., Altman, D.G., 2010b. Guidance for Developers of Health Research Reporting Guidelines. PLOS Medicine 7, e1000217. https://doi.org/10.1371/journal.pmed.1000217

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., PRISMA-P Group, 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Reviews 4, 1. https://doi.org/10.1186/2046-4053-4-1

Mohr, D.C., Ho, J., Hart, T.L., Baron, K.G., Berendsen, M., Beckner, V., Cai, X., Cuijpers, P., Spring, B., Kinsinger, S.W., Schroder, K.E., Duffecy, J., 2014. Control condition design and implementation features in controlled trials: a meta-analysis of trials evaluating psychotherapy for depression. Transl Behav Med 4, 407–423. https://doi.org/10.1007/s13142-014-0262-3

Moisset, X., Bouhassira, D., Avez Couturier, J., Alchaar, H., Conradi, S., Delmotte, M.H., Lanteri-Minet, M., Lefaucheur, J.P., Mick, G., Piano, V., Pickering, G., Piquet, E., Regis, C., Salvat, E., Attal, N., 2020. Pharmacological and non-pharmacological treatments for neuropathic pain: Systematic review and French recommendations. Rev Neurol (Paris) 176, 325–352. https://doi.org/10.1016/j.neurol.2020.01.361

Molassiotis, A., 2013. Evidence is in the eye of the beholder. Support Care Cancer 21, 3259–3260. https://doi.org/10.1007/s00520-013-1929-x

Molsted, S., Tribler, J., Snorgaard, O., 2012. Musculoskeletal pain in patients with type 2 diabetes. Diabetes Research and Clinical Practice 96, 135–140. https://doi.org/10.1016/j.diabres.2011.12.022

Montedori, A., Bonacini, M.I., Casazza, G., Luchetta, M.L., Duca, P., Cozzolino, F., Abraha, I., 2011. Modified versus standard intention-to-treat reporting: Are there differences in methodological quality, sponsorship, and findings in randomized trials? A cross-sectional study. Trials 12, 58. https://doi.org/10.1186/1745-6215-12-58

Moore, R.A., Derry, S., McQuay, H.J., Straube, S., Aldington, D., Wiffen, P., Bell, R.F., Kalso, E., Rowbotham, M.C., 2010. Clinical effectiveness: An approach to clinical trial design more relevant to clinical practice, acknowledging the importance of individual differences. PAIN® 149, 173–176. https://doi.org/10.1016/j.pain.2009.08.007

Moraska, A.F., Schmiege, S.J., Mann, J.D., Butryn, N., Krutsch, J.P., 2017. Responsiveness of Myofascial Trigger Points to Single and Multiple Trigger Point Release Massages: A Randomized, Placebo Controlled Trial. Am J Phys Med Rehabil 96, 639–645. https://doi.org/10.1097/PHM.0000000000000728

Moroz, A., Freed, B., Tiedemann, L., Bang, H., Howell, M., Park, J.J., 2013. Blinding Measured: A Systematic Review of Randomized Controlled Trials of Acupuncture [WWW Document]. Evidence-Based Complementary and Alternative Medicine. https://doi.org/10.1155/2013/708251

Moseley, A.M., Elkins, M.R., Van der Wees, P.J., Pinheiro, M.B., 2020. Using research to guide practice: The Physiotherapy Evidence Database (PEDro). Brazilian Journal of Physical Therapy 24, 384–391. https://doi.org/10.1016/j.bjpt.2019.11.002

Moustgaard, H., Clayton, G.L., Jones, H.E., Boutron, I., Jørgensen, L., Laursen, D.L.T., Olsen, M.F., Paludan-Müller, A., Ravaud, P., Savović, J., Sterne, J.A.C., Higgins, J.P.T., Hróbjartsson, A., 2020. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. BMJ 368. https://doi.org/10.1136/bmj.l6802

Movsas, B., Moughan, J., Owen, J., Coia, L.R., Zelefsky, M.J., Hanks, G., Wilson, J.F., 2007. Who Enrolls Onto Clinical Oncology Trials? A Radiation Patterns of Care Study Analysis. International Journal of Radiation Oncology*Biology*Physics 68, 1145–1150. https://doi.org/10.1016/j.ijrobp.2007.01.051

Musial, F., 2019. Acupuncture for the Treatment of Pain – A Mega-Placebo? Front. Neurosci. 13. https://doi.org/10.3389/fnins.2019.01110

Nagendran, M., Harding, D., Teo, W., Camm, C., Maruthappu, M., McCulloch, P., Hopewell, S., 2013. Poor adherence of randomised trials in surgery to CONSORT guidelines for non-pharmacological treatments (NPT): a cross-sectional study. BMJ Open 3, e003898. https://doi.org/10.1136/bmjopen-2013-003898

Naranjo, C., Del Reguero, L., Moratalla, G., Hercberg, M., Valenzuela, M., Failde, I., 2019. Anxiety, depression and sleep disorders in patients with diabetic neuropathic pain: a systematic review. Expert Review of Neurotherapeutics 19, 1201–1209. https://doi.org/10.1080/14737175.2019.1653760

Nathan, H.J., Poulin, P., Wozny, D., Taljaard, M., Smyth, C., Gilron, I., Sorisky, A., Lochnan, H., Shergill, Y., 2017. Randomized Trial of the Effect of Mindfulness-Based Stress Reduction on Pain-Related Disability, Pain Intensity, Health-Related Quality of Life, and A1C in Patients With Painful Diabetic Peripheral Neuropathy. Clinical Diabetes 35, 294–304. https://doi.org/10.2337/cd17-0077

National Institute for Health and Care Excellence, 2009. Medicines adherence: involving patients in decisions about prescribed medicines and supporting adherence, Clinical guideline [CG76] [WWW Document]. URL https://www.nice.org.uk/Guidance/CG76 (accessed 2.19.21).

National Institute for Health and Care Research, n.d. UK Standards for Public Involvement [WWW Document]. URL https://sites.google.com/nihr.ac.uk/pi-standards/standards (accessed 2.21.23).

Ng, J.Y., Boon, H.S., Thompson, A.K., Whitehead, C.R., 2016. Making sense of "alternative", "complementary", "unconventional" and "integrative" medicine: exploring the terms and meanings through a textual analysis. BMC Complement Altern Med 16, 134. https://doi.org/10.1186/s12906-016-1111-3

NICE, 2016. Low back pain and sciatica in over 16s: assessment and management (NICE guideline NG59), Assessment and non-invasive treatments: Methods, evidence and recommendations.

NICE, 2015. Overview | Type 2 diabetes in adults: management | Guidance | NICE [WWW Document]. URL https://www.nice.org.uk/guidance/ng28 (accessed 9.8.21).

Nicholas, M.K., 2022. The biopsychosocial model of pain 40 years on: time for a reappraisal? PAIN 10.1097/j.pain.0000000000002654. https://doi.org/10.1097/j.pain.0000000000002654

Nicholas, M.K., Asghari, A., Blyth, F.M., Wood, B.M., Murray, R., McCabe, R., Brnabic, A., Beeston, L., Corbett, M., Sherrington, C., Overton, S., 2013. Self-management intervention for chronic pain in older adults: A randomised controlled trial. PAIN® 154, 824–835. https://doi.org/10.1016/j.pain.2013.02.009

Nicholls, D.A., 2017. The End of Physiotherapy. Routledge, London. https://doi.org/10.4324/9781315561868

Nicholls, S.G., Carroll, K., Zwarenstein, M., Brehaut, J.C., Weijer, C., Hey, S.P., Goldstein, C.E., Graham, I.D., Grimshaw, J.M., McKenzie, J.E., Fergusson, D.A., Taljaard, M., on behalf of the Ethics of Pragmatic Trials project, 2019. The ethical challenges raised in the design and conduct of pragmatic trials: an interview study with key stakeholders. Trials 20, 765. https://doi.org/10.1186/s13063-019-3899-x

Nicholls, S.G., Zwarenstein, M., Hey, S.P., Giraudeau, B., Campbell, M.K., Taljaard, M., 2020. The importance of decision intent within descriptions of pragmatic trials. Journal of Clinical Epidemiology 125, 30–37. https://doi.org/10.1016/j.jclinepi.2020.04.030

Nielsen, J., 1994. Enhancing the explanatory power of usability heuristics, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 152–158.

Nielsen, J., Molich, R., 1990. Heuristic evaluation of user interfaces, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '90. Association for Computing Machinery, New York, NY, USA, pp. 249–256. https://doi.org/10.1145/97243.97281

Niemeijer, A., Lund, H., Stafne, S.N., Ipsen, T., Goldschmidt, C.L., Jørgensen, C.T., Juhl, C.B., 2020. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. Br J Sports Med 54, 1073–1080. https://doi.org/10.1136/bjsports-2018-100461

NIHR Oxford Biomedical Research Centre, n.d. Section 2 – What is Patient and Public Involvement?, in: PPI Researcher Guidance.

Noel-Storr, A., 2019. Working with a new kind of team: harnessing the wisdom of the crowd in trial identification. EFSA Journal 17, e170715. https://doi.org/10.2903/j.efsa.2019.e170715

Noel-Storr, A.H., Redmond, P., Lamé, G., Liberati, E., Kelly, S., Miller, L., Dooley, G., Paterson, A., Burt, J., 2021. Crowdsourcing citation-screening in a mixed-studies systematic review: a feasibility study. BMC Medical Research Methodology 21, 88. https://doi.org/10.1186/s12874-021-01271-4

Nyvang, J., Hedström, M., Gleissman, S.A., 2016. It's not just a knee, but a whole life: A qualitative descriptive study on patients' experiences of living with knee osteoarthritis and their expectations for knee arthroplasty. International Journal of Qualitative Studies on Health and Well-being 11, 30193. https://doi.org/10.3402/qhw.v11.30193

O'Connell, N.E., Moore, R.A., Stewart, G., Fisher, E., Hearn, L., Eccleston, C., Williams, A.C. de C., 2022. Investigating the veracity of a sample of divergent published trial data in spinal pain. PAIN 10.1097/j.pain.0000000000002659. https://doi.org/10.1097/j.pain.0000000000002659

Ojanen, M., 1994. Can the true effects of exercise on psychological variables be separated from placebo effects? International Journal of Sport Psychology 25, 63–80.

Onifer, S.M., Sozio, R.S., DiCarlo, D.M., Li, Q., Donahue, R.R., Taylor, B.K., Long, C.R., 2018. Spinal Manipulative Therapy Reduces Peripheral Neuropathic Pain in the Rat. Neuroreport 29, 191–196. https://doi.org/10.1097/WNR.0000000000000949

O'Rourke, P.P., Carrithers, J., Patrick-Lake, B., Rice, T.W., Corsmo, J., Hart, R., Drezner, M.K., Lantos, J.D., 2015. Harmonization and streamlining of research oversight for pragmatic clinical trials. Clinical Trials 12, 449–456. https://doi.org/10.1177/1740774515597685

Paci, M., Bianchini, C., Baccini, M., 2022. Reliability of the PEDro scale: comparison between trials published in predatory and non-predatory journals. Archives of Physiotherapy 12, 10. https://doi.org/10.1186/s40945-022-00133-6

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T., Mulrow, C.D., Shamseer, L., Moher, D., 2020. Mapping of reporting guidance for systematic reviews and meta-analyses generated a comprehensive item bank for future reporting guidelines. Journal of Clinical Epidemiology 118, 60–68. https://doi.org/10.1016/j.jclinepi.2019.11.010

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. Journal of Clinical Epidemiology 0. https://doi.org/10.1016/j.jclinepi.2021.03.001

Page, S.J., Persch, A.C., 2013. Recruitment, Retention, and Blinding in Clinical Trials. Am J Occup Ther 67, 154–161. https://doi.org/10.5014/ajot.2013.006197

Pagoto, S., McDermott, M.M., Reed, G., Greenland, P., Mazor, K.M., Ockene, J.K., Whited, M., Schneider, K., Appelhans, B., Leung, K., Merriam, P., Ockene, I., 2013. Can attention control conditions have detrimental effects in behavioral medicine randomized trials? Psychosom Med 75, 137–143. https://doi.org/10.1097/PSY.0b013e3182765dd2

Pallmann, P., Bedding, A.W., Choodari-Oskooei, B., Dimairo, M., Flight, L., Hampson, L.V., Holmes, J., Mander, A.P., Odondi, L., Sydes, M.R., Villar, S.S., Wason, J.M.S., Weir, C.J., Wheeler, G.M., Yap, C., Jaki, T., 2018. Adaptive designs in clinical trials: why use them, and how to run and report them. BMC Medicine 16, 29. https://doi.org/10.1186/s12916-018-1017-7

Panagopoulos, J., Hancock, M.J., Ferreira, P., Hush, J., Petocz, P., 2015. Does the addition of visceral manipulation alter outcomes for patients with low back pain? A randomized placebo controlled trial. European journal of pain (London, England) 19, 899–907. https://doi.org/10.1002/ejp.614

Pandey, A., Tripathi, P., Pandey, R., Srivatava, R., Goswami, S., 2011. Alternative therapies useful in the management of diabetes: A systematic review. J Pharm Bioallied Sci 3, 504–512. https://doi.org/10.4103/0975-7406.90103

Papa, L., Mandara, A., Bottali, M., Gulisano, V., Orfei, S., 2012. A randomized control trial on the effectiveness of osteopathic manipulative treatment in reducing pain and improving the quality of life in elderly patients affected by osteoporosis. Clinical Cases in Mineral and Bone Metabolism 9, 179–183.

Park, R., Mohiuddin, M., Poulin, P.A., Salomons, T., Edwards, R., Nathan, H., Haley, C., Gilron, I., 2020. Systematic scoping review of interactions between analgesic drug therapy and mindfulness-based interventions for chronic pain in adults: current evidence and future directions. Pain Rep 5. https://doi.org/10.1097/PR9.0000000000000868

Pascoe, S.C., Spoonemore, S.L., Young, J.L., Rhon, D.I., 2022. Proposing six criteria to improve reproducibility of “usual care” interventions in back pain trials: a systematic review. Journal of Clinical Epidemiology 149, 227–235. https://doi.org/10.1016/j.jclinepi.2022.05.002

Patrick, D.L., Deyo, R.A., Atlas, S.J., Singer, D.E., Chapin, A., Keller, R.B., 1995. Assessing Health-Related Quality of Life in Patients With Sciatica. Spine 20, 1899.

Pawson, R., 2019. The “pragmatic trial”: An essentially contested concept? Journal of Evaluation in Clinical Practice 25, 943–954. https://doi.org/10.1111/jep.13216

Pêgas de Oliveira, A., 2008. VERIFICATION OF THE ALTERATIONS OF GLUCEMIA AND INSULINEMIA IN PATIENTS WITH DIABETES TYPE II SUBMISSIVE TO A PROTOCOL OF OSTEOPATHIC TREATMENT [WWW Document].

Persaud, N., Mamdani, M.M., 2006. External validity: the neglected dimension in evidence ranking. Journal of Evaluation in Clinical Practice 12, 450–453. https://doi.org/10.1111/j.1365-2753.2006.00730.x

Petit-Zeman, S., Firkins, L., Scadding, J.W., 2010. The James Lind Alliance: tackling research mismatches. The Lancet 376, 667–669. https://doi.org/10.1016/S0140-6736(10)60712-X

Pirosca, S., Shiely, F., Clarke, M., Treweek, S., 2022. Tolerating bad health research: the continuing scandal. Trials 23, 458. https://doi.org/10.1186/s13063-022-06415-5

Pirosca, S., Treweek, S., 2022. WEBINAR - The Good, the Bad and the Ugly A look at 1640 trials included in 96 recent systematic reviews - Ms Stefania Pirosca & Prof. Shaun Treweek. HRB-TMRN. URL https://www.hrb-tmrn.ie/online-material-info/webinar-the-good-the-bad-and-the-ugly-a-look-at-1640-trials-included-in-96-recent-systematic-reviews-ms-stefania-pirosca-prof-shaun-treweek/ (accessed 3.23.23).

Pizzo, P.A., Clark, N.M., Carter-Pokras, O., Christopher, M., Farrar, J.T., Follett, K.A., 2011. Institutes of Medicine: Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research.

Plank, A., Rushton, A., Ping, Y., Mei, R., Falla, D., Heneghan, N.R., 2021. Exploring expectations and perceptions of different manual therapy techniques in chronic low back pain: a qualitative study. BMC Musculoskelet Disord 22, 444. https://doi.org/10.1186/s12891-021-04251-3

Plunkett, A., Fawkes, C., Carnes, D., 2021. UK osteopathic practice in 2019: a retrospective analysis of practice data. https://doi.org/10.1101/2021.01.28.21250601

Pottegård, A., Christensen, R. dePont, Houji, A., Christiansen, C.B., Paulsen, M.S., Thomsen, J.L., Hallas, J., 2014. Primary non-adherence in general practice: a Danish register study. Eur J Clin Pharmacol 70, 757–763. https://doi.org/10.1007/s00228-014-1677-y

Powell, C., 2003. The Delphi technique: myths and realities. Journal of Advanced Nursing 41, 376–382. https://doi.org/10.1046/j.1365-2648.2003.02537.x

Power, M., Hopayian, K., 2011. Exposing the evidence gap for complementary and alternative medicine to be integrated into science-based medicine. J R Soc Med 104, 155–161. https://doi.org/10.1258/jrsm.2011.100271

Public Health Scotland, 2021. Chronic pain waiting times - Quarter ending 30 September 2021 [WWW Document]. URL https://www.publichealthscotland.scot/publications/chronic-pain-waiting-times/chronic-pain-waiting-times-quarter-ending-30-september-2021/ (accessed 2.5.23).

Puhl, A.A., Reinhart, C.J., Doan, J.B., Vernon, H., 2017. The quality of placebos used in randomized, controlled trials of lumbar and pelvic joint thrust manipulation-a systematic review. Spine J 17, 445–456. https://doi.org/10.1016/j.spinee.2016.11.003

Qaseem, A., Wilt, T.J., McLean, R.M., Forciea, M.A., for the Clinical Guidelines Committee of the American College of Physicians, 2017. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. Ann Intern Med 166, 514. https://doi.org/10.7326/M16-2367

Qi, Z., Pang, Y., Lin, L., Zhang, B., Shao, J., Liu, X., Zhang, X., 2018. Acupuncture Combined with Hydrotherapy in Diabetes Patients with Mild Lower-Extremity Arterial Disease: A Prospective, Randomized, Nonblinded Clinical Study. Med Sci Monit 24, 2887–2900. https://doi.org/10.12659/MSM.909733

Queiroz, L.P., 2013. Worldwide epidemiology of fibromyalgia. Curr Pain Headache Rep 17, 356. https://doi.org/10.1007/s11916-013-0356-5

Racaru, S., Sturt, J., Celik, A., 2021. The Effects of Psychological Interventions on Diabetic Peripheral Neuropathy: A Systematic Review and Meta-Analysis. Pain Management Nursing 22, 302–311. https://doi.org/10.1016/j.pmn.2020.11.001

Rashid, A., Thomas, V., Shaw, T., Leng, G., 2017. Patient and Public Involvement in the Development of Healthcare Guidance: An Overview of Current Methods and Future Challenges. Patient 10, 277–282. https://doi.org/10.1007/s40271-016-0206-8

Raveendran, A.V., Chacko, E.C., Pappachan, J.M., 2018. Non-pharmacological Treatment Options in the Management of Diabetes Mellitus. Eur Endocrinol 14, 31–39. https://doi.org/10.17925/EE.2018.14.2.31

Rescher, N., 1998. Predicting the Future: An Introduction to the Theory of Forecasting. SUNY Press.

Reveiz, L., Chapman, E., Asial, S., Munoz, S., Bonfill, X., Alonso-Coello, P., 2015. Risk of bias of randomized trials over time. Journal of Clinical Epidemiology 68, 1036–1045. https://doi.org/10.1016/j.jclinepi.2014.06.001

Reynolds, R.F., Lem, J.A., Gatto, N.M., Eng, S.M., 2011. Is the Large Simple Trial Design Used for Comparative, Post-Approval Safety Research? Drug-Safety 34, 799–820. https://doi.org/10.2165/11593820-000000000-00000

Rice, A.S.C., Smith, B.H., Blyth, F.M., 2016. Pain and the global burden of disease. PAIN 157, 791. https://doi.org/10.1097/j.pain.0000000000000454

Rigoard, P., Basu, S., Desai, M., Taylor, R., Annemans, L., Tan, Y., Johnson, M.J., Van den Abeele, C., North, R., 2019. Multicolumn spinal cord stimulation for predominant back pain in failed back surgery syndrome patients: a multicenter randomized controlled trial. Pain 160, 1410–1420. https://doi.org/10.1097/j.pain.0000000000001510

Ritchie, J., Lewis, J., Lewis, P. of S.P.J., Nicholls, C.M., Ormston, R., 2013. Qualitative Research Practice: A Guide for Social Science Students and Researchers. SAGE.

Rocca, E., Anjum, R.L., 2020. Complexity, Reductionism and the Biomedical Model, in: Anjum, R.L., Copeland, S., Rocca, E. (Eds.), Rethinking Causality, Complexity and Evidence for the Unique Patient: A CauseHealth Resource for Healthcare Professionals and the Clinical Encounter. Springer International Publishing, Cham, pp. 75–94. https://doi.org/10.1007/978-3-030-41239-5_5

Rocha, C.B., Sanchez, T.G., 2012. Efficacy of myofascial trigger point deactivation for tinnitus control. Brazilian journal of otorhinolaryngology 78, 21–6.

Roddy, E.-//-C., 2019. Open-label randomised pragmatic trial (CONTACT) comparing naproxen and low-dose colchicine for the treatment of gout flares in primary care. Annals of the rheumatic diseases. https://doi.org/10.1136/annrheumdis-2019-216154

Roehr, B., 2013. The appeal of large simple trials. BMJ 346, f1317. https://doi.org/10.1136/bmj.f1317

Roland, M., Torgerson, D.J., 1998. Understanding controlled trials: What are pragmatic trials? BMJ 316, 285. https://doi.org/10.1136/bmj.316.7127.285

Rossen, J., Larsson, K., Hagströmer, M., Yngve, A., Brismar, K., Ainsworth, B., Åberg, L., Johansson, U.-B., 2021. Effects of a three-armed randomised controlled trial using self-monitoring of daily steps with and without counselling in prediabetes and type 2 diabetes-the Sophia

Step Study. Int J Behav Nutr Phys Act 18, 121. https://doi.org/10.1186/s12966-021-01193-w

Rossettini, G., Camerone, E.M., Carlino, E., Benedetti, F., Testa, M., 2020. Context matters: the psychoneurobiological determinants of placebo, nocebo and context-related effects in physiotherapy. Archives of Physiotherapy 10, 11. https://doi.org/10.1186/s40945-020-00082-y

Rothrock, J.F., Adams, A.M., Lipton, R.B., Silberstein, S.D., Jo, E., Zhao, X., Blumenfeld, A.M., FORWARD Study investigative group, 2019. FORWARD Study: Evaluating the Comparative Effectiveness of OnabotulinumtoxinA and Topiramate for Headache Prevention in Adults With Chronic Migraine. Headache 59, 1700–1713. https://doi.org/10.1111/head.13653

Rothwell, P.M., 2010. Commentary: External validity of results of randomized trials: disentangling a complex concept. Int J Epidemiol 39, 94–96. https://doi.org/10.1093/ije/dyp305

Rothwell, P.M., 2005a. External validity of randomised controlled trials: "To whom do the results of this trial apply?" The Lancet 365, 82–93. https://doi.org/10.1016/S0140-6736(04)17670-8

Rothwell, P.M., 2005b. External validity of randomised controlled trials: "To whom do the results of this trial apply?" The Lancet 365, 82–93. https://doi.org/10.1016/S0140-6736(04)17670-8

Rothwell, P.M., Watson, C.P.N., Castro-Lopez, J., 2009. External validity of randomized controlled trials: general principles and lessons from trials in neuropathic pain, in: Current Topics in Pain: 12th World Congress on Pain. IASP Press, Seattle, WA, pp. 199–220.

Roura, S., Alvarez, G., Hohenschurz-Schmidt, D., Sola, I., Nunez-Cortes, R., Bracchiglione, J., Fernández-Jané, C., Phalip, J., Gich, I., Sitjà-Rabert, M., Urrutia, G., 2023. A call for improving the Internal validity, generalizability and reporting of manual therapy trials self-labelled as pragmatic: A methodological review. J Clin Epidemiol.

Rowbotham, M.C., Gilron, I., Glazer, C., Rice, A.S.C., Smith, B.H., Stewart, W.F., Wasan, A.D., 2013. Can pragmatic trials help us better understand chronic pain and improve treatment? PAIN 154, 643. https://doi.org/10.1016/j.pain.2013.02.034

Rubinstein, S.M., de Zoete, A., van Middelkoop, M., Assendelft, W.J.J., de Boer, M.R., van Tulder, M.W., 2019. Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. BMJ 364, l689. https://doi.org/10.1136/bmj.l689

Safer, D.L., Hugo, E.M., 2006. Designing a Control for a Behavioral Group Therapy. Behavior Therapy 37, 120–130. https://doi.org/10.1016/j.beth.2005.06.001

Safren, S.A., Gonzalez, J.S., Wexler, D.J., Psaros, C., Delahanty, L.M., Blashill, A.J., Margolina, A.I., Cagliero, E., 2014. A Randomized Controlled Trial of Cognitive Behavioral Therapy for Adherence and Depression (CBT-AD) in Patients With Uncontrolled Type 2 Diabetes. Diabetes Care 37, 625–633. https://doi.org/10.2337/dc13-0816

Sajobi, T.T., Li, G., Awosoga, O., Wang, M., Menon, B.K., Hill, M.D., Thabane, L., 2018. A comparison of meta-analytic methods for synthesizing evidence from explanatory and pragmatic trials. Systematic Reviews 7, 19. https://doi.org/10.1186/s13643-017-0668-3

Sales, C., 2017. Seeing the Person in the Patient: Making the Case for Individualized PROMs in Mental Health Care. Current Psychiatry Reviews 13, 184–187. https://doi.org/10.2174/1573400513666170505111736

Saltaji, H., Armijo-Olivo, S., Cummings, G.G., Amin, M., Flores-Mir, C., 2017. Randomized clinical trials in dentistry: Risks of bias, risks of random errors, reporting quality, and methodologic quality over the years 1955–2013. PLOS ONE 12, e0190089. https://doi.org/10.1371/journal.pone.0190089

Savovic, J., Jones, H.E., Altman, D.G., Harris, R.J., Jüni, P., Pildal, J., Als-Nielsen, B., Balk, E.M., Gluud, C., Gluud, L.L., Ioannidis, J.P.A., Schulz, K.F., Beynon, R., Welton, N.J., Wood, L., Moher, D., Deeks, J.J., Sterne, J.A.C., 2012. Influence of Reported Study Design Characteristics on Intervention Effect Estimates From Randomized, Controlled Trials. Ann Intern Med 157, 429–438. https://doi.org/10.7326/0003-4819-157-6-201209180-00537

Savovic, J., Turner, R.M., Mawdsley, D., Jones, H.E., Beynon, R., Higgins, J.P.T., Sterne, J.A.C., 2018. Association Between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-Epidemiologic Study. Am J Epidemiol 187, 1113–1122. https://doi.org/10.1093/aje/kwx344

Scheliga, K., Friesike, S., Puschmann, C., Fecher, B., 2018. Setting up crowd science projects. Public Underst Sci 27, 515–534. https://doi.org/10.1177/0963662516678514

Schiller, S., 2021. The emergence of physiotherapy in Germany from the mid-19th to the mid-20th centuries: A “female profession” concerned with movement in the health care arena. Physiotherapy Theory and Practice 37, 359–375. https://doi.org/10.1080/09593985.2021.1887061

Schmidt, S., Grossman, P., Schwarzer, B., Jena, S., Naumann, J., Walach, H., 2011. Treating fibromyalgia with mindfulness-based stress reduction: Results from a 3-armed randomized controlled trial. PAIN® 152, 361–369. https://doi.org/10.1016/j.pain.2010.10.043

Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 340, c332. https://doi.org/10.1136/bmj.c332

Schwandt, T.A., 1998. Chapter 7. Constructivist, interpretivist approaches to human inquiry. The Landscape of Qualitative Research: Theories and Issues. Thousand Oaks, SAGE Publications.

Schwartz, D., Lellouch, J., 1967. Explanatory and pragmatic attitudes in therapeutical trials. Journal of Chronic Diseases 20, 637–648. https://doi.org/10.1016/0021-9681(67)90041-0

Schwerla, F., Bischoff, A., Nurnberger, A., Genter, P., Guillaume, J.-P., Resch, K.-L., 2008. Osteopathic treatment of patients with chronic non-specific neck pain: a randomised controlled trial of efficacy. Forschende Komplementarmedizin (2006) 15, 138–45. https://doi.org/10.1159/000132397

Scott, E.A., Black, N., 1991. When does consensus exist in expert panels? Journal of Public Health 13, 35–39. https://doi.org/10.1093/oxfordjournals.pubmed.a042575

Scott, W., Guildford, B.J., Badenoch, J., Driscoll, E., Chilcot, J., Norton, S., Kemp, H.I., Lee, M.J., Lwanga, J., Boffito, M., Moyle, G., Post, F.A., Campbell, L., Josh, J., Clift, P., Williams, A.C. de C., Rice, A., McCracken, L.M., 2021. Feasibility Randomized-Controlled Trial of Online Acceptance and Commitment Therapy for Painful Peripheral Neuropathy in People Living with HIV: The OPEN Study. Eur J Pain 25, 1493–1507. https://doi.org/10.1002/ejp.1762

Sekhon, R., Aminjavahery, N., Davis, C.N., Roswarski, M.J., Robinette, C., 2013. Compliance with opioid treatment guidelines for chronic non-cancer pain (CNCP) in primary care at a Veterans Affairs Medical Center (VAMC). Pain Med 14, 1548–1556. https://doi.org/10.1111/pme.12164

Selby, J.V., Forsythe, L., Sox, H.C., 2015. Stakeholder-Driven Comparative Effectiveness Research: An Update From PCORI. JAMA 314, 2235–2236. https://doi.org/10.1001/jama.2015.15139

Sepehrvand, N., Alemayehu, W., Das, D., Gupta, A.K., Gouda, P., Ghimire, A., Du, A.X., Hatami, S., Babadagli, H.E., Verma, S., Kashour, Z., Ezekowitz, J.A., 2019. Trends in the Explanatory or Pragmatic Nature of Cardiovascular Clinical Trials over 2 Decades. JAMA Cardiology 4, 1122–1128. https://doi.org/10.1001/jamacardio.2019.3604

Shafer, D., Gooing, J., Lee, V., Seffinger, M.A., 2020. Musculoskeletal Conditions in Patients With Diabetes: A Narrative Review. Journal of Osteopathic Medicine 120, 660–664. https://doi.org/10.7556/jaoa.2020.123

Shariff, N., 2015. Utilizing the Delphi survey approach: A review. J Nurs Care 4, 246.

Sheridan, S., Schrandt, S., Forsythe, L., Hilliard, T.S., Paez, K.A., 2017. The PCORI Engagement Rubric: Promising Practices for Partnering in Research. The Annals of Family Medicine 15, 165–170. https://doi.org/10.1370/afm.2042

Sherriff, B., Clark, C., Killingback, C., Newell, D., 2022. Impact of contextual factors on patient outcomes following conservative low back pain treatment: systematic review. Chiropractic & Manual Therapies 30, 20. https://doi.org/10.1186/s12998-022-00430-8

Sherwood, P., Deery, M., Jago, L., 2006. Not Another Survey! Exploring the Issues in Conducting a Webbased Delphi Survey. CAUTHE 2006: To the City and Beyond 1302.

Shrier, I., 2021. Consensus statements that fail to recognise dissent are flawed by design: a narrative review with 10 suggested improvements. Br J Sports Med 55, 545–549. https://doi.org/10.1136/bjsports-2020-102545

Simpson, J.K., Young, K.J., 2020. Vitalism in contemporary chiropractic: a help or a hinderance? Chiropr Man Therap 28, 35. https://doi.org/10.1186/s12998-020-00307-8

Singh, S.B., Singh, P.P., Singh, S., Aggarwal, R., Singh, J., 2012. Effect of Nerve Mobilization on Vibration Perception Threshold in Diabetic Peripheral Neuropathy. Indian Journal of Physiotherapy & Occupational Therapy-An International Journal 6, 189–195.

Sinha, I.P., Smyth, R.L., Williamson, P.R., 2011. Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials: Recommendations for the Future Based on a Systematic Review of Existing Studies. PLoS Med 8. https://doi.org/10.1371/journal.pmed.1000393

Sjøgren, P., Ekholm, O., Peuckmann, V., Grønbæk, M., 2009. Epidemiology of chronic pain in Denmark: An update. European Journal of Pain 13, 287–292. https://doi.org/10.1016/j.ejpain.2008.04.007

Skelly, A.C., Chou, R., Dettori, J.R., Turner, J.A., Friedly, J.L., Rundell, S.D., Fu, R., Brodt, E.D., Wasson, N., Kantner, S., Ferguson, A.J.R., 2020. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review Update. Agency for Healthcare Research and Quality (AHRQ), Rockville, MD. https://doi.org/10.23970/AHRQEPCCER227

Skivington, K., Matthews, L., Simpson, S.A., Craig, P., Baird, J., Blazeby, J.M., Boyd, K.A., Craig, N., French, D.P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., Moore, L., 2021.

A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 374, n2061. https://doi.org/10.1136/bmj.n2061

Smedslund, G., Kjeken, I., Musial, F., Sexton, J., Østerås, N., 2022. Interventions for osteoarthritis pain: A systematic review with network meta-analysis of existing Cochrane reviews. Osteoarthritis and Cartilage Open 4, 100242. https://doi.org/10.1016/j.ocarto.2022.100242

Smelt, A.F., Weele, G.M. van der, Blom, J.W., Gussekloo, J., Assendelft, W.J., 2010. How usual is usual care in pragmatic intervention studies in primary care? An overview of recent trials. Br J Gen Pract 60, e305–e318. https://doi.org/10.3399/bjgp10X514819

Smith, S., Normahani, P., Lane, T., Hohenschurz-Schmidt, D., Oliver, N., Davies, A.H., 2022a. Pathogenesis of Distal Symmetrical Polyneuropathy in Diabetes. Life 12, 1074. https://doi.org/10.3390/life12071074

Smith, S., Normahani, P., Lane, T., Hohenschurz-Schmidt, D., Oliver, N., Davies, A.H., 2022b. Prevention and Management Strategies for Diabetic Neuropathy. Life 12, 1185. https://doi.org/10.3390/life12081185

Smith, S.M., Dworkin, R.H., Turk, D.C., Baron, R., Polydefkis, M., Tracey, I., Borsook, D., Edwards, R.R., Harris, R.E., Wager, T.D., Arendt-Nielsen, L., Burke, L.B., Carr, D.B., Chappell, A., Farrar, J.T., Freeman, R., Gilron, I., Goli, V., Haeussler, J., Jensen, T., Katz, N.P., Kent, J., Kopecky, E.A., Lee, D.A., Maixner, W., Markman, J.D., McArthur, J.C., McDermott, M.P., Parvathenani, L., Raja, S.N., Rappaport, B.A., Rice, A.S.C., Rowbotham, M.C., Tobias, J.K., Wasan, A.D., Witter, J., 2017. The Potential Role of Sensory Testing, Skin Biopsy, and Functional Brain Imaging as Biomarkers in Chronic Pain Clinical Trials: IMMPACT Considerations. The Journal of Pain 18, 757–777. https://doi.org/10.1016/j.jpain.2017.02.429

Sobhani, S., Asayesh, H., Sharifi, F., Djalalinia, S., Baradaran, H.R., Arzaghi, S.M., Mansourian, M., Rezapoor, A., Ansari, H., Masoud, M.P., 2014. Prevalence of diabetic peripheral neuropathy in Iran: a systematic review and meta-analysis. Journal of Diabetes & Metabolic Disorders 13, 1–7.

Soliman, N., Haroutounian, S., Hohmann, A.G., Krane, E., Liao, J., Macleod, M., Segelcke, D., Sena, C., Thomas, J., Vollert, J., Wever, K., Alaverdyan, H., Barakat, A., Barthlow, T., Bozer, A.L.H., Davidson, A., Diaz-delCastillo, M., Dolgorukova, A., Ferdousi, M.I., Healy, C., Hong, S., Hopkins, M., James, A., Leake, H.B., Malewicz, N.M., Mansfield, M., Mardon, A.K., Mattimoe, D., McLoone, D.P., Noes-Holt, G., Pogatzki-Zahn, E.M., Power, E., Pradier, B., Romanos-Sirakis, E., Segelcke, A., Vinagre, R., Yanes, J.A., Zhang, J., Zhang, X.Y., Finn, D.P., Rice, A.S.C., 2021. Systematic review and meta-analysis of cannabinoids, cannabis-based medicines, and endocannabinoid system modulators tested for antinociceptive effects in animal models of injury-related or pathological persistent pain. Pain 162, S26–S44. https://doi.org/10.1097/j.pain.0000000000002269

Soliman, N., Hohmann, A.G., Haroutounian, S., Wever, K., Rice, A.S.C., Finn, D.P., 2019. A protocol for the systematic review and meta-analysis of studies in which cannabinoids were tested for antinociceptive effects in animal models of pathological or injury-related persistent pain: PAIN Reports 4, e766. https://doi.org/10.1097/PR9.0000000000000766

Somerville, S., Hay, E., Lewis, M., Barber, J., Windt, D. van der, Hill, J., Sowden, G., 2008. Content and outcome of usual primary care for back pain: a systematic review. Br J Gen Pract 58, 790–797. https://doi.org/10.3399/bjgp08X319909

Song, X.-J., Huang, Z.-J., Song, W.B., Song, X.-S., Fuhr, A.F., Rosner, A.L., Ndtan, H., Rupert, R.L., 2016. Attenuation Effect of Spinal Manipulation on Neuropathic and Postoperative Pain Through Activating Endogenous Anti-Inflammatory Cytokine Interleukin 10 in Rat Spinal Cord. Journal of Manipulative and Physiological Therapeutics 39, 42–53. https://doi.org/10.1016/j.jmpt.2015.12.004

Sox, H.C., Greenfield, S., 2009. Comparative Effectiveness Research: A Report From the Institute of Medicine. Ann Intern Med 151, 203–205. https://doi.org/10.7326/0003-4819-151-3-200908040-00125

Spall, H.G.C.V., Toren, A., Kiss, A., Fowler, R.A., 2007. Eligibility Criteria of Randomized Controlled Trials Published in High-Impact General Medical Journals: A Systematic Sampling Review. JAMA 297, 1233–1240. https://doi.org/10.1001/jama.297.11.1233

Spallone, V., Morganti, R., D'Amato, C., Greco, C., Cacciotti, L., Marfia, G.A., 2012. Validation of DN4 as a screening tool for neuropathic pain in painful diabetic polyneuropathy. Diabetic Medicine 29, 578–585. https://doi.org/10.1111/j.1464-5491.2011.03500.x

Staniszewska, S., Brett, J., Simera, I., Seers, K., Mockford, C., Goodlad, S., Altman, D.G., Moher, D., Barber, R., Denegri, S., Entwistle, A., Littlejohns, P., Morris, C., Suleman, R., Thomas, V., Tysall, C., 2017. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ 358, j3453. https://doi.org/10.1136/bmj.j3453

Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C.J., Cheng, H.-Y., Corbett, M.S., Eldridge, S.M., Emberson, J.R., Hernán, M.A., Hopewell, S., Hróbjartsson, A., Junqueira, D.R., Jüni, P., Kirkham, J.J., Lasserson, T., Li, T., McAleenan, A., Reeves, B.C., Shepperd, S., Shrier, I., Stewart, L.A., Tilling, K., White, I.R., Whiting, P.F., Higgins, J.P.T., 2019. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 366, l4898. https://doi.org/10.1136/bmj.l4898

Stochkendahl, M.J., Kjaer, P., Hartvigsen, J., Kongsted, A., Aaboe, J., Andersen, M., Andersen, M.Ø., Fournier, G., Højgaard, B., Jensen, M.B., Jensen, L.D., Karbo, T., Kirkeskov, L., Melbye, M., Morsel-Carlsen, L., Nordsteen, J., Palsson, T.S., Rasti, Z., Silbye, P.F., Steiness, M.Z., Tarp, S., Vaagholt, M., 2018. National Clinical Guidelines for non-surgical treatment of patients with recent onset low back pain or lumbar radiculopathy. Eur Spine J 27, 60–75. https://doi.org/10.1007/s00586-017-5099-2

Straus, S.E., Glasziou, P., Richardson, W.S., Haynes, R.B., 2018. Evidence-Based Medicine E-Book: How to Practice and Teach EBM. Elsevier Health Sciences.

Sugarman, J., Califf, R.M., 2014. Ethics and Regulatory Complexities for Pragmatic Clinical Trials. JAMA 311, 2381–2382. https://doi.org/10.1001/jama.2014.4164

Sumsion, T., 1998. The Delphi Technique: An Adaptive Research Tool. British Journal of Occupational Therapy 61, 153–156. https://doi.org/10.1177/030802269806100403

Surenkok, O., Aytar, A., Baltaci, G., 2009. Acute effects of scapular mobilization in shoulder dysfunction: a double-blind randomized placebo-controlled trial. Journal of Sport Rehabilitation 18, 493–501.

Tack, M., 2021. Problems with the MetaBLIND study: An examination of data on blinding patients in trials with patient-reported outcomes. J Health Psychol 13591053211059392. https://doi.org/10.1177/13591053211059391

Taljaard, M., Nicholls, S.G., Howie, A.H., Nix, H.P., Carroll, K., Moon, P.M., Nightingale, N.M., Giraudeau, B., Hey, S.P., Eldridge, S.M., Weijer, C., Zwarenstein, M., 2022. An analysis of published trials found that current use of pragmatic trial labels is uninformative. Journal of Clinical Epidemiology 0. https://doi.org/10.1016/j.jclinepi.2022.08.007

Taljaard, M., Weijer, C., Grimshaw, J.M., Ali, A., Brehaut, J.C., Campbell, M.K., Carroll, K., Edwards, S., Eldridge, S., Forrest, C.B., Giraudeau, B., Goldstein, C.E., Graham, I.D., Hemming, K., Hey, S.P., Horn, A.R., Jairath, V., Klassen, T.P., London, A.J., Marlin, S., Marshall, J.C., McIntyre, L., McKenzie, J.E., Nicholls, S.G., Alison Paprica, P., Zwarenstein, M., Fergusson, D.A., 2018. Developing a framework for the ethical design and conduct of pragmatic trials in healthcare: a mixed methods research protocol. Trials 19, 525. https://doi.org/10.1186/s13063-018-2895-x

Tate, R.L., Perdices, M., 2020. Research Note: Single-case experimental designs. J Physiother 66, 202–206. https://doi.org/10.1016/j.jphys.2020.06.004

Taylor, A.M., Phillips, K., Patel, K.V., Turk, D.C., Dworkin, R.H., Beaton, D., Clauw, D.J., Gignac, M.A.M., Markman, J.D., Williams, D.A., Bujanover, S., Burke, L.B., Carr, D.B., Choy, E.H., Conaghan, P.G., Cowan, P., Farrar, J.T., Freeman, R., Gewandter, J., Gilron, I., Goli, V., Gover, T.D., Haddox, J.D., Kerns, R.D., Kopecky, E.A., Lee, D.A., Malamut, R., Mease, P., Rappaport, B.A., Simon, L.S., Singh, J.A., Smith, S.M., Strand, V., Tugwell, P., Vanhove, G.F., Veasley, C., Walco, G.A., Wasan, A.D., Witter, J., 2016. Assessment of physical function and participation in chronic pain clinical trials: IMMPACT/OMERACT recommendations. PAIN 157, 1836. https://doi.org/10.1097/j.pain.0000000000000577

Teixeira, E., 2010. The Effect of Mindfulness Meditation on Painful Diabetic Peripheral Neuropathy in Adults Older Than 50 Years. Holistic Nursing Practice 24, 277–283. https://doi.org/10.1097/HNP.0b013e3181f1add2

Teys, P., Bisset, L., Vicenzino, B., 2008. The initial effects of a Mulligan's mobilization with movement technique on range of movement and pressure pain threshold in pain-limited shoulders. Manual Therapy 13, 37–42.

The Cochrane Collaboration, 2020. Review Manager 5 (RevMan 5).

The Health Foundation, 2019. NHS performance and waiting times - Priorities for the new government [WWW Document]. URL https://www.health.org.uk/publications/long-reads/nhs-performance-and-waiting-times (accessed 2.5.23).

Themistocleous, A.C., Ramirez, J.D., Shillo, P.R., Lees, J.G., Selvarajah, D., Orengo, C., Tesfaye, S., Rice, A.S.C., Bennett, D.L.H., 2016. The Pain in Neuropathy Study (PiNS): a cross-sectional observational study determining the somatosensory phenotype of painful and painless diabetic neuropathy. Pain 157, 1132–1145. https://doi.org/10.1097/j.pain.0000000000000491

Thomas, J., Noel-Storr, A., Marshall, I., Wallace, B., McDonald, Steven, Mavergames, C., Glasziou, P., Shemilt, I., Synnot, A., Turner, T., Elliott, J., Agoritsas, T., Hilton, J., Perron, C., Akl, E., Hodder, R., Pestridge, C., Albrecht, L., Horsley, T., Platt, J., Armstrong, R., Nguyen, P.H., Plovnick, R., Arno, A., Ivers, N., Quinn, G., Au, A., Johnston, R., Rada, G., Bagg, M., Jones, A., Ravaud, P., Boden, C., Kahale, L., Richter, B., Boisvert, I., Keshavarz, H., Ryan, R., Brandt,

L., Kolakowsky-Hayner, S.A., Salama, D., Brazinova, A., Nagraj, S.K., Salanti, G., Buchbinder, R., Lasserson, T., Santaguida, L., Champion, C., Lawrence, R., Santesso, N., Chandler, J., Les, Z., Schünemann, H.J., Charidimou, A., Leucht, S., Shemilt, I., Chou, R., Low, N., Sherifali, D., Churchill, R., Maas, A., Siemieniuk, R., Cnossen, M.C., MacLehose, H., Simmonds, M., Cossi, M.-J., Macleod, M., Skoetz, N., Counotte, M., Marshall, I., Soares-Weiser, K., Craigie, S., Marshall, R., Srikanth, V., Dahm, P., Martin, N., Sullivan, K., Danilkewich, A., Martínez García, L., Synnot, A., Danko, K., Mavergames, C., Taylor, M., Donoghue, E., Maxwell, L.J., Thayer, K., Dressler, C., McAuley, J., Thomas, J., Egan, C., McDonald, Steve, Tritton, R., Elliott, J., McKenzie, J., Tsafnat, G., Elliott, S.A., Meerpohl, J., Tugwell, P., Etxeandia, I., Merner, B., Turgeon, A., Featherstone, R., Mondello, S., Turner, T., Foxlee, R., Morley, R., van Valkenhoef, G., Garner, P., Munafo, M., Vandvik, P., Gerrity, M., Munn, Z., Wallace, B., Glasziou, P., Murano, M., Wallace, S.A., Green, S., Newman, K., Watts, C., Grimshaw, J., Nieuwlaat, R., Weeks, L., Gurusamy, K., Nikolakopoulou, A., Weigl, A., Haddaway, N., Noel-Storr, A., Wells, G., Hartling, L., O'Connor, A., Wiercioch, W., Hayden, J., Page, M., Wolfenden, L., Helfand, M., Pahwa, M., Yepes Nuñez, J.J., Higgins, J., Pardo, J.P., Yost, J., Hill, S., Pearson, L., 2017. Living systematic reviews: 2. Combining human and machine effort. Journal of Clinical Epidemiology 91, 31–37. https://doi.org/10.1016/j.jclinepi.2017.08.011

Thompson, S.G., Higgins, J.P.T., 2002. How should meta-regression analyses be undertaken and interpreted? Statistics in Medicine 21, 1559–1573. https://doi.org/10.1002/sim.1187

Thomson, O.P., MacMillan, A., 2023. What's wrong with osteopathy? International Journal of Osteopathic Medicine 48, 100659. https://doi.org/10.1016/j.ijosm.2023.100659

Thomson, O.P., Petty, N.J., Moore, A.P., 2014. A qualitative grounded theory study of the conceptions of clinical practice in osteopathy – A continuum from technical rationality to professional artistry. Manual Therapy 19, 37–43. https://doi.org/10.1016/j.math.2013.06.005

Thorlund, J.B., Juhl, C.B., Roos, E.M., Lohmander, L.S., 2015. Arthroscopic surgery for degenerative knee: systematic review and meta-analysis of benefits and harms. BMJ 350, h2747. https://doi.org/10.1136/bmj.h2747

Thorpe, K.E., Zwarenstein, M., Oxman, A.D., Treweek, S., Furberg, C.D., Altman, D.G., Tunis, S., Bergel, E., Harvey, I., Magid, D.J., Chalkidou, K., 2009. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol 62, 464–475. https://doi.org/10.1016/j.jclinepi.2008.12.011

Tian, T.Y., Zlateva, I., Anderson, D.R., 2013. Using electronic health records data to identify patients with chronic pain in a primary care setting. J Am Med Inform Assoc 20, e275–e280. https://doi.org/10.1136/amiajnl-2013-001856

Tölle, T., Xu, X., Sadosky, A.B., 2006. Painful diabetic neuropathy: a cross-sectional survey of health state impairment and treatment patterns. Journal of Diabetes and its Complications 20, 26–33. https://doi.org/10.1016/j.jdiacomp.2005.09.007

Tong, A., Sainsbury, P., Craig, J., 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. International Journal for Quality in Health Care 19, 349–357. https://doi.org/10.1093/intqhc/mzm042

Tough, E.A., White, A.R., Richards, S.H., Lord, B., Campbell, J.L., 2009. Developing and Validating a Sham Acupuncture Needle. Acupunct Med 27, 118–122. https://doi.org/10.1136/aim.2009.000737

Touma, C., Pannain, S., 2011. Does lack of sleep cause diabetes? Cleve Clin J Med 78, 549–558. https://doi.org/10.3949/ccjm.78a.10165

Tournebize, J., Gibaja, V., Muszczak, A., Kahn, J.-P., 2016. Are Physicians Safely Prescribing Opioids for Chronic Noncancer Pain? A Systematic Review of Current Evidence. Pain Pract 16, 370–383. https://doi.org/10.1111/papr.12289

Treweek, S., Bevan, S., Bower, P., Campbell, M., Christie, J., Clarke, M., Collett, C., Cotton, S., Devane, D., El Feky, A., Flemyng, E., Galvin, S., Gardner, H., Gillies, K., Jansen, J., Littleford, R., Parker, A., Ramsay, C., Restrup, L., Sullivan, F., Torgerson, D., Tremain, L., Westmore, M., Williamson, P.R., 2018. Trial Forge Guidance 1: what is a Study Within A Trial (SWAT)? Trials 19, 139. https://doi.org/10.1186/s13063-018-2535-5

Treweek, S., Zwarenstein, M., 2009. Making trials matter: pragmatic and explanatory trials and the problem of applicability. Trials 10, 37. https://doi.org/10.1186/1745-6215-10-37

Tucker, H.-R., Scaff, K., McCloud, T., Carlomagno, K., Daly, K., Garcia, A., Cook, C.E., 2020. Harms and benefits of opioids for management of non-surgical acute and chronic low back pain: a systematic review. Br J Sports Med 54, 664–664. https://doi.org/10.1136/bjsports-2018-099805

Turk, D.C., Dworkin, R.H., Allen, R.R., Bellamy, N., Brandenburg, N., Carr, D.B., Cleeland, C., Dionne, R., Farrar, J.T., Galer, B.S., Hewitt, D.J., Jadad, A.R., Katz, N.P., Kramer, L.D., Manning, D.C., McCormick, C.G., McDermott, M.P., McGrath, P., Quessy, S., Rappaport, B.A., Robinson, J.P., Royal, M.A., Simon, L., Stauffer, J.W., Stein, W., Tollett, J., Witter, J., 2003. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 106, 337–345. https://doi.org/10.1016/j.pain.2003.08.001

Turk, D.C., Dworkin, R.H., Revicki, D., Harding, G., Burke, L.B., Cella, D., Cleeland, C.S., Cowan, P., Farrar, J.T., Hertz, S., Max, M.B., Rappaport, B.A., 2008. Identifying important outcome domains for chronic pain clinical trials: An IMMPACT survey of people with pain. PAIN® 137, 276–285. https://doi.org/10.1016/j.pain.2007.09.002

Turner, J.A., Deyo, R.A., Loeser, J.D., Von Korff, M., Fordyce, W.E., 1994. The importance of placebo effects in pain treatment and research. JAMA 271, 1609–1614.

Turyahikayo, E., 2014. Resolving the qualitative-quantitative debate in healthcare research. Medical Practice and Reviews 5, 6–15.

UK Office for Health Improvement and Disparities, n.d. Creating a logic model for an intervention: evaluation in health and wellbeing [WWW Document]. GOV.UK. URL https://www.gov.uk/guidance/evaluation-in-health-and-wellbeing-creating-a-logic-model (accessed 2.21.23).

Underwood, M.R., Morton, V., Farrin, A., on behalf of the UK BEAM trial team, 2007. Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN32683578]. Rheumatology 46, 1297–1302. https://doi.org/10.1093/rheumatology/kem113

Unsgaard-Tøndel, M., Søderstrøm, S., 2021. Therapeutic Alliance: Patients' Expectations Before and Experiences After Physical Therapy for Low Back Pain—A Qualitative Study With 6-Month Follow-Up. Physical Therapy 101, pzab187. https://doi.org/10.1093/ptj/pzab187

Van De Ven, A.H., Johnson, P.E., 2006. Knowledge for Theory and Practice. AMR 31, 802–821. https://doi.org/10.5465/amr.2006.22527385

van Hecke, O., Hocking, L.J., Torrance, N., Campbell, A., Padmanabhan, S., Porteous, D.J., McIntosh, A.M., Burri, A.V., Tanaka, H., Williams, F.M.K., Smith, B.H., 2017. Chronic pain, depression and cardiovascular disease linked through a shared genetic predisposition: Analysis of a family-based cohort and twin study. PLoS ONE 12, e0170653. https://doi.org/10.1371/journal.pone.0170653

Vandvik, P.O., Lähdeoja, T., Ardern, C., Buchbinder, R., Moro, J., Brox, J.I., Burgers, J., Hao, Q., Karjalainen, T., Bekerom, M. van den, Noorduyn, J., Lytvyn, L., Siemieniuk, R.A.C., Albin, A., Shunjie, S.C., Fisch, F., Proulx, L., Guyatt, G., Agoritsas, T., Poolman, R.W., 2019. Subacromial decompression surgery for adults with shoulder pain: a clinical practice guideline. BMJ 364, l294. https://doi.org/10.1136/bmj.l294

Vase, L., Amanzio, M., Price, D.D., 2015. Nocebo vs. Placebo: The Challenges of Trial Design in Analgesia Research. Clinical Pharmacology & Therapeutics 97, 143–150. https://doi.org/10.1002/cpt.31

Vase, Lene, Baram, S., Takakura, N., Takayama, M., Yajima, H., Kawase, A., Schuster, L., Kaptchuk, T.J., Schou, S., Jensen, T.S., Zachariae, R., Svensson, P., 2015a. Can Acupuncture Treatment Be Double-Blinded? An Evaluation of Double-Blind Acupuncture Treatment of Postoperative Pain. PLOS ONE 10, e0119612. https://doi.org/10.1371/journal.pone.0119612

Vase, Lene, Vollert, J., Finnerup, N.B., Miao, X., Atkinson, G., Marshall, S., Nemeth, R., Lange, B., Liss, C., Price, D.D., Maier, C., Jensen, T.S., Segerdahl, M., 2015b. Predictors of the placebo analgesia response in randomized controlled trials of chronic pain: a meta-analysis of the individual data from nine industrially sponsored trials. PAIN 156, 1795–1802. https://doi.org/10.1097/j.pain.0000000000000217

Vermeersch, A., Quaghebeur, J., 2018. The effect of an osteopathic treatment on the blood glucose level in diabetic patients. [WWW Document].

Vernon, H.T., Triano, J.J., Ross, J.K., Tran, S.K., Soave, D.M., Dinulos, M.D., 2012. Validation of a novel sham cervical manipulation procedure. The spine journal : official journal of the North American Spine Society 12, 1021–8. https://doi.org/10.1016/j.spinee.2012.10.009

Vileikyte, L., Peyrot, M., Bundy, C., Rubin, R.R., Leventhal, H., Mora, P., Shaw, J.E., Baker, P., Boulton, A.J.M., 2003. The Development and Validation of a Neuropathy- and Foot Ulcer-Specific Quality of Life Instrument. Diabetes Care 26, 2549–2555. https://doi.org/10.2337/diacare.26.9.2549

Villamar, M.F., Contreras, V.S., Kuntz, R.E., Fregni, F., 2013. The reporting of blinding in physical medicine and rehabilitation randomized controlled trials: a systematic review. J Rehabil Med 45, 6–13. https://doi.org/10.2340/16501977-1071

Vitiello, M.V., McCurry, S.M., Shortreed, S.M., Balderson, B.H., Baker, L.D., Keefe, F.J., Rybarczyk, B.D., Von Korff, M., 2013. Cognitive-Behavioral Treatment for Comorbid Insomnia and Osteoarthritis Pain in Primary Care: The Lifestyles Randomized Controlled Trial. J Am Geriatr Soc 61, 947–956. https://doi.org/10.1111/jgs.12275

Vitiello, M.V., Rybarczyk, B., Von Korff, M., Stepanski, E.J., 2009. Cognitive behavioral therapy for insomnia improves sleep and decreases pain in older adults with co-morbid insomnia and

osteoarthritis. Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine 5, 355–62.

Vogel, S., 2021. W(h)ither osteopathy: A call for reflection; a call for submissions for a special issue. International Journal of Osteopathic Medicine 41, 1–3. https://doi.org/10.1016/j.ijosm.2021.09.001

von Wernsdorff, M., Loef, M., Tuschen-Caffier, B., Schmidt, S., 2021. Effects of open-label placebos in clinical trials: a systematic review and meta-analysis. Scientific Reports 11, 3855. https://doi.org/10.1038/s41598-021-83148-6

Wagner, J.A., Bermudez-Millan, A., Damio, G., Segura-Perez, S., Chhabra, J., Vergara, C., Feinn, R., Perez-Escamilla, R., 2016. A randomized, controlled trial of a stress management intervention for Latinos with type 2 diabetes delivered by community health workers: Outcomes for psychological wellbeing, glycemic control, and cortisol. Diabetes Research and Clinical Practice 120, 162–170. https://doi.org/10.1016/j.diabres.2016.07.022

Walach, H., 2001. The Efficacy Paradox in Randomized Controlled Trials of CAM and Elsewhere: Beware of the Placebo Trap. The Journal of Alternative and Complementary Medicine 7, 213–218. https://doi.org/10.1089/107555301300328070

Walach, H., Falkenberg, T., Fønnebø, V., Lewith, G., Jonas, W.B., 2006. Circular instead of hierarchical: methodological principles for the evaluation of complex interventions. BMC Med Res Methodol 6, 29. https://doi.org/10.1186/1471-2288-6-29

Walach, H., Loef, M., 2015. Using a matrix-analytical approach to synthesizing evidence solved incompatibility problem in the hierarchy of evidence. Journal of Clinical Epidemiology 68, 1251–1260. https://doi.org/10.1016/j.jclinepi.2015.03.027

Walker, B.F., Hebert, J.J., Stomski, N.J., Losco, B., French, S.D., 2013. Short-term usual chiropractic care for spinal pain: a randomized controlled trial. Spine 38, 2071–8. https://doi.org/10.1097/01.brs.0000435032.73187.c7

Walsh, E., Sheridan, A., 2016. Factors affecting patient participation in clinical trials in Ireland: A narrative review. Contemporary Clinical Trials Communications 3, 23–31. https://doi.org/10.1016/j.conctc.2016.01.002

Wang, C., Schmid, C.H., Hibberd, P.L., Kalish, R., Roubenoff, R., Rones, R., McAlindon, T., 2009. Tai Chi is effective in treating knee osteoarthritis: a randomized controlled trial. Arthritis and Rheumatism 61, 1545–53. https://doi.org/10.1002/art.24832

Wang, C., Schmid, C.H., Rones, R., Kalish, R., Yinh, J., Goldenberg, D.L., Lee, Y., McAlindon, T., 2010. A Randomized Trial of Tai Chi for Fibromyalgia. New England Journal of Medicine 363, 743–754. https://doi.org/10.1056/NEJMoa0912611

Ware, J.E., Gandek, B., 1998. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. Journal of Clinical Epidemiology 51, 903–912. https://doi.org/10.1016/S0895-4356(98)00081-X

Ware Jr., J.E., 1999. SF-36 Health Survey, in: The Use of Psychological Testing for Treatment Planning and Outcomes Assessment, 2nd Ed. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 1227–1246.

Wartolowska, K.A., Hohenschurz-Schmidt, D., Vase, L., Aronson, J.K., 2022. The importance of using placebo controls in non-pharmacological randomised trials. PAIN 10.1097/j.pain.0000000000002839. https://doi.org/10.1097/j.pain.0000000000002839

Wasan, A.D., 2014. Efficacy vs Effectiveness and Explanatory vs Pragmatic: Where Is the Balance Point in Pain Medicine Research? Pain Med 15, 539–540. https://doi.org/10.1111/pme.12420

Watson, C.P.N., 2007. External validity of pharmaceutical trials in neuropathic pain, in: Rothwell, P.M. (Ed.), Treating Individuals: From Randomised Trials to Personalised Medicine. Elsevier, London, pp. 121–130.

Watson, J.A., Ryan, C.G., Cooper, L., Ellington, D., Whittle, R., Lavender, M., Dixon, J., Atkinson, G., Cooper, K., Martin, D.J., 2019. Pain Neuroscience Education for Adults With Chronic Musculoskeletal Pain: A Mixed-Methods Systematic Review and Meta-Analysis. J Pain 20, P1140.e1-1140.e2. https://doi.org/10.1016/j.jpain.2019.02.011

Webster, R.K., Howick, J., Hoffmann, T., Macdonald, H., Collins, G.S., Rees, J.L., Napadow, V., Madigan, C., Price, A., Lamb, S.E., Bishop, F.L., Bokelmann, K., Papanikitas, A., Roberts, N., Evers, A.W.M., 2019. Inadequate description of placebo and sham controls in a systematic review of recent trials. European Journal of Clinical Investigation 49, e13169. https://doi.org/10.1111/eci.13169

Weinfurt, K.P., 2019. Clarifying the Meaning of Clinically Meaningful Benefit in Clinical Research: Noticeable Change vs Valuable Change. JAMA 322, 2381–2382. https://doi.org/10.1001/jama.2019.18496

Welch, M.J., Lally, R., Miller, J.E., Pittman, S., Brodsky, L., Caplan, A.L., Uhlenbrauck, G., Louzao, D.M., Fischer, J.H., Wilfond, B., 2015. The ethics and regulatory landscape of including vulnerable populations in pragmatic clinical trials. Clinical Trials 12, 503–510. https://doi.org/10.1177/1740774515597701

Welsing, P.M., Oude Rengerink, K., Collier, S., Eckert, L., van Smeden, M., Ciaglia, A., Nachbaur, G., Trelle, S., Taylor, A.J., Egger, M., Goetz, I., 2017. Series: Pragmatic trials and real world evidence: Paper 6. Outcome measures in the real world. Journal of Clinical Epidemiology 90, 99–107. https://doi.org/10.1016/j.jclinepi.2016.12.022

Wewege, M.A., Bagg, M.K., Jones, M.D., Ferraro, M.C., Cashin, A.G., Rizzo, R.R., Leake, H.B., Hagstrom, A.D., Sharma, S., McLachlan, A.J., Maher, C.G., Day, R., Wand, B.M., O'Connell, N.E., Nikolakopolou, A., Schabrun, S., Gustin, S.M., McAuley, J.H., 2023. Comparative effectiveness and safety of analgesic medicines for adults with acute non-specific low back pain: systematic review and network meta-analysis. BMJ 380, e072962. https://doi.org/10.1136/bmj-2022-072962

Whitty, C.J.M., MacEwen, C., Goddard, A., Alderson, D., Marshall, M., Calderwood, C., Atherton, F., McBride, M., Atherton, J., Stokes-Lampard, H., Reid, W., Powis, S., Marx, C., 2020. Rising to the challenge of multimorbidity. BMJ 368, l6964. https://doi.org/10.1136/bmj.l6964

Wibault, J., Öberg, B., Dedering, Å., Löfgren, H., Zsigmond, P., Peolsson, A., 2018. Structured postoperative physiotherapy in patients with cervical radiculopathy: 6-month outcomes of a randomized clinical trial. J Neurosurg Spine 28, 1–9. https://doi.org/10.3171/2017.5.SPINE16736

Wieland, L.S., Berman, B.M., Altman, D.G., Barth, J., Bouter, L.M., D'Adamo, C.R., Linde, K., Moher, D., Mullins, C.D., Treweek, S., Tunis, S., van der Windt, D.A., Zwarenstein, M., Witt, C., 2017. Rating of Included Trials on the Efficacy–Effectiveness Spectrum: development

of a new tool for systematic reviews. Journal of Clinical Epidemiology 84, 95–104. https://doi.org/10.1016/j.jclinepi.2017.01.010

Williams, A.C. de C., Fisher, E., Hearn, L., Eccleston, C., 2021. Evidence-based psychological interventions for adults with chronic pain: precision, control, quality, and equipoise. PAIN 160, 2149–2153. https://doi.org/10.1097/j.pain.0000000000002273

Williams, A.C. de C., Fisher, E., Hearn, L., Eccleston, C., 2020. Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database of Systematic Reviews. https://doi.org/10.1002/14651858.CD007407.pub4

Williams, C.M., Maher, C.G., Hancock, M.J., McAuley, J.H., McLachlan, A.J., Britt, H., Fahridin, S., Harrison, C., Latimer, J., 2010. Low Back Pain and Best Practice Care: A Survey of General Practice Physicians. Archives of Internal Medicine 170, 271–277. https://doi.org/10.1001/archinternmed.2009.507

Williams, H.C., Burden-Teh, E., Nunn, A.J., 2015. What Is a Pragmatic Clinical Trial? Journal of Investigative Dermatology 135, 1–3. https://doi.org/10.1038/jid.2015.134

Wilson, D.B., 2005. SPSS for Windows Meta-Analysis Macros.

Witt, C.M., Manheimer, E., Hammerschlag, R., Lüdtke, R., Lao, L., Tunis, S.R., Berman, B.M., 2012. How Well Do Randomized Trials Inform Decision Making: Systematic Review Using Comparative Effectiveness Research Measures on Acupuncture for Back Pain. PLOS ONE 7, e32399. https://doi.org/10.1371/journal.pone.0032399

Wright, P.J., Pinto, B.M., Corbett, C.F., 2021. Balancing Internal and External Validity Using Precis-2 and Re-Aim: Case Exemplars. West J Nurs Res 43, 163–171. https://doi.org/10.1177/0193945920940308

Yamamoto, J.M., Kellett, J.E., Balsells, M., García-Patterson, A., Hadar, E., Solà, I., Gich, I., Beek, E.M. van der, Castañeda-Gutiérrez, E., Heinonen, S., Hod, M., Laitinen, K., Olsen, S.F., Poston, L., Rueda, R., Rust, P., Lieshout, L. van, Schelkle, B., Murphy, H.R., Corcoy, R., 2018. Gestational Diabetes Mellitus and Diet: A Systematic Review and Meta-analysis of Randomized Controlled Trials Examining the Impact of Modified Dietary Interventions on Maternal Glucose Control and Neonatal Birth Weight. Diabetes Care 41, 1346–1361. https://doi.org/10.2337/dc18-0102

Yamato, T.P., Maher, C., Koes, B., Moseley, A., 2017. The PEDro scale had acceptably high convergent validity, construct validity, and interrater reliability in evaluating methodological quality of pharmaceutical trials. J Clin Epidemiol 86, 176–181. https://doi.org/10.1016/j.jclinepi.2017.03.002

Yonis, A., Taylor, R., Warren, F., Hay-Smith, J., Levack, W., Dean, S., 2020. Re: Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. BMJ 368. https://doi.org/10.1136/bmj.l6802

Yordanov, Y., Dechartres, A., Porcher, R., Boutron, I., Altman, D.G., Ravaud, P., 2015. Avoidable waste of research related to inadequate methods in clinical trials. BMJ 350, h809. https://doi.org/10.1136/bmj.h809

Younger, J., Gandhi, V., Hubbard, E., Mackey, S., 2012. Development of the Stanford Expectations of Treatment Scale (SETS): A tool for measuring patient outcome expectancy in clinical trials. Clinical Trials 9, 767–776. https://doi.org/10.1177/1740774512465064

Yu, S.P., Ferreira, M.L., Duong, V., Caroupapoullé, J., Arden, N.K., Bennell, K.L., Hunter, D.J., 2021. Responsiveness of an activity tracker as a measurement tool in a knee osteoarthritis

clinical trial (ACTIVe-OA study). Ann Phys Rehabil Med 101619. https://doi.org/10.1016/j.rehab.2021.101619

Zarin, D.A., Goodman, S.N., Kimmelman, J., 2019. Harms From Uninformative Clinical Trials. JAMA 322, 813–814. https://doi.org/10.1001/jama.2019.9892

Zhang, W., Robertson, J., Jones, A.C., Dieppe, P.A., Doherty, M., 2008. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. Annals of the Rheumatic Diseases 67, 1716–1723. https://doi.org/10.1136/ard.2008.092015

Zhang, Y.-Q., Jiao, R.-M., Witt, C.M., Lao, L., Liu, J.-P., Thabane, L., Sherman, K.J., Cummings, M., Richards, D.P., Kim, E.-K.A., Kim, T.-H., Lee, M.S., Wechsler, M.E., Brinkhaus, B., Mao, J.J., Smith, C.A., Gang, W.-J., Liu, B.-Y., Liu, Z.-S., Liu, Y., Zheng, H., Wu, J.-N., Carrasco-Labra, A., Bhandari, M., Devereaux, P.J., Jing, X.-H., Guyatt, G., 2022. How to design high quality acupuncture trials—a consensus informed by evidence. BMJ 376, e067476. https://doi.org/10.1136/bmj-2021-067476

Zhu, G.-C., Chen, Y.-W., Tsai, K.-L., Wang, J.-J., Hung, C.-H., Schmid, A.B., 2022. Effects of Neural Mobilization on Sensory Dysfunction and Peripheral Nerve Degeneration in Rats With Painful Diabetic Neuropathy. Physical Therapy 102, pzac104. https://doi.org/10.1093/ptj/pzac104

Ziegler, D., Tesfaye, S., Spallone, V., Gurieva, I., Al Kaabi, J., Mankovsky, B., Martinka, E., Radulian, G., Thy Nguyen, K., Stirban, A.O., Tankova, T., Varkonyi, T., Freeman, R., Kempler, P., Jm Boulton, A., 2021. Screening, diagnosis and management of diabetic sensorimotor polyneuropathy in clinical practice: International expert consensus recommendations. Diabetes Res Clin Pract 109063. https://doi.org/10.1016/j.diabres.2021.109063

Zuccaro, S.M., Vellucci, R., Sarzi-Puttini, P., Cherubino, P., Labianca, R., Fornasari, D., 2012. Barriers to Pain Management. Clin Drug Investig 32, 11–19. https://doi.org/10.2165/11630040-000000000-00000

Zuidgeest, M.G., Goetz, I., Grobbee, D.E., 2017. PRECIS-2 in perspective: what is next for pragmatic trials? Journal of Clinical Epidemiology 84, 22–24. https://doi.org/10.1016/j.jclinepi.2016.02.027

Zuidgeest, Mira G.P., Goetz, I., Groenwold, R.H.H., Irving, E., van Thiel, G.J.M.W., Grobbee, D.E., 2017. Series: Pragmatic trials and real world evidence: Paper 1. Introduction. Journal of Clinical Epidemiology 88, 7–13. https://doi.org/10.1016/j.jclinepi.2016.12.023

Zuidgeest, Mira G. P., Welsing, P.M.J., van Thiel, G.J.M.W., Ciaglia, A., Alfonso-Cristancho, R., Eckert, L., Eijkemans, M.J.C., Egger, M., 2017. Series: Pragmatic trials and real world evidence: Paper 5. Usual care and real life comparators. Journal of Clinical Epidemiology 90, 92–98. https://doi.org/10.1016/j.jclinepi.2017.07.001

Zwarenstein, M., Thorpe, K., Treweek, S., Loudon, K., 2020. PRECIS-2 for retrospective assessment of RCTs in systematic reviews. J Clin Epidemiol 126, 202–206. https://doi.org/10.1016/j.jclinepi.2020.06.023

Zwarenstein, M., Treweek, S., 2009. What kind of randomized trials do we need? CMAJ 180, 998–1000. https://doi.org/10.1503/cmaj.082007

Zwarenstein, M., Treweek, S., Gagnier, J.J., Altman, D.G., Tunis, S., Haynes, B., Oxman, A.D., Moher, D., CONSORT group, Pragmatic Trials in Healthcare (Practihc) group, 2008. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 337, a2390. https://doi.org/10.1136/bmj.a2390

Zwarenstein, M., Treweek, S., Loudon, K., 2017. PRECIS-2 helps researchers design more applicable RCTs while CONSORT Extension for Pragmatic Trials helps knowledge users decide whether to apply them. Journal of Clinical Epidemiology 84, 27–29. https://doi.org/10.1016/j.jclinepi.2016.10.010