Development of a core set of outcomes for foot and ankle disorders in rheumatic and musculoskeletal diseases (COMFORT)

Lara Suzanne Chapman

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Medicine

December 2025

Intellectual property and publication statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the researcher and other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapters 1 and 6 include work from the following published paper:

Chapman LS, Jones J, Redmond AC, Flurey CA, Richards P, Hofstetter C, Smith TO, Arnold JB, Hannan MT, Maxwell LJ, Menz HB, Shea B, Golightly YM, Tugwell P, Beaton D, Conaghan PG, Helliwell PS, Siddle HJ. Developing a core outcome set for foot and ankle disorders in rheumatic and musculoskeletal diseases: A scoping review and report from the OMERACT 2022 foot and ankle special interest group session. Semin Arthritis Rheum. 2023 Aug;61:152210. https://doi.org/10.1016/j.semarthrit.2023.152210

Chapters 2-6 include work from the following published protocol:

Chapman LS, Redmond AC, Flurey CA, Richards P, Smith TO, Arnold JB, Beaton D, Conaghan PG, Golightly YM, Hannan MT, Hofstetter C, Maxwell LJ, Menz HB, Shea B, Tugwell P, Helliwell PS, Siddle HJ. Developing an Outcome Measures in Rheumatology (OMERACT) Core set of Outcome Measures for FOot and ankle disorders in RheumaTic and musculoskeletal diseases (COMFORT): core domain set study protocol. Trials. 2023 Jan 28;24(1):65. https://doi.org/10.1186/s13063-023-07104-7

Chapter 3 includes work from the following published paper:

Chapman LS, Flurey CA, Redmond AC, Richards P, Hofstetter C, Tapster B, Emmel J, Helliwell PS, Menz HB, Hannan MT, Shea B, Siddle HJ. Living with foot and ankle disorders in rheumatic and musculoskeletal diseases: A systematic review of qualitative studies to inform the work of the OMERACT Foot and Ankle Working Group. Semin Arthritis Rheum. 2023 Aug;61:152212. https://doi.org/10.1016/j.semarthrit.2023.152212

Chapter 4 includes work from the following published paper:

Chapman LS, Alcacer-Pitarch B, Pauling JD, Flurey CA, Redmond AC, Richards P, Herrick AL, Merkel PA, Proudman S, Menz HB, Helliwell PS, Hannan MT, Domsic RT, Saketkoo LA, Shea B, Siddle HJ. Patients' perspectives on systemic sclerosis-related Raynaud's phenomenon in the feet: A qualitative study from the OMERACT Foot and Ankle Working Group. Semin Arthritis Rheum. 2024 Apr;65:152372. https://10.0.3.248/j.semarthrit.2024.152372

Chapter 5 includes work from the following published paper:

Chapman LS, Flurey CA, Richards P, Redmond AC, Soliman E, Moshrif A, Malone L, Joyce C, Arnold JB, Golightly YM, Hofstetter C, Helliwell PS, Menz HB, Hannan MT, Rahman MN, Shea BJ, Smith TO, Siddle HJ. What outcomes are important to people with foot and ankle disorders in rheumatic and musculoskeletal diseases? An OMERACT qualitative interview study across four continents. Semin Arthritis Rheum. 2025 Jun;72:152671.

https://10.1016/j.semarthrit.2025.152671

The candidate is the first and corresponding author on all of the above papers. The researcher led all aspects of the work reported including study conception, study design, data collection, data analysis, data interpretation, drafting the manuscripts and responding to peer reviewers' comments. The supervisory team and project advisory group contributed to the revision of the papers, and approved final versions. Their contributions to specific aspects of the papers are described below:

Scoping review paper

JJ: Data interpretation

CAF: Study conception, study design and data interpretation

PR: Study conception, study design and data interpretation

CH: Study conception, study design and data interpretation

TOS: Study conception, study design and data interpretation

JBA: Study design and data interpretation

MTH: Study conception, study design and data interpretation

LHM: Study design

HBM: Study conception, study design and data interpretation

BS: Study design

YMG: Study design

PT: Study design

DB: Study design

PGC: Study design

PSH: Study conception, study design and data interpretation

HJS: Study conception, study design and data interpretation

Protocol paper

CAF: Study design

PR: Study design

TOS: Study design

JBA: Study design

DB: Study design

PGC: Study design

YMG: Study design

MTH: Study design

CH: Study design

LJM: Study design

HBM: Study design

BS: Study design

PT: Study design

PSH: Study conception and study design

HJS: Study conception and study design

Qualitative synthesis paper

CAF: Study conception, study design and data interpretation

ACR: Study conception and study design

PR: Study design and data interpretation

CH: Study design and data interpretation

BT: Study design

JE: Study design

PSH: Study conception and study design

HBM: Study design

MTH: Study design

BS: Study design

HJS: Study conception, study design and data interpretation

Secondary analysis paper

BAP: Data interpretation

JDP: Original study conception

CAF: Study design and data interpretation

PR: Data interpretation

ALH: Original study conception

PAM: Study conception

SP: Original study conception

HBM: Study design and data interpretation

PSH: Study design and data interpretation

MTH: Study design and data interpretation

RTD: Original study conception

BS: Study design

HJS: Study conception, study design and data interpretation

Qualitative interview paper

CAF: Study conception, study design and data interpretation

PR: Study conception, study design and data interpretation

ACR: Study design and data interpretation

ES: Study design and data interpretation

AM: Study design and data interpretation

LM: Data interpretation

CJ: Data interpretation

JBA: Study design and data interpretation

CH: Study design and data interpretation

PSH: Study design and data interpretation

HBM: Study design and data interpretation

MTH: Study design and data interpretation

MNR: Data interpretation

TOS: Study design and data interpretation

HJS: Study conception, study design and data interpretation

Acknowledgements

Firstly, I would like to thank all of the participants who gave up their time to take part in the studies within my PhD, and HEE/NIHR for funding my fellowship.

Next, much, much thanks to:

H. From all of the coffees and PARs and chats, to hysterically knocking on my door at 4am, getting stranded on an island and running all over the globe, it really has been the most extraordinary adventure. Thank you for all of the opportunities you've given me and for making this, and all of the work we do together, feel like a hobby. You are and will always be the most brilliant person I've ever worked with.

T. You have taught me so much. I would never dare to stupidly split an infinitive now. You must know that almost every day of my PhD genuinely has been the best day ever, because of you. Thanks for believing I could do this and for always having my spine. I think it is fair to say that when you emerge from the merch shop holding the goods, it will all be worth it. However, that chicken *was* overcooked.

My external supervisors, Caroline and Pam. You have made every supervision meeting feel like a catch up with old friends. Thank you both for championing me since day one. Caroline, thank you for sharing your extensive knowledge with me, especially all things qualitative. Pam, thank you for sharing your experiences. You have both made this work an absolute joy to complete.

Everyone involved with the OMERACT work, particularly Professor Philip Helliwell, Professor Marian Hannan, Professor Hylton Menz, Professor Toby Smith, Cathie Hofstetter, Professor Philip Conaghan, Dr John Arnold, Shawna Grosskleg, Professor Peter Tugwell, and Dr Beverley Shea. You have all been supportive beyond measure. Thank you.

My mentor, Professor Cathy Bowen. Your positivity is contagious. Thank you for all of the wise words and encouragement over the last decade.

Dr Mike Backhouse. Thank you for giving me an early opportunity to do research at Leeds, and for all of your support and kindness ever since.

All of my FASTER buddies. Thank you for your friendship. Particularly you, Chris Drake. The dream we had in 2018 to be office mates came true, and no other office mate will ever compare. Thank you for maintaining your cool, calm exterior as our office got taken over by Lego, action figures, stickers, and panini crumbs. You have helped me so much, especially with every moment of self-doubt. Your enhanced understanding of how people think and behave, and how you make them feel about themselves, is your gift. Thank you for sharing it with me.

My dear Mum. I could tell your patience was starting to waver somewhat when I embarked upon the tenth new life plan, but here we are. Thank you for all of your love and support.

Nic. You have certainly put up with a lot since I signed up for a podiatry degree on a whim in 2013. Thank you for looking after me. And thank you for doing all of the chores as I continually reiterated what a crucial point it was in my PhD whilst re-reading all of the Harry Potter books. There is no way I would have been able to finish this (or start it) without you.

Finally, I would like to dedicate this thesis to two people. To my FIL, Les, who was sincerely relieved when I put down the scalpel. And to my Dad, who died just before I started my fellowship, but was chuffed to bits that I finally managed to wing one.

Abstract

Background: Rheumatic and musculoskeletal diseases (RMDs) frequently affect the foot and ankle, leading to pain, disability, and a reduction in quality of life, but there is a lack of high-quality evidence to determine the effectiveness of treatments for foot and ankle disorders. A contributing factor is heterogeneity in the domains (outcomes) measured across research studies in this area. This hinders the ability to compare study findings. Additionally, the domains measured may not be meaningful or relevant to patients.

Aim: To develop a core set of outcomes for foot and ankle disorders in rheumatic and musculoskeletal diseases (COMFORT).

Methods: COMFORT was developed with multidisciplinary, multi-contributor groups using a rigorous evidence-based approach, which employed a mix of methods. A systematic review and thematic synthesis of existing qualitative studies (n=34 studies) explored patients' symptoms and their impact, and identified gaps in the literature. Patients' experiences of foot and ankle problems in under-researched areas were further explored in a secondary analysis of focus group data (n=40 patients), and a primary qualitative interview study in eight countries (n=56 patients). Domains of potential importance identified in the qualitative work were presented to patients, healthcare professionals, and researchers in a modified Delphi consensus study, consisting of four rounds of online surveys. Participants rated the importance of each domain on a 9-point scale.

Findings: A total of 206 participants representing 22 countries completed at least one round of the Delphi study. Consensus was reached on the inclusion of five core domains: pain intensity, pain when weightbearing, physical function (activities and participation), joint movement, and treatment satisfaction. Two circumstance-specific domains were also included: structural pathology and healthcare expenses. Finally, four domains were considered important for future consideration: emotional wellbeing, sleep, gait, footwear.

Conclusion: This project developed a standardised, internationally agreed set of domains for measurement in future research. Its uptake has the potential to improve the quality of evidence for foot and ankle RMD treatments. Future work will focus on standardising outcome measurement instruments for the core domains. Crucially, the involvement of patients in the design and conduct of each phase of this project ensures that future research findings will reflect what outcomes are meaningful and relevant to those with foot and ankle disorders in RMDs.

Contents

Figures

Tables

Abbreviations

AI	Artificial intelligence
AS	Ankylosing spondylitis
AT	Achilles tendinopathy
BRC	Biomedical Research Centre
BSR	British Society for Rheumatology
CASP	Critical Appraisal Skills Programme
CERQUAL	Confidence in the Evidence from Reviews of Qualitative research
COMFORT	Core set of OutcoMes for FOot and ankle disorders in RheumaTic and musculoskeletal diseases
COMIT'iD	Core Outcome Measures in Tinnitus International Delphi
CONQual	Confidence in the Qualitative synthesis findings
COREQ	Consolidated Criteria for Reporting Qualitative Research
COMET	Core Outcome Measures in Effectiveness Trials
COS	Core outcome set
COS-STAD	Core Outcome Set STAndards for Development
COSMIN	COnsensus-based Standards for the selection of health Measurement Instruments
COSMOS	Core Outcome Set for Multimorbidity Studies
CPPD	Calcium pyrophosphate deposition disease
CROWN	Core Outcomes in Women's health
CTD	Connective tissue disease
EDI	Equality, diversity and inclusion
ENTREQ	Enhancing transparency in reporting the synthesis of qualitative research
GASTROS	GAstric cancer Surgery Trials Reported Outcomes Standardisation
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HCP	Healthcare professional

HOME	Harmonising Outcome Measures for Eczema
ICF	International Classification of Functioning, Disability and Health
IMMPACT	Initiative of Methods, Measurement and Pain Assessment in Clinical Trials
IPA	Interpretative phenomenological analysis
JBI	Joanna Briggs Institute
JIA	Juvenile idiopathic arthritis
LMIC	Lower and middle income country
MSK	Musculoskeletal
MTP	Metatarsophalangeal
NGT	Nominal group technique
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health and Care Research
OA	Osteoarthritis
OMERACT	Outcome Measures in Rheumatology
OMI	Outcome measurement instrument
PAG	Project advisory group
PHP	Plantar heel pain
PICO	Population, intervention, comparator, outcome
PIS	Participant information sheet
PPI	Patient and public involvement
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROM	Patient-reported outcome measure
PROQOLID	Patient-Reported Outcome and Quality of Life Instruments Database
PROSPERO	International Prospective Register of Systematic Reviews
PRP	Patient Research Partner
PsA	Psoriatic arthritis

PTTD	Posterior tibial tendon dysfunction
RA	Rheumatoid arthritis
RCT	Randomised controlled trial
RMD	Rheumatic and musculoskeletal disease
ROM	Range of motion
RP	Raynaud's phenomenon
SLE	Systemic lupus erythematosus
SpA	Spondyloarthropathy
SRQR	Standards for Reporting Qualitative Research
SSc	Systemic sclerosis
TAR	Total ankle replacement
UK	United Kingdom
USA	United States of America
WHO	World Health Organisation

Chapter 1 Introduction

"The feet are your foundation." - COMFORT participant

1.1 Introduction

This thesis reports a doctoral research project involving the development of a core domain set for foot and ankle disorders in rheumatic and musculoskeletal diseases (RMDs). This chapter summarises the background to the project, and outlines its aim and objectives. An overview of subsequent thesis chapters and the terminology that will be used throughout this work is then provided. Section 1.5.1 of this chapter is informed by a scoping review that was conducted, prior to PhD enrolment, in preparation for this project. Findings have been published as:

> **Chapman LS**, Jones J, Redmond AC, Flurey CA, Richards P, Hofstetter C, Smith TO, Arnold JB, Hannan MT, Maxwell LJ, Menz HB, Shea B, Golightly YM, Tugwell P, Beaton D, Conaghan PG, Helliwell PS, Siddle HJ. Developing a core outcome set for foot and ankle disorders in rheumatic and musculoskeletal diseases: A scoping review and report from the OMERACT 2022 foot and ankle special interest group session. Semin Arthritis Rheum. 2023 Aug;61:152210. https://doi.org/10.1016/j.semarthrit.2023.152210

1.2 The foot and ankle

The foot and ankle are complex and often neglected areas of the body. Composed of 26 bones, 33 joints, and over 100 muscles, tendons and ligaments, the foot and ankle bear weight, transmit the forces generated during movement, and allow for locomotion and balance. The ankle joint is made up of the tibia and fibula in the lower leg, and the talus. The foot is often subdivided for functional purposes into the rearfoot, midfoot and forefoot. The rearfoot is composed of the talus and the calcaneus (heel bone); the articulation of these two bones is the subtalar joint. The midfoot comprises of the navicular, cuboid, and three cuneiforms. The articulations of the calcaneus and the cuboid, and the talus with the navicular, are known as the midtarsal joints, whilst the articulation of the cuneiforms and cuboid with the metatarsal bones are the tarsometatarsal joints. Finally, the forefoot consists of the metatarsals and the phalanges, which articulate at the metatarsophalangeal joints (1: p. 1520-1529). Foot anatomy is presented in Figure 1.1.

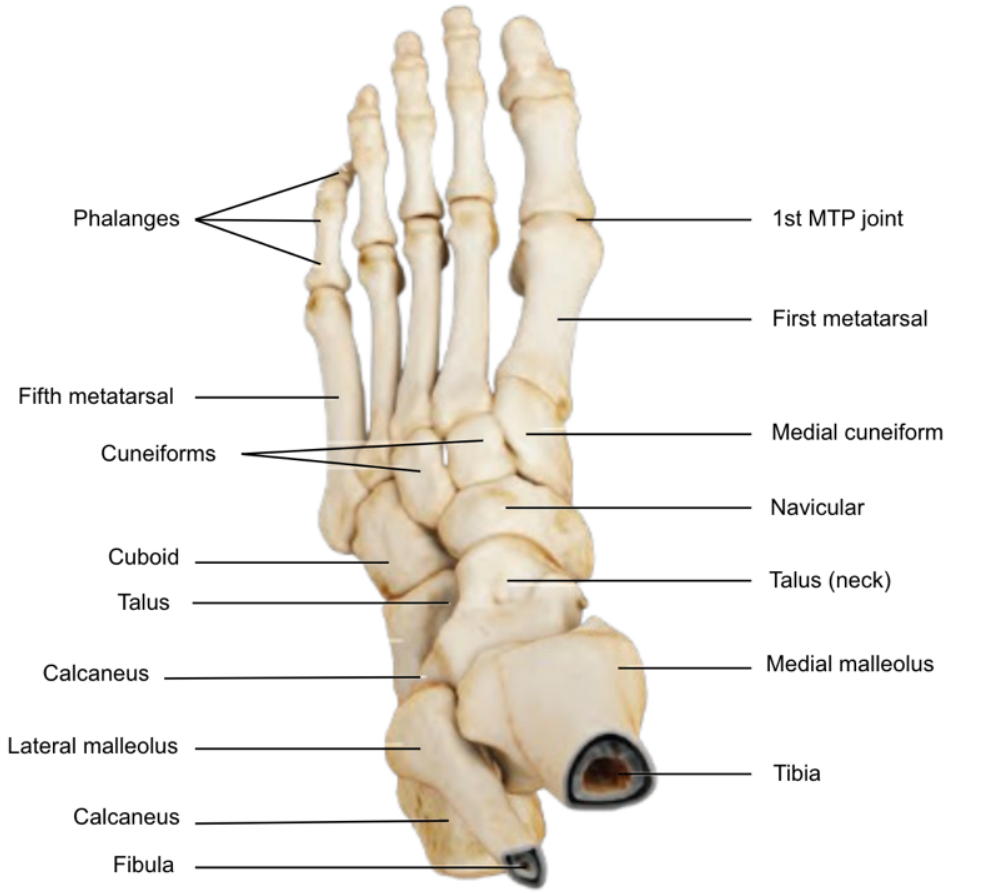


Figure 1.1: Anatomy of the foot and ankle.

Figure modified from Figure 82.9-F in Gray's Anatomy, 43rd ed. (1: p.1529), with the author's permission.

Key structures in the foot and ankle include the plantar fascia, Achilles tendon, tibialis posterior tendon, and peroneal tendons. The plantar fascia is a thick band of connective tissue that originates from the calcaneus and inserts into the toes. The plantar fascia supports the arch of the foot and absorbs shock during weightbearing activity. The Achilles tendon, which is the thickest tendon in the human body, connects the calf muscles of the leg to the calcaneus and allows for propulsive action. The tibialis posterior tendon connects the tibialis posterior muscle in back of the lower leg to the midfoot, stabilising the arch, whilst the peroneal tendons are located on the outside of the ankle, maintaining ankle stability (1: p. 1524-1536).

1.3 Rheumatic and musculoskeletal diseases (RMDs)

RMDs are chronic systemic diseases that commonly affect the joints. There are more than 200 different RMDs and over 1.63 billion people worldwide are estimated to live with an RMD worldwide (2). RMDs significantly impact on the economy and healthcare resources (3-7), and their prevalence is projected to continue to increase due to population growth and obesity (8). RMDs frequently affect the foot and ankle, leading to significant personal and economic burden; patients with foot and ankle problems in RMDs report high levels of pain and suffer severe, ongoing disability, leading to an overwhelming reduction in quality of life (9-16). The prevalence and characteristics of foot and ankle involvement in certain RMDs will now be discussed.

1.3.1 Inflammatory arthritis (IA)

1.3.1.1 Rheumatoid arthritis (RA)

RA is a chronic, progressive autoimmune condition causing synovial inflammation of the joints (17). Around 90% of people with RA experience foot involvement during the course of the disease, including rearfoot and forefoot deformity, tibialis posterior dysfunction (where inflammation of the tibialis posterior tendon leads to loss of function), peripheral arthritis, and subluxation or full dislocation of the MTP joints (9, 18). The development and severity of foot problems in RA increases with the duration of active inflammation (19). Mechanical stresses, as a result of weightbearing, also contribute to damage and deformity (20).

1.3.1.2 Adult juvenile idiopathic arthritis (adult JIA)

JIA is an umbrella term for different types of IA that occur in children. JIA can persist into adulthood, where the condition, including foot and ankle manifestations, presents similarly to RA (21: p.163-164).

1.3.1.3 Spondyloarthropathies (SpA)

SpA is an umbrella term for a group of inflammatory conditions primarily affecting the sacroiliac joints and cervical spine. The SpA umbrella includes ankylosing spondylitis (AS) (predominantly affecting the spine), psoriatic arthritis (PsA) (the arthritis associated with psoriasis), reactive arthritis (a form of arthritis that develops in response to an infection), and enteropathic arthritis (the arthritis of inflammatory bowel disease) (22: p.103). In SpA, foot problems include peripheral arthritis, dactylitis

(severe, uniform swelling of a toe), and enthesitis (inflammation of the entheses, where tendons and ligaments attach to bone), particularly at the Achilles tendon, plantar fascia and tibialis posterior tendon insertions (22: p.108-109). In PsA specifically, the foot is affected in up to 95% of people with the condition, with foot deformity and pain reported (23, 24).

1.3.1.4 Extra-articular features in IA

Extra-articular features of IA, including peripheral neuropathy and entrapment neuropathies, can also manifest in the foot (25, 26), and the risk of peripheral arterial disease is increased (27). Steroids are used in IA to control inflammation rapidly (28), but long-term steroid use can contribute to poor tissue viability in the foot and when combined with joint deformity and poor vascular supply, the risk of tissue breakdown is significantly increased (29). Foot ulcers are an important consideration in IA and immunosuppression, which can occur as a result of medications used to treat autoimmune conditions (e.g. steroids, disease-modifying antirheumatic drugs, and biologic therapies), increases the risk of potentially serious infection in the foot (30, 31).

1.3.2 Osteoarthritis (OA)

OA is classically characterised by a focal loss of articular cartilage, degradation of subchondral bone, and the development of osteophytes, although it is now recognised that OA is a whole-joint organ disease and any of the tissues of the affected organ can influence aetiology and progression (32). Symptomatic radiographic foot OA affects almost 17% of people aged over 50 (33), whilst the population prevalence of symptomatic radiographic ankle OA in the over 50s is 3.4% (34). The most commonly affected area in the foot is the 1st MTP joint, followed by the midfoot (33). Symptoms of OA include pain and stiffness at the affected joint, increasing with activity, as well as a progressive limitation in joint range of movement (32).

1.3.3 Crystal arthropathies

1.3.3.1 Gout

Gout is a disorder of purine metabolism, whereby the presence of monosodium urate crystals clinically manifests as acute attacks of peripheral synovitis, causing joint pain and swelling, but also has well-recognised chronic manifestations including an erosive, destructive arthropathy and a pathogenetic link with OA (35). Gout affects 1-4% of

adults worldwide (36). Foot involvement, particularly the 1st MTP joint, is the clinical hallmark of gout; almost 90% of patients with gout have 1st MTP joint involvement at some point in the disease course (33).

1.3.3.2 Calcium pyrophosphate deposition disease (CPPD)

CPPD, also known as pseudogout, involves deposition of calcium pyrophosphate crystals in articular cartilage. It presents similarly to acute gout, with signs of synovitis. The ankle joint is the third most commonly affected CPPD site, and acute inflammation of the Achilles secondary to CPPD can also occur (37: p.73-74).

1.3.4 Connective tissue diseases (CTDs)

1.3.4.1 Raynaud's phenomenon (RP)

RP is characterised by episodic vasospasm causing discoloration and pain in the digits, occurring upon exposure to cold temperatures or emotional stress (38, 39). Primary RP has an overall prevalence of ~5% in the general population (38, 39). Secondary RP is caused by an underlying systemic condition, as discussed below.

1.3.4.2 Systemic sclerosis (SSc)

SSc is a multi-system CTD characterised by excessive collagen production. An estimated 1.47 million people worldwide are affected by SSc (40, 41). The condition results in microvascular and macrovascular damage, with fibrosis of the skin and internal organs (42). RP, which occurs in the fingers and toes, is the hallmark feature of SSc, affecting ~96% of patients (43). RP can progress to tissue loss/ulceration (44). The foot can also be affected by callus formation and subcutaneous calcinosis, tendinopathies, joint space narrowing, bone demineralisation, joint subluxation, joint margin erosions, and degenerative changes in SSc (45-47).

1.3.4.3 Systemic lupus erythematosus (SLE)

SLE is a systemic autoimmune disease with multi-organ involvement, affecting at least 3.4 million people worldwide (48). Whilst foot and ankle involvement in SLE has not been extensively studied, a survey investigating patterns of self-reported foot complaints indicated that 77% of 131 respondents with SLE reported experiencing foot

pain (49). Extra-articular foot complaints were also reported, including cold feet, swelling and numbness. Approximately one third of SLE patients experience RP (45).

1.3.5 Non-systemic musculoskeletal (MSK) conditions

There are multiple chronic MSK conditions affecting the foot and ankle in the absence of systemic disease. Three MSK conditions commonly seen in clinical practice will be discussed in this section. These are plantar heel pain, tendinopathies, and hallux valgus.

1.3.5.1 Plantar heel pain (PHP)

PHP, also known as plantar fasciitis, is broadly characterised by pain under the heel, which is worse when people take their first steps in the morning and after longer periods of weightbearing (50). Around 1 in 10 people aged over 50 experience PHP (51). The most frequent aetiology is mechanical overload (50, 52).

1.3.5.2 Tendinopathy

Tendinopathy is a broad term encompassing various tendon problems, including tendinitis (inflammation) and tendinosis (degeneration), and causes pain, swelling, and thickening of the tendon. In the foot and ankle, tendinopathy most frequently occurs in the Achilles, peroneal, and tibialis posterior tendons. For Achilles tendinopathy (AT), an incidence rate of 2.35 per 1,000 has been reported in the adult population (21-60 years) (53), although data for other types of tendinopathy is lacking. Tibialis posterior tendinopathy can progress to a weakened tendon, loss of muscle power and destabilisation of the foot, known as tibialis posterior tendon dysfunction (TPPD), posterior tibial tendon dysfunction (PTTD), or adult acquired flatfoot (54: p.145).

1.3.5.3 Hallux valgus

Hallux valgus is characterised by the lateral deviation of the hallux towards the lesser toes, disrupting the alignment of the 1st MTP joint, eventually causing subluxation of the 1st MTP joint and formation of an osseus prominence (bunion). Hallux valgus affects 23% in people aged 18-65 years and 36% in people aged over 65 years (55). Over 10,000 hallux valgus surgeries are carried out in England every year (56).

1.4 Interventions for foot and ankle disorders in RMDs

Foot and ankle disorders in RMDs have gained attention in recent years, with an increase in the number of clinical trials and observational studies investigating interventions for these conditions. Pharmacological (e.g. local steroid injections and systemic drug therapies), conservative (e.g. footwear, insoles, and exercises), and surgical interventions have been shown to reduce foot and ankle symptoms in people with RMDs (57, 58), although there is a lack of clinical trials assessing the effectiveness of treatments for foot and ankle disorders in some RMDs (e.g. SpA, CTDs).

National Institute for Health and Care Excellence (NICE) guidelines for the management of RA state that functional insoles (example in Figure 1.2) and therapeutic footwear (example in Figure 1.3) should be available for all adults with RA if indicated (59). Recent British Society for Rheumatology (BSR) guidelines for the management of foot health in IA (including RA, SpA, and JIA) similarly recommend the use of customised insoles and specialist footwear, as well as localised steroid injections and targeted exercises (60).

Figure 1.2: Example of functional insoles.

Candidate's own image.

Figure 1.3: Example of therapeutic footwear.

Candidate's own image.

These foot and ankle RMD guidelines have been informed primarily by consensus among experts, due to either an absence of clinical trials for a specific intervention or a lack of high-quality evidence from the clinical trials that have been conducted.

1.5 Quality of evidence

As outlined above, in RMDs where clinical trials have been conducted, there is a lack of high-quality evidence relating to the efficacy of interventions for foot and ankle disorders and existing research outputs are not always translated into practice (61, 62). A key contributing issue relates to outcome measurement. Outcome measurement encompasses domains (*what* is being measured) and outcome measurement instruments (OMIs) (*how* a domain is measured) (63). Multiple systematic reviews have highlighted how diversity of domains and OMIs used across trials, and selective outcome reporting, hinders the ability to pool data (57, 58, 64). Additionally, it is widely agreed that if research findings are to influence policy and practice, the selection of study domains needs to be important, relevant and feasible for all key contributors, including patients (63, 65-68). However, domains are often specified unilaterally by researchers and healthcare professionals (HCPs), reflecting their priorities, and previous studies have shown perspectives and priorities differ between patients and HCPs (69-72).

1.5.1 Domain and OMI heterogeneity

The extent of domain heterogeneity in foot and ankle RMD research to date has been explored in a scoping review of domains in clinical trials and observational studies, undertaken in preparation for the research within this thesis (73). In 150 interventional studies involving people with RA, SpA, OA, gout, and CTDs, a total of 63 different domains were identified. Foot or ankle pain was the most commonly measured domain (n = 118, 78% of studies). There was considerable heterogeneity in the other domains measured. An overview of the domains identified in the scoping review is presented in Table 1.1. Domains are organised according to core areas within the OMERACT framework (63): pathophysiological manifestations (signs, symptoms, biomarkers), life impact, death, and societal/resource use. This framework will be discussed in further detail in Chapter 2.

Table 1.1: Domains measured in existing research

Pathophysiological manifestation domains (n studies)
Signs
Joint range of motion (30)
Alignment (24)
Global disease activity/assessment of overall condition by clinician (16)
Joint swelling (8)
Foot/ankle disease activity (6)
Presence of deformity (5)
Presence of callosities (5)
Pain upon palpation (4)
Joint tenderness (4)
Pressure-pain threshold (2)
Muscle strength (2)
Joint stability (2)
Muscle activity (1)
Joint girth (1)
Joint temperature (1)

Clinician-assessment of gait (1)
Symptoms
Foot/ankle pain (118)
Pain during weightbearing (14)
Pain during non-weightbearing (9)
Foot/ankle pain severity/intensity (7)
Pain at night (4)
Pain on provocation (4)
General pain (21)
Joint stiffness (10)
Stiffness after rest (7)
Stiffness during weightbearing (5)
Fatigue (8)
Patient global change in foot/ankle symptoms (6)
Joint catching (4)
Joint grinding (4)
Patient-reported disease severity/assessment of overall condition by patient (2)
Biomarkers
Disease progression/deformity on imaging (51)
Gait (39)
Temporal spatial parameters (21)
Plantar pressure (24)
Kinematics (18)
Kinetics (12)
Disease activity on imaging (9)
Disease activity on laboratory markers (6)
Life impact domains (n studies)
Foot/ankle function or disability (102)

Global function or disability (50)
Overall quality of life/health status (37)
Social function (12)
Emotional status (10)
Sports participation (8)
Footwear requirements (7)
Foot/ankle related quality of life/health status (5)
Pain interference (1)
Treatment satisfaction (36)
Death domains (n studies)
Survival (1)
Adverse events (96)
Societal/resource use domains (n studies)
Healthcare utilisation (28)
Need for revision/incidence of device failure/implant survivorship (18)
Hospitalisation/investigative procedures (1)
Length of hospital stay (3)
Hospital readmissions (1)
Visits to healthcare providers (11)
Accident and Emergency (A&E) attendances (1)
Medication/rescue medication use/need for secondary interventions (15)
Direct/indirect costs (11)
Return to work time (1)
Days unable to work (1)
Cost-effectiveness (2)

Table modified from Chapman et. al (73).

Whilst the focus of the scoping review was domains, the OMIs used to measure domains were also extracted from the included studies and published in the supplementary material of the paper (73: supplementary table 2). The extracted data demonstrates extensive OMI heterogeneity, with a wide range of instruments used to measure each domain across the studies. This mirrors available systematic reviews of OMIs for foot and ankle disorders in RMDs; in eight systematic reviews registered within the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) database for OMI measurement property studies (74-81), a total of 87 different OMIs for foot and ankle disorders in RMDs were included for assessment (Appendix A). The COSMIN database will be discussed in Chapter 2.

1.5.2 Core outcome set as a solution

A significant factor contributing to the issues presented above is the absence of a core outcome set (COS) for clinical studies involving patients with foot and ankle disorders in RMDs. A COS, also known as a core domain set, is defined as “an agreed standardised set of outcomes/domains that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or healthcare” (65: pg.1). A core domain set focuses on *what* to measure. Core OMIs can then be selected for these core domains (*how* to measure) (63). A key factor when developing a COS is the involvement of all key contributors (e.g. patients, HCPs, and researchers), to ensure clinically relevant outcomes are identified (63, 65-68). There are three key benefits to developing and using a COS. Firstly, when trials addressing similar clinical questions measure the same outcome, using the same tool, results can be combined in systematic reviews and meta-analyses, leading to more conclusive evidence. Secondly, use of a COS can prevent selective reporting of outcomes, as the outcomes to be reported are predetermined. This reduces the introduction of bias. Finally, as COS measure outcomes that are relevant to different research users, study results are more clinically meaningful and trials translate into improved clinical care (65, 66, 68, 82).

1.6 Aims and objectives

The overall aim of this project, 'Development of a Core set of OutcoMes for FOot and ankle disorders in RheumaTic and musculoskeletal diseases' (COMFORT), is to develop a core domain set for foot and ankle disorders in RMDs.

The objectives are to:

1. Identify and explore domains of importance to patients with foot and ankle disorders in RMDs.
2. Achieve multidisciplinary, multi-contributor, and expert international consensus on a core domain set for foot and ankle disorders in RMDs.

It is anticipated that the core domain set developed within the COMFORT project will increase consistency in outcome measurement for research investigating the effectiveness or efficacy of interventions for foot and ankle disorders in RMDs, and improve the transferability of findings into practice, leading to improved evidence on treatments for patients.

1.7 Exploratory hypothesis

An exploratory hypothesis was developed to guide interpretation throughout this thesis. The hypothesis was that by using a mix of methods and with input from multiple contributors (including patients) internationally, it would be possible to develop a core set of domains to be measured and reported in future research involving foot and ankle disorders in a range of RMDs.

1.8 Thesis structure

Chapter 2: Narrative literature review

This chapter provides an overview of core domain set development methods, encompassing methods to identify potentially important domains, consensus methods and OMI selection methods. An introduction to the COMFORT project is then presented.

Chapter 3: Systematic review and thematic synthesis of existing qualitative studies

This chapter reports a systematic review and qualitative synthesis of 34 existing qualitative studies involving patients with foot and ankle disorders in RMDs. This study aimed to address project objective 1 by exploring domains of potential importance to patients, and to identify any gaps in the existing qualitative literature base. Data were analysed using thematic synthesis.

Chapter 4: Secondary analysis of focus groups

This chapter reports a secondary qualitative analysis of transcripts from six focus groups, conducted in two countries, involving 28 patients with one of the of the under-researched RMDs identified in Chapter 3 (systemic sclerosis-related Raynaud's phenomenon affecting the feet), also addressing project objective 1. Data were analysed using directed content analysis.

Chapter 5: Qualitative interview study

This chapter details an international, qualitative study involving 56 individual, semi-structured interviews with patients, with a focus on RMDs, geographic locations, ethnicities and first languages that have been under-represented in previous qualitative research. It aimed to further address objective 1 by identifying any additional domains of potential importance to patients living with foot and ankle disorders in RMDs, through a direct exploration of their experiences. Data were analysed using the framework method.

Chapter 6: Online Delphi consensus study

An international, four-round, online modified Delphi study, which addressed project objective 2, is presented in Chapter 6. Round 1 consisted of a long list of domains derived from the findings from existing literature, including the qualitative work in Chapters 3-5. Patients, HCPs, and researchers rated the importance of each domain and generated additional domains. Quantitative data were analysed using descriptive statistics, whilst qualitative data were analysed using content analysis. In the final round, participants rated domains as 'in' or 'out' and consensus on a final core set of domains was achieved.

Chapter 7: Discussion and conclusion

This chapter discusses key findings of the overall project, strengths and limitations, and implications for future research and clinical practice. The project is then concluded.

1.5 Terminology

The terminology used by COS developers varies. Within this thesis, the term *domain* is used to refer to the health outcome being measured, whilst the term *OMI* is used to refer to the instrument measuring a domain. The terms *core domain set* (a core set of domains to be measured) is used throughout this thesis in relation to the COMOFRT project (63), whilst the term COS is used when referring to the general concept or discussing other studies where the developers have used this term.

In recognition of the recent shift away from the term *stakeholde*r (83, 84), *key contributor* is used throughout this thesis to denote different groups of research participants (e.g. patients, HCPs, researchers) who are actively working towards achieving the goal of developing the COS. In contrast, whilst there has also been a shift from the term *patient* to *person* when referring to an individual with a health condition, with the latter term emphasising an individual's broader identity (85), the term *patients* will be used throughout this thesis when referring specifically to this group of key contributors in the context of COS development. Finally, the term *patient and public involvement* (PPI) will be used throughout this work, following the National Institute for Health and Care Research (NIHR) definition, unless the name of a specific PPI group using a different term for PPI is referred to (e.g. the OMERACT Patient Research Partner (PRP) Network). NIHR defines PPI as "research that is being carried out 'with' or 'by' patients or members of the public rather than 'to', 'about' or 'for them" (86).

Chapter 2 Narrative literature review

2.1 Introduction

Chapter 1 provided an overview of foot and ankle disorders in RMDs. It highlighted a paucity of high-quality evidence for treatments for these disorders, linking to inconsistency in the domains measured in research in this area to date. Development of a COS was identified as a solution to this problem. The following chapter provides an overview of COS development methods and introduces COMFORT, the project that forms this thesis. A protocol for the COMFORT project is published in *Trials* (87):

> **Chapman LS**, Redmond AC, Flurey CA, Richards P, Smith TO, Arnold JB, Beaton D, Conaghan PG, Golightly YM, Hannan MT, Hofstetter C, Maxwell LJ, Menz HB, Shea B, Tugwell P, Helliwell P, Siddle HJ. Developing an Outcome Measures in Rheumatology (OMERACT) Core set of Outcome Measures for FOot and ankle disorders in RheumaTic and musculoskeletal diseases (COMFORT): core domain set study protocol. Trials. 2023 Jan 28;24(1):65. https://doi:10.1186/s13063-023-07104-7

2.2 COS development initiatives

COS development and methodology has been driven by three key initiatives to date: the Outcome Measures in Rheumatology (OMERACT) initiative, the Core Outcome Measures in Effectiveness Trials (COMET) Initiative, and the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative.

OMERACT is a global, volunteer-driven organisation that aims to improve outcome measurement in rheumatology. Through global working groups, OMERACT supports the development of COS, identifying patient and disease-relevant outcome domains and corresponding outcome measurement instruments (OMIs), primarily for use in clinical trials (63). OMERACT working groups are composed of multi-contributor representatives from multiple continents and focus on different areas, including specific rheumatic diseases, body regions, imaging, biomarkers, and methodologies. Since its establishment in 1992, OMERACT has played a critical role in developing COS through data-driven, iterative consensus processes. OMERACT hosts biennial meetings, bringing contributors together to advance COS methodology. OMERACT has published

resources for COS developers, including a handbook (63). OMERACT focuses on both *what* to measure (domains) and *how* to measure (OMIs).

The COMET Initiative unites researchers who are developing COS, collating resources in a website (88). COMET was formed in 2010 and has six specific objectives:

- To raise awareness of current problems with outcomes in clinical trials;
- To encourage COS development and uptake;
- To promote PPI in COS development;
- To provide resources to facilitate these aims;
- To avoid unnecessary duplication of effort;
- To encourage evidence-based COS development.

The COMET website provides a searchable online database; a repository of studies relevant to COS development. The website indicates that there are currently over 800 published and ongoing COS across various areas of health. The COMFORT project is published in the COMET database (89). The COMET website also provides relevant publications related to COS methodology, plain language summaries, PPI resources, and links to relevant networks and events. COMET has also published a handbook that presents different methods of developing a COS, with a specific focus on domains (65). The handbook, published in 2017, does not recommend any specific methods, highlighting the limited empirical evidence in this area. COMET's work to address the evidence gaps in COS methodology is ongoing.

Whereas COMET focuses on domains, COSMIN aims to improve the selection of OMIs in research (including the selection of OMIs for domains within COS) and clinical practice (90). COSMIN develops methodology and tools for selecting the most appropriate OMI amongst competing instruments, by collating systematic reviews of measurement properties of available OMIs, and providing checklists for measurement properties and standards for evaluating these properties. Collaborations between COSMIN, COMET, and OMERACT are well established.

While COS development and methodology has primarily been advanced by the initiatives discussed above, other COS initiatives have been established in various areas of health. For example, following advances in COS development in rheumatology, the Harmonising Outcome Measures for Eczema (HOME) initiative has focused on developing core outcomes for eczema trials (91). Additionally, the Initiative of Methods, Measurement and Pain Assessment in Clinical Trials (IMMPACT) aims to

develop consensus reviews and recommendations for acute and chronic pain trials (92), and the Core Outcomes in Women's health (CROWN) initiative is developing core outcome sets for key conditions in women's health (93). These initiatives draw on methodological aspects from OMERACT, COMET, and COSMIN.

2.3 Patient and public involvement (PPI)

PPI input into COS development has evolved significantly over time (94). PPI contributors can improve the relevance of a COS by offering their perspectives during the COS development process, by contributing to study design, advising on and helping with participant recruitment, data collection and analysis, and disseminating results (95).

PPI has been a fundamental part of OMERACT since 2002. OMERACT PPI contributors, known as patient research partners, are active research team members with RMDs who work with researchers to add their experiential knowledge during COS development (63). OMERACT PPI contributors are integrated at each phase of the COS development process, with equal opportunities to fully participate. OMERACT principles and recommendations for PPI involvement in COS development have been developed through consensus (96); these include the unconditional requirement for incorporating the patient perspective when developing disease specific COS and patient-reported outcome measures (PROMs). Inclusion of a minimum of two PPI contributors is recommended for each OMERACT working group (63). The impact of PPI involvement on the scope and conduct of COS development in rheumatology is routinely evaluated by OMERACT (97-100). PPI involvement has widened OMERACT's research agenda by identifying previously neglected domains such as fatigue, sleep disturbances and flares, led to the inclusion of patient-relevant domains in COS, and enhanced PROMs.

In accordance with OMERACT, COMET suggests seven key benefits of including PPI contributors in COS development (65):

- Provision of advice on the best ways of identifying and accessing particular patient populations;
- Informing discussions about ethical aspects of the study;
- Facilitating the design of more appropriate study information;
- Promoting the development of more relevant materials to promote the study;
- Enabling ongoing troubleshooting opportunities for patient participation issues during the study, e.g. recruitment and retention issues of study participants;

- Informing the development of a dissemination strategy of COS study results for patient participants and the wider patient population;
- Ensuring the COS is relevant to patients, and that patients can trust that the development process has genuinely taken account of the patient perspective.

COMET has produced a toolkit for COS developers to consider when involving PPI in their work (101). The toolkit covers planning, supporting, and recording/evaluating PPI. OMERACT has also produced a PPI toolkit (102), but this is primarily aimed at PPI contributors who are taking part in developing OMERACT COS, rather than at COS developers. Future work is needed to understand researchers' and PPI contributors' experiences of using these toolkits and to measure their impact.

2.4 Key contributor group engagement

If research findings are to translate to clinical practice, domains measured in trials and other studies need to reflect what is important to all key contributor groups. Engagement with key contributors is important throughout COS development; uptake of the end product by a specific contributor group is unlikely if it does not represent their perspectives and priorities. Typically, patients, HCPs, and researchers (including trialists and methodologists) are key contributors in any COS development process, but input from clinical commissioners, guideline developers, journal editors, industry representatives, and commissioners may also be considered (65). A COMET-led systematic review of studies assessing the impact of patient participation on COS development across multiple health settings identified that COS involving patients were more likely to include domains relating to life impact than those developed without patient participation (94).

To facilitate comparison of studies in a specific area of health, a COS needs to be applicable and adopted across relevant settings and disciplines (63, 65). Therefore, global contributor engagement is important. OMERACT states that contributor representation must come from at least three continents; frequently, the continents involved are North America, Europe, and Australia/Oceania. A systematic review conducted in 2020 found that only 74 (20%) of 370 COS studies included participants from low and middle income countries (LMICs) (predominantly China, Brazil, and South Africa), and only four of these were initiated from LMIC settings (103). In a systematic review of COS Delphi studies, patients from LMICs were under-represented compared with HCPs (104). Limited engagement of contributors from LMICs in COS development may limit its use in these countries, given notable differences in disease patterns, burden, healthcare systems, and priorities among patients, HCPs, and researchers

(105). Challenges of including LMIC contributors in COS development have been postulated, including lack of resources and networks to support involvement (103, 105). A cross-sectional survey with LMIC respondents proposed strategies to improve LMIC contributor engagement, such as enhancing collaborations (105). Recently, various COS developers have started to focus on LMICs, particularly during COS consensus processes. For example, the Gastric cancer Surgery Trials Reported Outcomes Standardisation (GASTROS) study developed a COS for gastric cancer surgery and grouped results according to region (East or West) and country income classification, to determine how groups compared with respect to their categorisation of 56 potentially important domains. However, the researchers found little variation in domain prioritisation between regions or income classification (106). Similarly, a COS for burn care demonstrated considerable agreement between participants from LMICs and high-income countries across domains during a consensus process (107). Nevertheless, diversity among the key contributors inputting into the COS development process remains important, ensuring that all perspectives are represented.

Engagement with contributors from LMICs can also be considered at an earlier stage of the COS development process. For example, the Core Outcome Set for Multimorbidity Studies (COSMOS) study developed two COS for trials of interventions for multimorbidity prevention and treatment, specific to LMIC contexts. This involved patient interviews conducted by local teams, in local languages, in South Asia, Africa and Latin America, to identify relevant domains for prioritisation (108).

2.5 COS development methods

There is no gold standard process for developing a COS, but 11 minimum standards, known as Core Outcome Set STAndards for Development (COS-STAD), have been derived through an international consensus exercise with experienced COS developers, methodologists, journal editors, potential users of COS (clinical trialists, systematic reviewers, and clinical guideline developers), and patient representatives (109).

COS-STAD focus on defining the scope of the COS, deciding which key contributors to involve and how to involve them, and ensuring transparency in the consensus process. When these standards were developed, there was no consensus on how to identify potential domains, or how to select OMIs for domains. Notwithstanding, all COS initiatives agree that, after defining the COS scope (the applicable area of health to which the COS is to be applied), COS development is a two-step process that involves

determining what domains to measure, and determining which OMIs should be used. The scope of a COS can be described in the population, intervention, comparator, outcome (PICO) format, as per clinical trials, specifying the type of patients and interventions/comparators that a clinical trial using the COS would include (e.g. the health condition of interest and specific/any interventions). The context/setting (e.g. research and/or clinical practice, study type) to which the COS will be applied should also be confirmed (65). The next phase of COS development involves identifying domains that should be considered for inclusion in the COS. Most frequently, this has involved reviews of existing quantitative and qualitative literature, and primary qualitative research with key contributors. The result is a list of domains that are prioritised using consensus methods, resulting in a core set of domains (65). The methodological considerations within each phase of the core domain set development process will now be discussed.

In addition to standards for COS development, guidance for COS reporting has also been produced (110). The Core Outcome Set-STAndards for Reporting (COS-STAR) reporting guideline consists of an 18-item checklist, developed through consensus.

2.5.1 Review of existing quantitative studies

Potential domains for inclusion in a COS can be identified by reviewing existing and planned clinical trials and other quantitative studies in the area of health specified in the COS scope (65). Additionally, a review can identify domain heterogeneity, thus justifying the need for a COS. COS developers may conduct a systematic review or scoping review to generate a list of domains previously measured and reported.

A systematic review can be defined as using “explicit, systematic methods that are selected with a view to minimising bias, thus providing more reliable findings from which conclusions can be drawn and decisions made” (111: section 1.1). Following a structured and predefined process, systematic reviews are routinely used to inform practice and policy by identifying and synthesising existing evidence that is relevant to a specific question, and establishing the quality of this evidence. Researchers are encouraged to register systematic reviews in the international prospective register of systematic reviews (PROSPERO), which is an international database for review protocols. This aims to promote transparency, reduce reporting bias, and prevent review duplication (112). Comparatively, a scoping review also follows a structured and predefined process, but differs in terms of its aims and methodology. Originally introduced by Arksey and O’Malley in 2005 (113), this methodology provides a broad overview of a topic by comprehensively mapping existing literature, regardless of its

quality. Scoping reviews typically aim to identify key concepts and establish gaps in knowledge (113).

In the context of COS development, a systematic review of domains can be undertaken to synthesise evidence contributing to the COS, but is less inclusive in terms of the breadth of identified domains. In contrast, a scoping review aims to be broad, increasing understanding of the range and scope of domains reported in existing literature, rather than assessing risk of bias (114). In both cases, COS developers must consider the search strategy and inclusion/exclusion criteria, aligning with the scope of the COS, in addition to screening, selection, and data extraction processes, and data synthesis methods. COS developers may limit the review timeframe to determine which domains have been reported in recent studies only, or to certain study types, e.g. solely randomised controlled trials (RCTs) (115). However, potentially important domains may be missed, as domains measured in different types of study can differ (116). COS development guidance proposes that when conducting a scoping or systematic review of domains, the domains should be extracted verbatim, and the extraction and classification of domains should be conducted independently by at least two researchers (65).

2.5.2 Qualitative research

Domains measured in clinical trials and other research studies typically reflect those that are important to researchers, and sometimes HCPs, but these are not always domains of relevance or importance to patients. Patients are not always consulted regarding what domains are measured in studies, therefore their priorities may be overlooked if reviews of domains in existing studies are solely used to inform the development of a COS. Qualitative research can lead to the identification of additional or different domains compared to those identified in a scoping or systematic review of quantitative literature, particularly when quantitative studies lacked PPI input, or used PROMs that were developed with limited or no patient participation. This has been consistently acknowledged in COS development studies (72, 117). COMET and OMERACT both recommend that qualitative research is conducted during COS development, to identify domains for potential inclusion in a COS that may have been overlooked by researchers (63, 65). In addition to identifying relevant domains, qualitative research can also increase understanding of *why* certain domains are important to contributors, which can facilitate consensus later in the process. Additional motives to conduct qualitative research during COS development have been proposed (118), including the ability to determine the scope of domains. This ensures the full

breadth of a domain (e.g. the severity, frequency, type, and impact of pain experienced) is captured. Qualitative research also facilitates the identification of appropriate language for a subsequent domain prioritisation study; patients are able to conceptualise domains by describing their experiences and priorities in their own language.

In qualitative research, epistemology, the philosophical study of knowledge, must be considered. In qualitative research, epistemology influences the research question, how data is collected, the role of the researcher and the relationship with participants, and interpretation of findings. The most common epistemological paradigms in qualitative research include interpretivism, constructivism, and pragmatism. Interpretivism is based on the idea that knowledge is socially constructed, and multiple subjective realities exist. The researcher interprets meaning from participants' experiences and perspectives (119-121). Constructivism is similar, but assumes that the researcher and what they are researching are linked and interact; knowledge is co-constructed by the researcher and participant. Pragmatism is a practical epistemological paradigm that focuses on problem solving and producing knowledge that is meaningful in relation to the research problem; the researcher focuses on addressing a research question in the real world using whatever approaches work best to do this (122). COS development aligns to a pragmatic epistemology, as it uses methods that focus on solving a problem: lack of standardisation in outcome measurement. As the development of a COS involves identifying and prioritising domains that are meaningful and acceptable to different key contributor groups, an approach that utilises a mix of methods can capture different types of data, balance differing perspectives and give a more comprehensive understanding of domains and OMIs overall (65). Qualitative research conducted to inform a COS can include a review of existing qualitative studies, a secondary analysis of existing qualitative data, and/or primary qualitative research (63, 65).

2.5.2.1 Review of existing qualitative studies

There is increasing interest in qualitative data synthesis, particularly for informing healthcare policy and practice, as it can be used to explore existing literature, and combine multiple primary qualitative studies related to a specific research question to provide additional insight that would not be possible from an individual study alone (123). Whilst systematic review methods are well established in quantitative research (e.g. RCTs), development of systematic review methods for qualitative research is ongoing (124), as is the debate as to whether qualitative research *should* be

synthesised, given its overarching aim to provide depth of understanding rather than general conclusions. Qualitative studies do not aim to be generalisable, and are heavily influenced by the context, researcher, and participant (125). Instead, their value lies in transferability: the extent to which findings can be applied in other contexts (126).

Some qualitative researchers therefore argue that synthesis can lead to misinterpretations, and as each qualitative study typically reflects a unique social reality, it is not appropriate to combine multiple studies. As qualitative research can involve many different frameworks, designs and participants, methodological diversity also presents an issue when considering synthesis (127). The main counter argument is that reviewing and synthesising qualitative studies systematically, enhances their potential to inform the development of evidence-based treatments, healthcare policy, and practice (123). An overview of the most commonly used qualitative synthesis methods are presented in Table 2.1.

Table 2.1: Qualitative synthesis methods

Synthesis method	**Overview**
Meta-ethnography	Meta-ethnography aims to re-interpret conceptual data from each primary qualitative study. Primary data where terms and concepts are similar is taken into consideration and 'translated' to create higher order constructs, generating new theory (128).
Grounded theory	Data from multiple qualitative studies are synthesised with the aim of generating a new, higher-level grounded theory that extends beyond the findings of any single study. This process involves comparing and integrating themes, categories, and concepts across different studies to identify overarching patterns and theoretical insights (129).
Critical interpretive synthesis	Involves an iterative approach to refining the research question, searching and selecting from the literature, using theoretical sampling, and defining and applying codes and categories. It adapts meta-ethnography and grounded theory techniques, with a focus on likely contribution to theory development as opposed to methodological characteristics as a means of determining the quality of individual papers (125).

Meta-aggregation	Meta-aggregation synthesises qualitative findings from multiple studies without reinterpreting them, to produce a set of recommendations or action statements for healthcare practice or policy. (130)
Meta-narrative	Meta-narrative highlights the similar and contrasting ways in which researchers have conceptualised the same or a similar topic (131). Comparable studies are grouped together according to their underlying narratives, assumptions, and conceptual frameworks (132).
Meta-study	Meta-study seeks to reveal similarities and discrepancies among accounts of a particular phenomenon through three elements of analysis: meta-data-analysis (the analysis of findings), meta-method (the analysis of methods), and meta-theory (the analysis of theory). A new interpretation is created which accounts for the results of all three elements of analysis (128).
Textual narrative synthesis	Studies are arranged into more homogenous groups. Study characteristics, context, quality and findings are typically reported according to a standard format and similarities and differences are compared across studies. Structured summaries may also be developed, elaborating on and putting into context the extracted data (133).
Thematic synthesis	Free codes of findings from different studies are organised into descriptive themes, then further interpreted to form analytical themes. Findings are summarised in a narrative (134).
Framework synthesis	Framework synthesis is a highly structured approach to organising and analysing data, using a predefined framework informed by existing literature and team discussions to extract and synthesise findings across studies. Framework synthesis is a largely deductive approach, but allows new topics from the data to be incorporated (135).

Quality assessment

Quality assessment is another key consideration for COS developers conducting qualitative syntheses. A systematic review of qualitative studies can include a quality assessment of primary studies (136). Critical appraisal tools are commonly used to

evaluate the methods applied to minimise biases in quantitative studies (137), although it is still debated as to whether or not critical appraisal is appropriate in qualitative research. In parallel to the issues around synthesising qualitative studies, critical appraisal in this context has been challenged by some qualitative researchers, particularly as there is no consensus on how to judge quality (128). However, being able to critically appraise the methodological strengths and limitations of qualitative research with a set of criteria, can permit a systematic and transparent evaluation of a study's contribution to knowledge (136). Many tools are available for critically appraising the quality of qualitative studies, although there is no consensus on how to judge the methodological quality of qualitative studies being synthesised; a review by Munthe-Kaas et al. (2019) identified 102 unique critical appraisal tools, with many lacking rationale for their development (138). The authors developed a framework to compare criteria across tools, and identified that no single tool covered all themes. Dalton et al. (2017) reviewed 145 qualitative evidence syntheses and found that 92% of 145 included papers reported a quality appraisal of included studies, with 30 different tools identified (139). The most commonly reported tools were the Critical Appraisal Skills Programme (CASP) checklist (49 reviews) and the Joanna Briggs Institute Qualitative Review and Assessment Instrument (18 reviews). The CASP tool has been acknowledged by the Cochrane Handbook for use in qualitative evidence synthesis (136).

Confidence in qualitative synthesis findings

The value of assessing the confidence in synthesised findings from qualitative systematic reviews has been recognised by the World Health Organisation (WHO) (124), although in line with the argument against synthesising qualitative research presented above, reducing context-sensitive insights to a single confidence rating may lead to misinterpretation and risks diminishing nuances. Two key approaches that focus on systematically evaluating the trustworthiness of synthesised findings from qualitative systematic reviews will now be discussed: Grading of Recommendations Assessment, Development and Evaluation (GRADE-CERQual) and Confidence in the Qualitative synthesis findings (ConQual).

GRADE-CERQual is well established in the provision of guidance to assess evidence in quantitative research (140). In parallel, the GRADE-Confidence in the Evidence from Reviews of Qualitative research (GRADE-CERQual) approach is a systematic method for assessing how much confidence can be placed in findings from qualitative evidence syntheses, which is particularly pertinent when a synthesis is being used to inform guidelines, policy or practice (141). GRADE-CERQual is comprised of four components

(Table 2.2). Based on these components, a confidence level is assigned to each finding from the review (Table 2.3).

Table 2.2: GRADE-CERQual component definitions

GRADE-CERQual component	**Definition**
Methodological limitations	The extent to which there are concerns about the design or conduct of the primary studies that contributed evidence to an individual review finding.
Coherence	An assessment of how clear and cogent the is fit between the data from the primary studies and a review finding that synthesises the data.
Adequacy	An overall determination of the degree of richness and the quantity of data supporting a review finding.
Relevance	The extent to which the body of data from the primary studies supporting a review finding is applicable to the context specified in the review question.

Table reproduced from Lewin et al. (141).

Table 2.3: GRADE-CERQual confidence levels

Confidence level	**Definition**
High confidence	It is highly likely that the review finding is a reasonable representation of the phenomenon of interest.
Moderate confidence	It is likely that the review finding is a reasonable representation of the phenomenon of interest.
Low confidence	It is possible that the review finding is a reasonable representation of the phenomenon of interest.
Very low confidence	It is not clear whether the review finding is a reasonable representation of the phenomenon of interest.

Table reproduced from Lewin et al. (141).

Limitations of GRADE-CERQual include subjectivity in judgements of the component and confidence levels, which can mean that different reviewers report different confidence levels for the same finding, as there is no numeric score to combine judgements (142). Additionally, the process can take a lot of time and effort, depending on how many papers are included, and is dependent upon the quality and completeness of reporting in the primary qualitative studies (143).

An alternative to the GRADE-CERQual approach is ConQual, developed by the Joanna Briggs Institute (JBI) (144). This primarily aims to assess the confidence in synthesised findings from meta-aggregation, but is less flexible for use with interpretive methods (e.g. meta-ethnography or grounded theory synthesis). ConQual does not assess relevance or adequacy as explicitly as GRADE-CERQual.

Reporting in qualitative reviews

The enhancing transparency in reporting the synthesis of qualitative research (ENTREQ) framework was developed to facilitate transparent and comprehensive conduct and reporting of reviews synthesising qualitative studies, so that the processes involved in developing the qualitative synthesis can be clearly understood (145). The ENTREQ checklist contains 21 items covering five domains (introduction, methodology, literature search and selection, appraisal, and synthesis of findings). The development of ENTREQ is line with reporting guidelines for quantitative systematic reviews (e.g. the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (146), although it is acknowledged that a standardised set of procedures is less appropriate for qualitative reviews and the relevance of ENTREQ guidance depends on the focus of a specific review (145). Nevertheless, in a meta-review of 1,695 qualitative reviews, improved reporting quality of qualitative reviews following publication of the ENTREQ checklist was identified, although only 28% of included studies demonstrated uptake of ENTREQ (147).

Qualitative reviews in COS development

It is clear that in some contexts, combining multiple qualitative studies is useful to provide insights into a specific phenomenon, and can identify gaps in the existing literature where certain perspectives and experiences have not been explored. In COS development, reviewing existing qualitative studies can identify the extent to which domains from clinical trials and observational studies map to those that are important to patients, and can highlight where qualitative research is lacking (e.g. for specific conditions). A rapid review and synthesis of qualitative studies of patients' views and

experiences of type 2 diabetes identified domains relating to life impact that had not been identified in a preceding systematic review of clinical trials (148). Similarly, in rheumatology, the OMERACT COS for shoulder disorders conducted a thematic synthesis of existing qualitative studies exploring patients' experiences and identified a domain, cognitive dysfunction, that had not been identified in a previous systematic review of clinical trials (149). However, while a review of existing qualitative studies can provide insight into domains of potential importance, it is based on data generated by different researchers with varying research questions, and can be limited by interpretation bias (134).

2.5.2.2 Primary qualitative research

Primary qualitative research allows COS developers to collect data that is more relevant to the purpose of their work, facilitating exploration of potentially important domains in more depth. The most frequently used primary qualitative methods in COS development are individual interviews and focus groups.

An individual (1:1) interview generates data through an interaction between a researcher or HCP and a participant. Interviews can be structured, where an interviewer asks a pre-determined set of questions in a specific order; or semi-structured, where important aspects are covered whilst allowing flexibility for concepts of interest to be further explored. Individual qualitative interviews can provide an in-depth understanding of an individual's perspectives and experiences of a specific research topic, and allow a researcher and participant to build rapport (121). In COS development, a 1:1 interview permits a participant's journey, and therefore the range and scope of domains that are important to them, to be fully explored.

A focus group involves bringing together a group of people to discuss a specific topic, with facilitation from a researcher (150). Focus groups facilitate synergistic discussion between participants, and convergences and divergences between participants may lead to additional data in comparison to an interview (150). In COS development, focus groups allow participants to discuss, challenge, and clarify domains that are important to them, and can give additional insight into the scope of domains and the language that participants use to describe domains. However, a focus group participant has considerably less speaking time than an interview participant, which reduces the range of domains that can be fully explored (118).

There is currently no consensus on the qualitative methods that should be used when developing a COS, and advantages and disadvantages to both methods of data collection have been identified. Focus groups and 1:1 interviews can attract different

types of participant; some participants do not feel comfortable sharing personal experiences in a group setting (121: p.113), whilst others may find 1:1 settings intimidating and prefer support and validation from other participants (151). A challenge within focus groups can be dominant voices; participants who take over the conversation, leading to less participation from others and subsequently a distorted representation of perspectives (150, 152). Face-to-face focus groups can also present financial and logistical challenges, especially when participants are in different geographical locations and therefore different time zones, whereas 1:1 interviews can be scheduled at a mutually convenient time for the interviewer and participant (121: p.113). Interviews and focus groups can also be held online, and interviews can be held by telephone. Online focus groups have been shown to produce data similar in richness to face-to-face focus groups (153). Likewise, telephone interviews and face-to-face interviews can produce similarly rich data (154).

Generally, triangulation of qualitative methods is advocated as a strategy to achieve a more comprehensive understanding of a phenomena (155), and some COS developers have collected data using both interviews and focus groups (156). COS developers conducting qualitative research must consider which contributor group(s) to involve in their qualitative research, and the sample size. Qualitative research can be conducted with any key contributor group(s), but COMET and OMERACT recommend the involvement of patient participants (63, 65). OMERACT suggests that 1:1 interviews or focus groups should involve a minimum of 30 participants with experience of the relevant condition, and should include representation from at least three continents (63). COMET does not make any recommendations regarding sample size, but highlights the appropriateness of aiming for maximum diversity in a sample (65). Whilst recognising that sample size can vary depending on the research question, type of analysis, richness of data, and how much each participant contributes (157), Braun and Clarke (2013) recommend that 10-20 interviews, or 3-6 focus groups, is appropriate for a study within a larger project. For large projects, 20+ interviews, or 10+ focus groups, are recommended (121: p.48-50).

Qualitative sampling techniques

Sampling technique is an important component of qualitative research. The most common qualitative sampling techniques are presented in Table 2.4.

Table 2.4: Common qualitative sampling techniques

Sampling technique	Description
Purposive sampling	Intentional selection of participants, with the aim of gaining information about the phenomenon of interest that is as valuable as possible.
Convenience sampling	Selection of participants based on their accessibility and availability.
Snowball sampling	Identifying participants who meet the eligibility criteria and asking them to refer other potential participants who fit this criteria.
Theoretical sampling	Selection of participants, cases or contexts based on their potential to contribute to the development and refinement of theoretical concepts.

Table based on Creswell et al. (158) and Braun and Clarke (121).

COMET recommends that qualitative sampling for COS development should be purposive, with the aim of recruiting participants with a diversity of perspectives who have direct experience of the condition and/or treatment that is relevant to the COS being developed (65). OMERACT does not make any specific recommendations relating to qualitative sampling technique, but emphasises the need to capture diversity. Most COS developers have adopted purposive sampling in the qualitative studies underpinning the development of their COS (159-161).

Qualitative analysis methods

Primary qualitative research can be analysed in various ways; key analysis approaches are shown in Table 2.5. Neither COMET nor OMERACT give specific recommendations on which qualitative analysis method to employ, although many OMERACT COS developers have used thematic analysis (159-163). COMET highlights the need for COS developers to contextualise the data during interpretation, rather than simply cataloguing data extracts, and recommend involving PPI in the interpretation (65). OMERACT suggest that qualitative research conducted to inform COS development should follow reporting standards such as the consolidated criteria for reporting qualitative research (COREQ) (section 2.5.2.2) (63).

Table 2.5: Qualitative analysis methods

Approach	Description
Grounded theory	Generates a novel theory that explains a social process/action based on the perspectives of many participants.
Constant comparative method	Organises data into groups through a structured process of constantly revisiting and comparing codes and categories, and building findings into the emerging theory. This method is closely associated with grounded theory.
Phenomenology	Seeks to understand the meaning of several participants' lived experiences of a particular phenomenon, focusing on commonalities in their experiences. This can be approached with methods such as interpretative phenomenological analysis (IPA) and hermeneutic analysis.
Thematic analysis	Identifies, analyses, and reports patterns across a dataset, allowing flexibility in theoretical frameworks and research questions.
Framework analysis	Uses a structured matrix to organise and compare data by themes or concepts.
Content analysis	Systematically categorises verbal or textual data to identify patterns, frequencies, and meanings, often quantifying aspects of the data.
Narrative	Explores the life story of one or a small number of participants in detail, often involving chronological ordering of their experiences.
Ethnography	Describes and interprets shared patterns among an entire group of participants from the same culture, often through prolonged observation and immersion.
Case study	Describes and analyses in detail one or more clearly defined cases, such as a person, organisation or decision process.
Qualitative description	Comprehensively summarises a specific phenomenon in a way that closely reflects participants' perspectives.

Table based on Cresswell and Poth (164), Starks and Trinidad (165), and Braun and Clarke (121).

Reporting in primary qualitative research

Many tools are available for judging the quality of qualitative research and qualitative research reporting; in a systematic search for checklists and guidelines, Santiago-Delefosse et al. (166) identified 133 different tools. Two of the most commonly used tools are the Standards for Reporting Qualitative Research (SRQR) checklist (167) and COREQ (168). SRQR was developed in 2014 and consists of 21 items that the authors consider essential for complete, transparent reporting of qualitative research. COREQ is a 32-item checklist that was developed to facilitate explicit and comprehensive reporting of qualitative studies (in-depth interviews and focus groups). It has three domains: research team and reflexivity, study design, and data analysis and reporting. COREQ was developed through a reviewing and consolidating 22 existing tools, as well as author consensus. The use of COREQ as a rigid, generic and universally applicable tool, the use of which supposedly resulting in better reporting, has been identified by some researchers as a problem. For example, Braun and Clarke (2024) argue that use of COREQ can contribute to methodologically incongruent reporting, given the diversity of qualitative research (169). They have proposed Reflexive Thematic Analysis Reporting Guidelines, which provide guidance for reporting reflexive thematic analysis with methodological coherence and reflexive openness, and recommend that these are used instead of SRQR or COREQ for reporting reflexive thematic analysis (170).

Rigour

Rigour is also key consideration in qualitative research. There are various definitions of rigour. Lincoln and Guba's terminology and criteria are well established (171); they refer to rigour as 'trustworthiness', and propose four criteria for establishing the overall rigour of qualitative research findings. These are outlined in Table 2.6.

Table 2.6: Criteria for establishing rigour

Criteria	Description
Credibility (internal validity)	Findings are an accurate representation of the phenomenon that was studied.
Transferability (external validity/generalisability)	Detailed information is provided about the study design, context and participants, so that readers can determine whether the findings are applicable to other contexts.
Dependability (reliability)	Research processes are described in sufficient detail, so that the work could be repeated.
Confirmability (objectivity)	Findings are based on and reflective of the information gathered from the participants, rather than the researcher's interpretations or bias.

Table based on Johnson et al. (172).

2.5.2.3 Consensus methods

Consensus methods are recommended to prioritise domains identified in literature review(s) and qualitative research, for inclusion in a COS (63, 65). The most frequently used consensus methods in COS development are nominal group technique (NGT), the Delphi technique, and consensus workshops. A combination of these methods can also be used to obtain consensus on core domains and OMIs (65). Consensus methods are discussed in detail in relation to the COMFORT project in Chapter 6. There is no gold standard consensus process for COS development, although overarching COS-STAD state that consideration should be given to both health professionals' and patients' views. Additionally, these standards state that a consensus definition, and criteria for including, dropping or adding new domains, should be described *a priori*, and ambiguity in the language used when describing domains should be avoided (e.g. by presenting both lay and medical terms) (109). The three most frequently used consensus methods in COS development will now be discussed.

Nominal group technique (NGT)

NGT was developed in the 1960s and involves a structured, face-to-face meeting to obtain information from a group of experts in a focus group setting (173). NGT can be used to identify problems, develop solutions, and establish priorities. A moderator with expertise in the topic being discussed asks pre-determined questions about a specific

topic, ideas are generated confidentially, and a discussion relating to each idea takes place, followed by a ranking exercise. NGT involves at least two rounds, where ideas are further discussed and re-ranked. Rankings are aggregated statistically to determine consensus about which ideas to keep. NGT has been adapted and modified for different purposes, but the key steps (Table 2.7) have remained constant.

Table 2.7: Key steps in NGT

NGT steps
1. Introduction to the meeting and explanation of the purpose of research
2. Silent generation of ideas
3. Round-robin on flip chart (sharing ideas)
4. Group discussion of ideas on flip chart to clarify, elaborate, defend or dispute ideas
5. Ranking priorities: private ranking to produce top 10
6. Voting of top 10
7. Discussion on vote
8. Re-ranking and rating priorities individually
9. Conclusion

Table based on McMillan et al. (174).

NGT has been used in various areas of health to elicit priorities, including within COS development (175, 176). This method permits a considerable amount of information to be collected in a relatively short time period compared to the Delphi technique, which involves multiple surveys and has more potential for participant burden. The information is available as it is created when using NGT, which can be less burdensome to researchers than multiple Delphi analyses (173). In addition, NGT facilitates discussion when opinions differ; in COS development, this permits greater understanding of why certain domains are important. The involvement of geographically diverse participants is logistically challenging in NGT, however, and has obvious cost implications (177). Additionally, the small size of the group limits the number of expert contributors involved, negatively impacting on representation and potentially reducing the reliability of NGT (178).

Consensus meetings

A consensus meeting (also known as a consensus conference or workshop) typically involves a face-to-face meeting with a panel of experts, although online consensus meetings have also been adopted (179). The precise structure of a consensus meeting can vary, but usually involves presentation of relevant evidence, followed by a discussion among meeting attendees. During COS development, a consensus meeting often follows other methods of consensus, particularly when there is an excess number of domains or instruments indicated for inclusion and further discussion and prioritisation is warranted. COS developers planning a consensus meeting must consider whether to host a single meeting that involves all contributors, which may introduce power imbalance and influence voting, or separate meetings. A consensus meeting can involve anonymous votes by all attendees using pre-defined consensus criteria. For example, OMERACT historically specified that COS development must involve a consensus workshop, and a consensus threshold of ≥70% for each core domain is applied (63), although this process recently changed to an online 'module' whereby evidence for each domain is presented on a pre-recorded video, and participants anonymously vote whether or not they agree that the COS has been developed according to robust methodology (180-182).

Delphi technique

The Delphi technique was developed by the RAND Corporation in 1948 and is defined as "a method for achieving consensual agreement among expert panellists through repeated iterations of anonymised opinions and of proposed compromise statements from the group moderator" (183: p.49). Consensus is achieved from the opinions and feedback of participants who complete multiple rounds of surveys. In a classical Delphi, a survey is presented to a panel of experts, their anonymous responses are analysed, and a second survey is designed based on these responses. A key part of this process involves repeatedly giving each panel member feedback on other panel members' responses, and asking them to consider this feedback and review their initial responses (184). Key methodological elements of the Delphi technique have remained consistent since its inception, although modifications to the classical technique have been adopted. Variations are termed "modified Delphi technique", although there is no standard definition of a modified Delphi; different researchers apply different modifications. Typically, the first round of a modified Delphi study includes pre-determined items, as opposed to open-ended questions.

The Delphi technique has been used extensively in health research to achieve consensus and is frequently used to prioritise domains for a COS (185). In this context,

a modified Delphi technique is employed, initially presenting a list of domains identified from literature reviews and qualitative research. Participants prioritise these, but also have the option to generate additional domains in the initial Delphi round. New domain suggestions are then discussed and potentially added to the second round of the survey (63, 65).

The Delphi technique enables large numbers of participants from geographically diverse locations, allowing for multi-contributor and global representation at relatively low costs (174). The Delphi technique avoids the influence of overly dominant participants, as Delphi responses are anonymous and have equal weight (186). Additionally, Delphi participants can take part remotely and at a time of their choice, and take time to think about their responses (184). The time taken to understand the importance of a domain from other participants' perspectives can influence a change in Delphi domain ratings (187), particularly for HCPs, who have been shown to change their ratings as a result of patients' ratings (187). Key considerations for the Delphi technique as part of COS development include sample size, feedback, consensus definitions, scale, and attrition (and subsequently recruitment and retention methods).

Sample size

There is no standardised method for determining the sample size for a Delphi study. OMERACT recommends initially including a minimum of 100 participants from each key contributor group (63). Indeed, higher numbers of Delphi participants has been shown to increase the reliability of data in research outside of COS development (188), although this increase can result in a lesser degree of consensus due to lack of face-to-face interactions (189).

Feedback

A COS Delphi exercise typically involves multiple contributor groups, including patients, HCPs, and researchers. Consideration must therefore be given to how responses are fed back to participants after each round; these can be pooled or presented separately to each contributor group. In RCTs nested within the development of three COS for various cancers, provision of separate group responses increased consensus between patients and HCPs (190). In concurrence, a qualitative study with participants who had completed COS Delphi surveys indicated that attempting to understand the perspectives of other participants was a common reason for revising scores, especially among HCPs (187). In the development of a COS for prostate cancer, Delphi

participants from two contributor groups (HCPs and patients) were randomised to receive feedback from peers only, multiple contributor groups separately, or multiple contributor groups combined. In contrast to previous evidence, there was no evidence that different feedback strategies influenced consensus, in terms of the number of domains retained or reduction in variability of opinion (191), although the authors suggested that this may have been explained by the high level of agreement in Round 1 and called for further methodological work in this area.

There is also a lack of evidence on the most effective way to present feedback to Delphi participants. Fish et al. (192) conducted a simulated two-round Delphi study with eight participants, using think-aloud interviews. Feedback was presented as a single summary statistic (median score), then as two different graphs (histogram, pie chart). Participants' understanding of a median score was limited, whereas the graphs were deemed more understandable. The histogram was preferred overall. Further supporting these findings, Hall et al. indicated that 76% of participants were satisfied to some degree with interpreting the graphical display of Delphi feedback (193).

Consensus threshold

The threshold for consensus is another important consideration for COS developers conducting Delphi studies. The choice of consensus threshold can affect the overall Delphi study findings (194). A threshold that is too low can result in an extensive list of domains, which goes against the 'minimal' concept of a COS. However, if the consensus threshold is too high, some key domains might be excluded. There is extensive variability in Delphi consensus criteria definitions, and the choice of criteria for defining consensus is rarely justified (195). Most commonly, descriptive statistics (e.g. percentage agreements or measures of central tendency) are employed (196, 197). When percentage agreement is used, the threshold is usually 70% or higher (196, 197). A systematic review by Diamond et al. (197) assessed the consensus threshold in a random sample of 100 Delphi studies and identified that the most common definition of consensus was percentage agreement (25% of studies), with 75% being the median threshold to define consensus. Other COS developers have reported consensus thresholds between 60% to 90% (198-200).

As previously outlined, COS-STAD state that a consensus definition should be described a priori, but there are no standards relating to consensus scoring for domains to be included or excluded from a COS. OMERACT recommends a consensus threshold of ≥70% from both groups (patients, other key contributors) for domains rated as critically important (scores 7-9, as discussed below) to include in a

COS. However, whilst OMERACT recommend that if ≥70% of both groups rate a domain as unimportant, it should be excluded, some COS developers have implemented a threshold of 70% of participants scoring domains between 7-9 *and* ≤15% scoring domains between 1-3 for both key contributor groups (patients and HCPs), or ≥ 90% scoring the domain as 7-9 from any single group (201, 202). The argument here is that most participants rate the domain highly for inclusion in the COS, with only a small minority considering it to be of little or no importance (67).

Rating scale

The type of rating scale used in a Delphi study can influence consensus. A systematic review by Barrington et al. (203) identified that a 1-9 scoring system was most commonly used. Indeed, a 9-point scale is used in DelphiManager (a programme designed by COMET to facilitate e-Delphi surveys) (204). This scoring system also in line with Grading of Recommendations Assessment, Development and Evaluation (GRADE), which recommends a 9-point scale to assess the importance of evidence (205). Similarly, OMERACT states that Delphi participants should rate the importance of each domain on a scale of 1-9, where 1-3 represents domains that are unimportant, 4-6 represents domains that are 'important but not critical' and 7-9 represents 'critically important' domains for inclusion in a COS (63).

In the first round of a Delphi study for an incontinence-associated dermatitis COS, a 9-point scale resulted in more domains being rated as important compared to a 3-point rating scale (206). In contrast, the Pelvic Girdle Pain Core Outcome Set (PGP-COS) study compared a 5-point to a 9-point rating scale and concluded that the 5-point rating scale resulted in twice as many domains reaching consensus (207). Biggane et al. (185) explored the experiences of 24 patients and HCPs who had participated in COS Delphi surveys, and identified that perspectives of a 9-point scale differed; some patient participants indicated the need for additional support and guidance for the 9-point scale and felt they would prefer a 5-point scale, whereas others were satisfied. Alwin and Koswick (208) found that 9-point scales demonstrate maximal reliability.

Domain order

The order of domains presented in a Delphi study can also affect consensus. In the development of a COS for oesophageal surgery, Brooks et al. (209) randomised patients and HCPs to receive a Delphi survey where the patient-reported outcomes section was presented first or last. Findings demonstrated that patients inflated the importance of patient-reported outcomes when rating them last, whereas HCPs inflated

the importance of clinical items when rating them last, suggesting that question order in Delphi surveys should be randomised.

Piloting

Piloting a survey can highlight ambiguity in the wording of questions, allowing researchers to improve clarity prior to survey launch (210). This is particularly pertinent when conducting a Delphi survey to develop a COS, as domains may be perceived as redundant (e.g. overlapping with other domains or not comprehensive enough), and domain definitions may include medical jargon and have the potential to be interpreted in different ways depending on the participant. The duration of survey completion, flow, navigation, ease of use, and accessibility of the survey can also be tested during a pilot, including the rating scale and how this scale is explained to participants, and how results from the previous round are presented. COMET and OMERACT both recommend piloting during COS development (63, 65), although Barrington et al. (203) indicated that only approximately one third of COS development studies report piloting their Delphi surveys.

Attrition

Attrition refers to loss of participants over the course of a study. High attrition rates can impact the validity and reliability of Delphi findings if the priorities of participants who drop out differ systematically from the priorities of those who complete the study (65), leading to overestimation of consensus degree in a final COS. The number of rounds, number of domains, recruitment and retention methods, and language used in a Delphi study can impact on attrition.

Number of rounds

There is no agreement on how many rounds a Delphi should have; more rounds can increase consensus as participants are provided with more feedback, but the potential for participant burden also increases (184, 211). COMET states that a minimum of two rounds should take place. In comparison, OMERACT recommends a minimum of three rounds of Delphi surveys; a fourth round, where participants can select and rank a maximum of ten most important domains and provide comments on why a domain should be included or excluded, is recommended when multiple domains have reached consensus for inclusion in a COS at the end of the third round (63).

A recent trial randomised participants to a standard, three-round Delphi or a real-time Delphi (212). In the real-time Delphi, participants initially rated a domain without seeing other participants' responses. The survey page was then refreshed and response graphs from all contributor groups were presented. Participants could subsequently change their rating before continuing to rate further domains. Seventy-three per cent of the domains were rated the same across both survey arms, with greater convergence in the real-time Delphi. The benefits of shorter completion time and reduction of dropout in real-time Delphi were highlighted.

Consideration should also be given to the duration of and time between Delphi rounds. A long delay between rounds may cause loss of interest among participants, leading to dropout (193), whilst rounds in close succession may feel burdensome. The duration of and time between rounds is not always reported in Delphi studies (213), although COMET suggests that each Delphi round typically remains open for 2-3 weeks, with a period of 2-3 weeks for analysis before the next round opens (65).

Number of domains

A key consideration in a Delphi study is the number of domains to include; too many domains can be burdensome for participants, affecting the length of the surveys, which can affect recruitment and retention (213). An association between attrition and a higher number of Delphi items has also been identified (213). OMERACT recommends including less than 70 items in the first round of a Delphi survey (63). In contrast, too few domains may lead to key domains being excluded.

The granularity of each domain, referring to how broad or narrow a domain definition is, must also be considered during the Delphi process (63, 65). Domains may be defined as broad aspects of health, whilst subdomains are defined as aspects of domains that are more specific and granular than domains (214). OMERACT use the terms 'target domains' and 'broad domains', with target domains referring to the specific concept that will be measured in a clinical trial and as part of a COS, and broad domains referring to a less specific concept (215). For example, pain can be considered a broad domain, whereas pain intensity and pain when weightbearing can be considered target domains. Presenting only broad domains in a Delphi can reduce the number of domains included, but these may not be fully understood by participants, highlighting the need to compromise between domains being "specific enough to be meaningful, but broad enough to be practical" (214: p.3). Kottner et al. (2024) developed guidance

relating to domain granularity (214). The authors recognise that the appropriate level of granularity of domains will differ depending on the specific COS being developed, negating the need for overly prescriptive guidance. They recommend that the level of domain granularity should be explicitly planned, and that clear rationale about the level of granularity should be given by COS developers. One key point in this guidance refers to avoiding overlapping levels of granularity; for example, presenting a broad domain such as ‘clinical signs’ in addition to target domains covering different clinical signs. This overlap can confuse participants.

Recruitment and retention methods

Recruitment and retention methods have not been well reported in Delphi studies for COS development. To further understanding in this area, Hall et al. (2018) (193), evaluated the effectiveness of recruitment and retention methods used in the Core Outcome Measures in Tinnitus International Delphi (COMiT’ID) study, using a 20-item online questionnaire, with free text comments. The authors found that personalised email invitations were the most effective method of recruitment in terms of numbers of consented participants. Other Delphi studies with high retention rates also sent personalised emails (213, 216). COMET recommends that personalised emails sent from distinguished researchers in the field can enhance retention (65). Acknowledging participants who complete all Delphi rounds in the publication arising from the study is another strategy (65, 193). Extending the closure date of each Delphi round is also recommended to increase retention (65, 193).

Other methods of recruitment and retention include selecting motivated participants by inviting them to participate prior to launching the first round of the Delphi (184), and creating a brand identity with dedicated websites and social media pages for information and promotion, to maximise engagement with participants (216). For example, the COMiT’ID study also produced an animated video featuring an overview of the project and a call for participants, which was showcased at relevant conferences and public engagement events, and kept participants updated about progress through emails and social media (193). These updates, in addition to thanking participants for their participation throughout the process, aimed to build rapport with participants, and included information on when to expect the next round. The COMiT’ID study also involved sending reminder emails emphasising the importance of completing the full Delphi process for the results to be meaningful (193). In contrast, some participants may feel harassed by emails (217).

A systematic review by Barrington et al. (203) explored how recruitment methods influenced patient participation in COS development. The authors suggested that studies where patients were recruited from clinics appeared to have higher Delphi Round 2 response rates.

PPI input

Structured PPI input into Delphi studies can address the challenge of recruiting and retaining patients, considering that the process can be time consuming and repetitive, with the concept of a COS and its long-term benefits potentially difficult to understand. Methods and recommendations to involve PPI have included recruitment to the project steering group, reviewing and reducing the list of candidate domains, naming domains, creating a plain language description for each domain, and piloting the Delphi survey for accessibility (65, 203). In the development of a COS for pressure ulcer prevention trials, a pre-Delphi workshop with a group of people with or at elevated risk of a pressure ulcer and informal caregivers discussed their understanding of the domains identified in a previous scoping review. The COS developers found that the group had some issues understanding the concept of an outcome, and in distinguishing between outcomes and interventions. The group was subsequently involved in creating and reviewing the introductory text and domain descriptions for the Delphi survey (218).

There is a paucity of research regarding PPI input in COS development and its benefits (95). Smith et al. (219) involved patients with tinnitus in creating a list of candidate domains for a Delphi study and reflected on the benefits of this. Two patients co-designed the survey, and an independent patient review panel was consulted. The authors concluded that involving patients reduced the list of domains included in the survey and made domain names and associated descriptions clearer. The need for further methodological guidance on how to involve PPI in COS development studies has been identified (156). Young and Bagley (95) conducted three interactive workshops with patients, HCPs, and researchers to discuss how to make COS development more meaningful and accessible for patients, ensuring they have a genuine say in the development process. The suggested benefits of involving PPI included helping recruitment and retention by improving the accessibility of the study. Barrington et al. (203) found no association between PPI and patient participation rates, although the minimal reporting of PPI was acknowledged as a limiting factor.

As discussed in section 2.3, COMET has developed a PPI toolkit for COS developers, which supports transparent and meaningful PPI by provided a structured framework to plan, record and reflect on PPI activities (101). The toolkit encourages COS developers

to consider how PPI contributors were involved and supported. Additionally, COMET has developed lay resources aimed at patients and the public to support PPI in COS development.

Survey accessibility and readability

Inaccessible surveys (e.g. surveys that are difficult to complete, confusing or ambiguous questions, dense text, certain fonts and colours) may deter potential participants (220). COS developers must also consider language use. Plain language, rather than medical jargon, can encourage broader participation and may reduce attrition among participants (193); many adults have limited health literacy levels and struggle with health information, particularly if it includes both text and numbers (221). Text that is too difficult for participants to read and understand can potentially impact on recruitment to research studies (222). It is important, therefore, that participants in a Delphi for COS development understand the aim of the study, its relevance, and what they are being asked to do, to avoid demotivation and dropout. Conversely, if participants don't fully understand what they are being asked to do, but complete the survey anyway, the results will lack validity. Biggane et al. (223) conducted semi-structured interviews with patients and HCPs who participated in COS studies that had involved Delphi surveys, to explore their experiences and understanding of COS. Participants who had taken part in multiple COS generally understood the purpose, but some participants who were taking part for the first time indicated that they did not fully understand. This highlights the importance of making sure patients are involved in the development of Delphi surveys, to ensure accessibility.

2.6 Conceptual frameworks

COS developers categorise core domains into a taxonomy, which is a conceptual classification framework. A well-established taxonomy is the WHO International Classification of Functioning, Disability and Health (ICF) (224), which categorises domains as follows:

- body functions and structures, and impairments thereof (functioning at the level of the body);
- activities (functioning at the level of the individual) and the activity limitations experienced;
- participation or involvement of people in all areas of life, and the participation restrictions experienced (functioning of a person as a member of society);

- environmental factors which affect these experiences (and whether these factors are facilitators or barriers).

Although the ICF framework has been used by COS developers to categorise domains (225-227), it does not consider domains related to resource use, adverse events or death, which are often pertinent to measure in clinical trials. To address this issue, conceptual frameworks specifically for COS development have been proposed. OMERACT has developed the OMERACT Filter (Figure 2.1) to guide the development of COS in rheumatology (228). The OMERACT Filter draws on the key aspects of the ICF in four core areas: pathophysiological manifestations (signs, symptoms, biomarkers), life impact, death, and societal/resource use. This framework proposes that at least one domain should be specified for each of the four core areas, although societal/resource use is recommended rather than mandatory. Adverse events and important contextual factors (variables that need to be recognised and measured to understand the results of a study) are also considered.

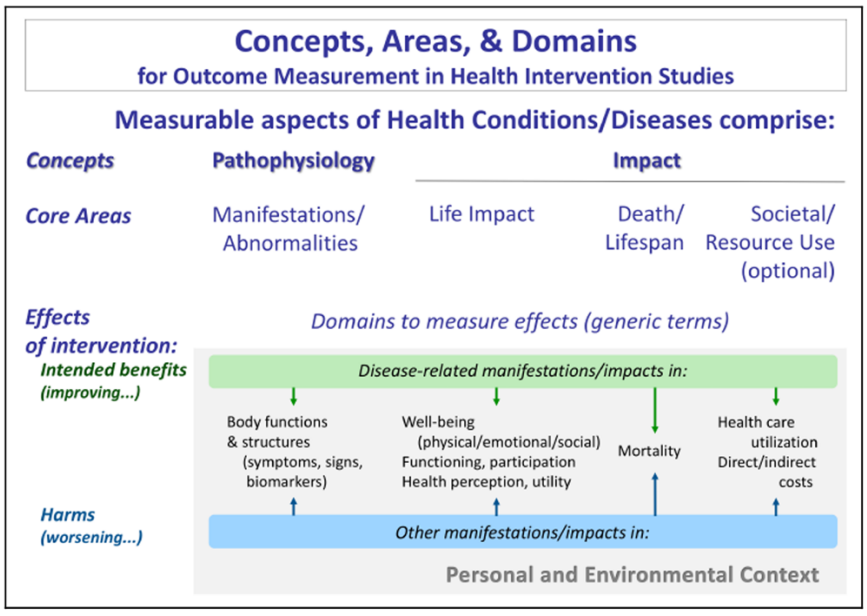


Figure 2.1 OMERACT 2.1 Filter

Figure reproduced without modification from Figure 2 in Boers et al. (228) under the terms of the Creative Commons Attribution 4.0 International License (229).

Similarly, COMET has developed a taxonomy, consistent with the core areas within the OMERACT Filter, but more detailed (230). The 38-item classification system was developed in 2018, but there is limited published work demonstrating its use (231). The

HOME initiative has also published a methodological framework for COS development, primarily mirroring the OMERACT framework, but with an additional roadmap for implementation (91, 232).

OMERACT has produced a conceptual model, known as the OMERACT onion, to structure domains in a core domain set. Historically, the three layers of the OMERACT onion included core/mandatory domains, for measurement in all studies in a specific area of health (the inner core), important but optional domains (the middle circle), and a research agenda, representing domains that are not yet well-defined and need further research (the outer layer). An example of a the OMERACT onion for the PsA core domain set is shown in Figure 2.2. More recently, OMERACT have amended the onion, with the middle layer now representing circumstance-specific domains, and the outer layer representing domains for future consideration (Figure 2.3) (233).

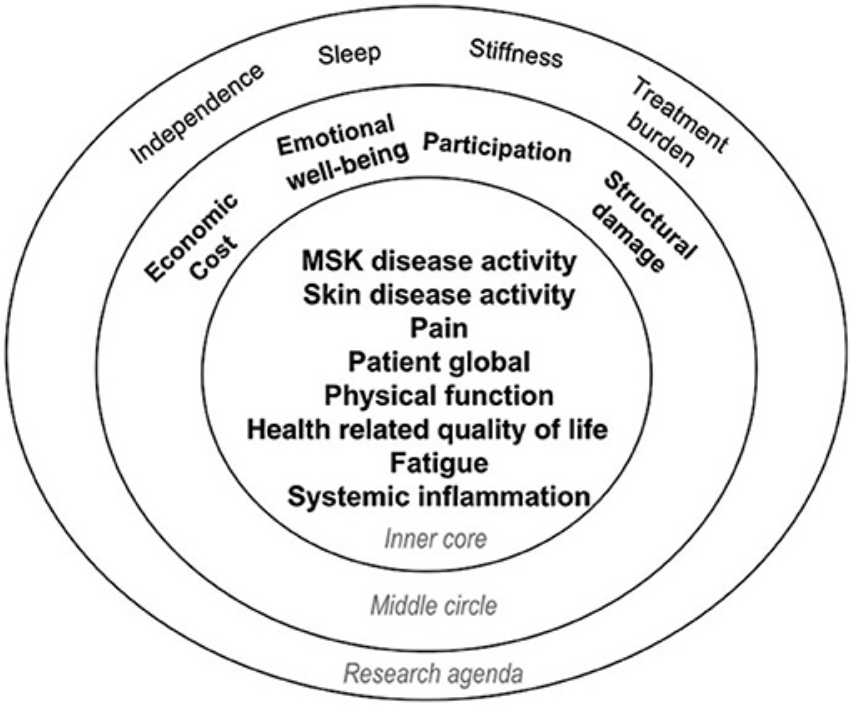


Figure 2.2: OMERACT onion for the PsA core domain set.

Figure reproduced without modification from Figure 2 in Orbai et al. (233) (https://doi.org/10.1186/1748-5908-6-42) under the terms of the Creative Commons Attribution 4.0 International License (229).

Domains for future consideration	Important to only one collaborator group or domains needing more research.
Circumstance-dependent core domains	Should be measured in all clinical trials when specific circumstances are present.
Core domains	Should be measured in all clinical trials or longitudinal observational studies, covering the OMERACT core areas.

Figure 2.3: Updated OMERACT onion.

Figure reproduced without modification and with permission from OMERACT (234).

2.7 Instrument selection methods

Once a core set of domains (*what* to measure) has been finalised, COS developers recommend an OMI for each domain (*how* to measure each domain). Failure to standardise OMIs for core domains continues to limit evidence synthesis (63, 65). Existing relevant OMIs can be identified through systematic or scoping reviews, and through OMI databases, such as the COSMIN database of systematic reviews of OMIs and the Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID). Each database will now be discussed.

2.7.1 COSMIN database of systematic reviews of OMIs

The COSMIN initiative aims to improve the selection of OMIs for research and clinical practice through the development of methodology and tools for selecting the most appropriate OMI. To facilitate these aims, COSMIN systematically collects systematic reviews of OMIs in a searchable database, which is free to access (235). Systematic reviews are included in this database when they fulfil the following criteria:

- A search in at least one electronic database was performed;
- The review aims to identify all OMIs of interest and to summarise the evidence on their measurement properties;
- The construct of interest of the review should be aspects of health status, defined as biological and physiological processes, symptoms, physical

functioning, social/psychological functioning, general health perceptions, or health-related quality of life (236);

- The population of included studies should contain humans (patients or general population);
- The instruments of interest should be OMIs (defined as instruments that can be/are applied in longitudinal studies to monitor changes in health over time;
- The review should evaluate and report on at least one or more measurement properties of the included instruments.

A limitation of the COSMIN database is that it depends on what systematic reviews have been done. For some OMIs, populations or domains, there may not be any relevant reviews. Additionally, as a result of resource limitations, the database has not been updated since March 2024.

2.7.2 PROQOLID

The COSMIN database of systematic reviews is connected to the PROQOLID, which was also developed to facilitate the selection process of PROMs in clinical research (237, 238). Each systematic review entry within the COSMIN database lists the OMIs covered within the review, with a link to the PROQOLID database entry if applicable. The PROQOLID database contains a record of over 8,000 individual OMIs, including development and validation information. The database helps researchers to find OMIs, and to understand practical details such as how to obtain the OMI, locating accompanying manuals, translations, and copyright/licensing information. However, whilst basic content is free, many features of PROQOLID require a paid subscription.

2.7.3 OMI selection

Once candidate OMIs have been identified, selection for core domains can draw on several methods, including conceptual alignment of OMIs to core domains, a feasibility assessment, a systematic review of the psychometric properties of OMIs, and consensus processes. Selection of methods often depends on the timeframe and resources available to COS developers. A guideline recommending methods for selecting OMIS for a COS was published in 2016 (239). This was a collaboration between COSMIN and COMET, and was based on the findings of a literature review of existing studies providing guidance on OMI selection, and on expert opinion. Minor differences exist between COSMIN and OMERACT methods for OMI selection. These methods will now be discussed.

2.7.3.1 Conceptual alignment

OMERACT recommends that conceptual alignment is the first step of assessing an OMI (240). This process addresses whether an OMI appears to be a good match for a core domain, in terms of content and face validity (Table 2.8). Similarly, COSMIN highlights that content validity should be assessed first during an assessment of OMI measurement properties, and if it is poor or unknown, the OMI should not be considered further (239).

2.7.3.2 Feasibility assessment

OMERACT recommends that the feasibility of an OMI is assessed following conceptual alignment, to determine if its use is practical (240). Consideration is given to cost, burden, access, format and mode of administration. Input from researchers, HCPs, and patients at this stage is recommended to confirm feasibility. Comparatively, COSMIN recommends that feasibility aspects of an OMI should be assessed as the second step of a quality assessment, following an evaluation of methodological quality (239).

2.7.3.3 Quality assessment

A unique feature of the OMERACT process of developing a COS is that an OMI cannot progress to the quality assessment stage if conceptual alignment is poor, or if it is deemed unfeasible (240). In comparison, quality assessment is a necessary and earlier step of OMI selection in COSMIN recommendations (239). Both initiatives state that evidence determining the measurement properties of each candidate OMI should be ascertained; this includes an evaluation of the methodological quality of existing studies assessing candidate OMIs and an evaluation of the quality of the OMI itself. Measurement properties can be assessed using the COSMIN taxonomy (Table 2.8). The overall quality of each candidate OMI is constructed based on a best-evidence synthesis, incorporating the number and methodological quality of studies assessing the OMI, and the consistency of measurement property results.

Table 2.8: COSMIN taxonomy

Measurement property	**Definition according to the COSMIN taxonomy**
Content validity (including face validity)	The degree to which the content of a measurement instrument is an adequate reflection of the construct to be measured.
Reliability	The degree to which the measurement is free from measurement error.
Responsiveness	The ability of a measurement instrument to detect change over time in the construct to be measured.
Internal consistency	The degree of interrelatedness among the items.
Structural validity	The degree to which the scores of a measurement instrument are an adequate reflection of the dimensionality of the construct to be measured
Measurement error	The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured.
Hypotheses testing	The degree to which the scores of a measurement instrument are consistent with hypotheses based on the assumption that the measurement instrument validly measures the construct to be measured.
Criterion validity	The degree to which the scores of a measurement instrument are an adequate reflection of a "gold standard."
Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted measurement instrument is an adequate reflection of the performance of the items of the original version of the measurement instrument.

Table based on Prinsen et al. (239).

2.7.3.4 Consensus processes for OMI selection

COSMIN does not make any recommendations relating to achieving consensus on what OMIs to use for a COS. In comparison, OMIs selected during an OMERACT COS development process are subject to a consensus meeting and subsequent survey involving key contributors, again with a consensus threshold of ≥70%. Each core domain is typically considered in turn. For example, in the development of a COS for PsA, candidate OMIs (the Health Assessment Questionnaire-Disability Index and SF-36 Short Form-Physical Function) for the core domain of physical function underwent conceptual alignment, a feasibility assessment, and a quality assessment, followed by a consensus meeting and subsequent online poll. Further work to standardise OMIs for other PsA core domains (enthesitis, fatigue, structural damage) is ongoing (241).

2.8 COS uptake

The theoretical benefits of a COS (facilitation of meta-analyses, reduced outcome reporting bias, relevance/transferability of findings to clinical practice) were discussed in Chapter 1. For these benefits to manifest, the COS must be used by researchers. Lack of uptake of a COS after its development leads to research waste, an issue that COS development intends to overcome. COS uptake varies widely across different areas of health. A systematic review assessing the uptake of COS in 24 RCTs across five different disease categories identified that uptake rates varied from 0-82% (242). In rheumatology specifically, assessments of uptake of the COS for RA have demonstrated that over 80% of RA trials are using the COS (243, 244). In contrast, an assessment of the uptake of the COS for hip and knee OA trials found decreasing adoption of the COS over a 20-year period (245). OMERACT and COMET are continuing to attempt to improve COS uptake (246, 247). Key barriers to uptake have been identified through qualitative interviews with trialists and include lack of key contributor involvement in COS development and lack of awareness of the COS, highlighting the importance of a COS dissemination strategy (248).

2.9 The COMFORT project

COMFORT involves a mix of methods, using qualitative and quantitative research to identify and prioritise domains for a COS. These methods will be detailed throughout Chapters 3-6 of this thesis. Given the involvement of RMDs, the project is guided by OMERACT methodology, including the OMERACT conceptual framework. However, it also draws upon COMET and COSMIN resources, and follows COS-STAD.

2.9.1 COMFORT scope

An overview of the prevalence and characteristics of foot and ankle disorders in a range of RMDs, treatments for these disorders, and findings from a scoping review of domains measured in clinical trials and observational studies investigating the effectiveness of interventions for these conditions, was presented in Chapter 1.

The scope of the COS developed in the COMFORT project was determined based on this existing literature (e.g. by considering which RMDs have evidence of foot and ankle involvement) and on clinical experience (e.g. by considering the conditions in which foot and ankle outcomes were likely to be similar enough for a core set across multiple RMDs to be feasible). The proposed scope was developed with expert input and advice from a project advisory group and the wider OMERACT Foot and Ankle Working Group, including PPI contributors. Experiences of managing and conducting research with patients with foot and ankle disorders in different RMDs, and of living with different RMDs, were taken into consideration when making a decision about the scope. The proposed COS is applicable to measuring the efficacy or effectiveness of pharmacological, conservative and surgical interventions in RCTs, controlled clinical trials, and observational studies involving patients with RMDs and foot and ankle disorders. In the context of the COMFORT project, RMDs encompass those discussed in Chapter 1: inflammatory arthritis, OA, spondyloarthropathies, connective tissue diseases, crystal arthropathies, and musculoskeletal disorders affecting the foot and ankle in the absence of systemic disease. Based on how under-represented foot and ankle research is, and based on clinical experience that HCPs and patients rarely distinguish between these two areas of the body, the foot and ankle were considered together within this project. The inclusion criteria for the COS scope was under review throughout this project: findings from each phase were discussed with the PAG to determine whether any of the RMDs initially included needed excluding (e.g. if domains of importance in one RMD were potentially too different to those of importance in others).

The following conditions were excluded, based on clinical and research experience: acute foot and ankle injuries (e.g. fracture, rupture, sprain), diabetes, peripheral neuropathy, primary neurological conditions (e.g. multiple sclerosis). The outcomes measured in trials involving these conditions, and indeed the outcomes important to patients and with these conditions, are likely to be different to those specified in the inclusion criteria, and require their own COS.

The development of this work commenced in 2018 with the formation of the OMERACT Foot and Ankle Working Group. At this point, no other COS for foot and ankle disorders

in RMDs were registered in the COMET database. Searches for “foot” and “ankle” revealed planned COS for diabetic neuropathy, diabetic foot ulceration, clubfoot, and ankle fractures. Whilst many other COS in rheumatology exist, including those for RMDs such as RA (249), AS (250), and PsA (233), no specific attention has been given to measuring the effectiveness of treatments for foot and ankle disorders in these conditions.

2.9.2 COMFORT project advisory group (PAG)

A multidisciplinary PAG oversaw the COMFORT project. The PAG included the candidate, her two University of Leeds supervisors (HJS, ACR, both podiatrists), an external supervisor/qualitative expert (CAF) two PPI contributors (PR, CH; PR was also a PhD supervisor), two rheumatologists (PSH, PGC), an epidemiologist (MTH), two podiatrists (HBM, JBA), and two physiotherapists (TOS, YMG). The candidate had a leading role within the PAG, steering all aspects of the work. PAG members included offered advice on all aspects of research design, and monitored the progress of the work. The PAG met approximately every three months for the duration of the project, and contact between meetings was maintained through emails. The PAG overlapped with a wider OMERACT Foot and Ankle Working Group, who provided additional input into OMERACT-specific decisions. This included OMERACT methodologists (PT, BS).

2.9.3 COMFORT PPI

As outlined above, two PPI contributors were members of the COMFORT PAG, one of whom was a PhD supervisor for the candidate. Both PPI contributors had experienced foot and ankle involvement in RMDs, and had previously been involved in developing COS for RMDs. To ensure all patient-facing materials in the qualitative interviews and Delphi study were accessible, additional PPI input was sought from the NIHR Leeds Biomedical Research Centre (BRC) PPI group, who were unfamiliar with COS. The UK Standards for Public Involvement were used as guidance for meaningful PPI (251). All PPI contributors were provided with a description of their expected role and time commitment. Their training needs were assessed and they received ongoing generalised PPI training and support through dedicated OMERACT and NIHR Leeds BRC leadership teams. OMERACT PPI training also included COS-specific training, which was supported by lay summaries (written and video-recorded) of the COMFORT project. Local PPI contributors were invited to contribute to the Delphi study, where lacking knowledge about COS development was an advantage. To ensure their

perspectives would more closely reflect those of Delphi patient participants, they did not receive any COS or COMFORT-specific training.

All PPI contributors were provided with feedback detailing the impact of their contributions and how aspects of the project changed as a result of their input. PPI contributors were reimbursed for their input according to NIHR PPI guidelines (252). Both of the formal PPI groups involved in this project are actively striving to increase equality, diversity and inclusion (EDI) amongst their members, by seeking and encouraging participation from PPI contributors a wide range of disease, cultural, and geographic backgrounds (102, 253). The PPI activities employed throughout the COMFORT project were guided by the COMET and OMERACT PPI toolkits and will be discussed in further detail throughout Chapters 3-6. The candidate also reflected on PPI input during each phase of the project a using PPI reflection log (101). This encouraged ongoing reflexivity and helped to identify any challenges to effective participation.

2.9.4 Key contributor group engagement

A variety of key contributor groups were involved in developing COMFORT as participants. Justification for the inclusion of each group and details of their participation throughout this work is outlined in Table 2.9.

Table 2.9: Overview of COMFORT contributor groups

Key contributor group	**Justification for inclusion**
Patients	Patients with RMDs have valuable insights into the experience of living with foot and ankle disorders.
HCPs	HCPs from different disciplines (e.g. medicine, podiatry, physiotherapy, prosthetics and orthotics, occupational therapy, orthopaedic surgery) have insight into the manifestation of symptoms, prognosis and management of patients.
Researchers	Researchers with expertise in foot and ankle disorders in RMDs have insight into the feasibility of measuring outcomes in the context of clinical research.

2.10 Conclusion

This chapter provided a comprehensive overview of COS development methods, including literature reviews, qualitative research, consensus methods, and instrument selection methods. It then introduced the COMFORT project, including its scope, PPI and the key contributor groups involved. The next chapter details the process of identifying domains, from existing qualitative literature, that should be considered for inclusion in the COMFORT core domain set.

Chapter 3 Systematic review and thematic synthesis of qualitative studies

3.1 Introduction

As discussed in Chapter 2, qualitative research with patients is a recommended step when developing a COS, given its potential for identifying meaningful domains that researchers may not have anticipated (63, 65). No previous studies have synthesised the findings of existing qualitative studies exploring the lived experiences of people with foot and ankle disorders in RMDs. The following chapter reports a systematic review and thematic synthesis of existing qualitative research involving people with RMDs who have experienced foot and ankle disorders. Findings inform the design of a primary qualitative study (Chapter 5) and contribute to a Delphi consensus study (Chapter 6). Findings have been published as:

> **Chapman LS**, Flurey CA, Redmond AC, Richards P, Hofstetter C, Tapster B, Emmel J, Helliwell PS, Menz HB, Hannan MT, Shea B, Siddle HJ. Living with foot and ankle disorders in rheumatic and musculoskeletal diseases: A systematic review of qualitative studies to inform the work of the OMERACT Foot and Ankle Working Group. Semin Arthritis Rheum. 2023 Aug;61:152212. https://doi.org/10.1016/j.semarthrit.2023.152212

3.2 Aim

The overall aim of this review was to identify, from existing qualitative studies, domains that are important to patients with RMDs who have experienced foot and/or ankle disorders, and should therefore be considered for inclusion in a COS. The objectives were to:

- Establish participants' self-reported foot and ankle symptoms, and any broader symptoms occurring as a result of foot and ankle disorders;
- Explore experiences of these symptoms and their impact on participants' lives;
- Identify any gaps in existing qualitative studies relating to participant demographics and characteristics;
- Assess the quality of existing qualitative research in this area, and confidence in review findings.

3.3 Methods

The systematic review was registered on PROSPERO (ID CRD42021299523). Reporting of the systematic review follows the ENTREQ framework (145), which was discussed in Chapter 2 (section 2.5.2.1). ENTREQ was selected because it has been shown to improve the quality of qualitative systematic reviews (147).

3.3.1 Search strategy and study selection

A literature search was performed using Ovid MEDLINE, Ovid Embase, Cumulative Index to Nursing and Allied Health Literature (CINAHL), PROSPERO, Ovid PsycINFO, Cochrane Database of Systematic Reviews, and Cochrane Central Register of Controlled Trials (CENTRAL) from inception to March 2022. The search strategy was conducted with guidance from two experienced information scientists, with additional input from the PAG, including PPI contributors. The full search strategy is included in Chapman et al. (254: supplementary table 1). Systematic reviews were also included in the initial database searches and screened for potentially relevant articles prior to their exclusion. A manual search of reference lists and related citations of relevant articles was also conducted, in addition to forward citation tracking using Scopus. Finally, the wider OMERACT Foot and Ankle Working Group were asked if they were aware of any qualitative literature that had not been included, to minimise the likelihood of overlooking any additional relevant material.

Studies in which the authors had used qualitative interviewing or focus group methods to explore the perceptions and experiences of adults (≥ 18 years) living with foot and ankle disorders in RMDs were eligible for inclusion. RMD inclusion and exclusion criteria aligned to the scope of the COS, as presented in Chapter 2 (section 2.9.1).

Full articles in the English language that were published in peer-reviewed journals were eligible for inclusion. Conference abstracts were excluded, as were qualitative surveys, as the intention of this study was to conduct an in-depth analysis. Mixed methods studies reporting quantitative and qualitative data were eligible for inclusion only if the qualitative interview/focus group data could be extracted separately. Studies including both eligible and ineligible participants (e.g. HCPs, or patients with diagnoses of other diseases) were included only if the data on eligible participants could be separated from the data on ineligible participants.

Independent screening was undertaken by two reviewers (LSC and HJS) under the supervision of an experienced qualitative methodologist (CAF). Studies retrieved from

the database searches were imported into EndNote (EndNote X9.3.3, Clarivate, 2021). After removing duplicates, titles were reviewed, irrelevant literature was excluded, and abstracts were reviewed. Once further irrelevant literature was excluded, full texts of the studies identified as being potentially eligible for inclusion were assessed against the eligibility criteria above. Any disputes during title screening, abstract screening or full text review were settled through discussion or by a third reviewer (CAF) when necessary.

3.3.2 Data extraction

One reviewer (the candidate) extracted the following data from full texts eligible for inclusion using a standardised data collection form in Microsoft Excel: study details (lead author, year of publication), design, country, RMD, sample size, participant demographics (gender, age range, disease duration, ethnicity), intervention type, data collection method, data analysis method, and study findings (including themes, subthemes, verbatim participant quotes, and verbatim authors' interpretations). The risk of misinterpreting the authors' original interpretations and transferring their assumptions and bias into the review was considered, but triangulation of participant quotes and authors' interpretations was undertaken with the aim of adding depth to this qualitative synthesis. This method captures analytic insights, leading the identification of nuances that may not have been recognised from participant quotes alone (136).

3.3.3 Quality appraisal

Tools designed to critically appraise the quality of qualitative studies were presented in Chapter 2 (section 2.5.2.1). To summarise, two of the most commonly used quality appraisal tools in qualitative syntheses are the CASP Qualitative Studies Checklist (255) and the JBI checklist (256). Whilst the suitability of using quality appraisal in qualitative research has been debated, it was considered appropriate in this project because COS are ultimately used in quantitative research, and uptake may depend on trialists trusting the development process. As the candidate had limited previous experience in appraising the quality of qualitative studies, the CASP was selected to assess the quality of included studies in this review due to its frequent use in healthcare, suitability for a novice researcher, and endorsement by Cochrane (136).

Prior to undertaking the quality appraisal, the candidate received training from an experienced qualitative researcher (CAF), and assessed the quality of the first two papers under supervision. Two reviewers (the candidate and HJS) independently

assessed the quality of the included studies using the CASP checklist. Discrepancies were resolved through discussion or by a third reviewer (CAF). The CASP includes a question about data saturation, the appropriateness of which is now highly debatable in qualitative research (181). Therefore, this question was not considered in the appraisal.

3.3.4 Data synthesis and analysis

Methods of synthesising qualitative data were identified in Chapter 2 (Table 2.1). Framework synthesis (135) and thematic synthesis (134) were both considered for the current systematic review. Framework synthesis uses an a priori framework to extract and synthesise findings across different qualitative studies. This framework could have been based on findings from a scoping review of domains in existing clinical trials and observational studies (73). However, as the purpose of this review is to identify domains that researchers may not have considered, an inductive approach was deemed most appropriate. The thematic synthesis method described by Thomas and Harden (2008) was selected for the current review as it aligned with the aim to synthesise multiple qualitative studies in a narrative, identifying key domains across the data (134). Thematic synthesis is a flexible and transparent method that has been used in other qualitative reviews in healthcare, including in COS development (257).

All extracted data from the findings section of each study were considered in the synthesis. One reviewer (the candidate) read each article multiple times to achieve immersion, then performed line-by-line coding of the data in Microsoft Excel, to search for concepts. Codes were independently verified by two other reviewers (HJS, CAF). Following comparisons of common convergent and divergent concepts within and across studies, codes were organised into related areas to construct descriptive themes and subthemes. This was achieved through an iterative process of translating concepts from one study to another by adding coded text to existing concepts and creating new concepts when deemed necessary. An example of the coding and categorisation carried out in Excel is presented in Appendix B. A conceptual map of the descriptive themes was developed to understand how they related to each other. Descriptive themes were then inductively analysed further to construct analytical themes (domains for consideration in a COS), to ‘go beyond’ the findings reported in the included studies and generate additional understanding relating to the COMFORT project’s specific objectives (134). Domains identified were categorised according to the core areas of the OMERACT framework (pathophysiological manifestations (signs, symptoms, biomarkers), life impact, death, and societal/resource use) (63), and added to a list of domains identified in the previous scoping review (73).

Each included article was subsequently re-read by one reviewer (the candidate) to ensure themes were represented in the primary data, and illustrative verbatim quotations were incorporated. The proposed descriptive and analytical themes were subsequently presented, discussed, and finalised with the PAG.

3.3.5 Confidence in findings

The GRADE-CERQual framework was discussed in Chapter 2 (section 2.5.2.1). To recap, it was developed to assess the confidence in the evidence from reviews of qualitative research (141). GRADE-CERQual is a structured and commonly used method in health research (258). It was chosen to assess the confidence in findings for this review to indicate how much emphasis to place on each finding when making decisions regarding the domains to include in a COS.

Two reviewers (the candidate and HJS) independently assessed the confidence in the findings using the GRADE-CERQual approach (141, 143, 259-262). Four components were considered to formulate an overall assessment of confidence in the synthesised qualitative findings: methodological limitations (using CASP, as discussed in Chapter 2, section 2.5.2.1), coherence of data, adequacy of data, and relevance of the studies. Full definitions of each GRADE-CERQual component and confidence ratings were presented in Chapter 2 (Table 2.2 and Table 2.3). Any disagreements in confidence ratings were resolved through discussion or by a third reviewer (CAF). Key review findings, confidence judgements for each finding, and an explanation of each judgement are presented in a ‘summary of qualitative findings’ table.

3.3.6 PPI

The PPI contributors within the PAG (PR, CH) were actively involved throughout this review. They contributed to of the design of the study, inputting into the search strategy by identifying search terms that the candidate had overlooked. These included terms relating to different types of toe deformity (e.g. “hammer toe”). Both PPI contributors inputted into the analysis by contributing to ongoing discussions about the findings, identifying and naming descriptive themes, and interpreting descriptive themes to construct domains for potential inclusion in a COS. PPI contributors recognised their own experiences in the review findings.

3.4 Results

3.4.1 Search strategy

The searches yielded 1,443 records, of which 42 were retrieved for full-text screening. Thirty-four studies representing 32 data sets met the inclusion criteria. The full selection process is presented in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 flow diagram (Figure 3.1) (146).

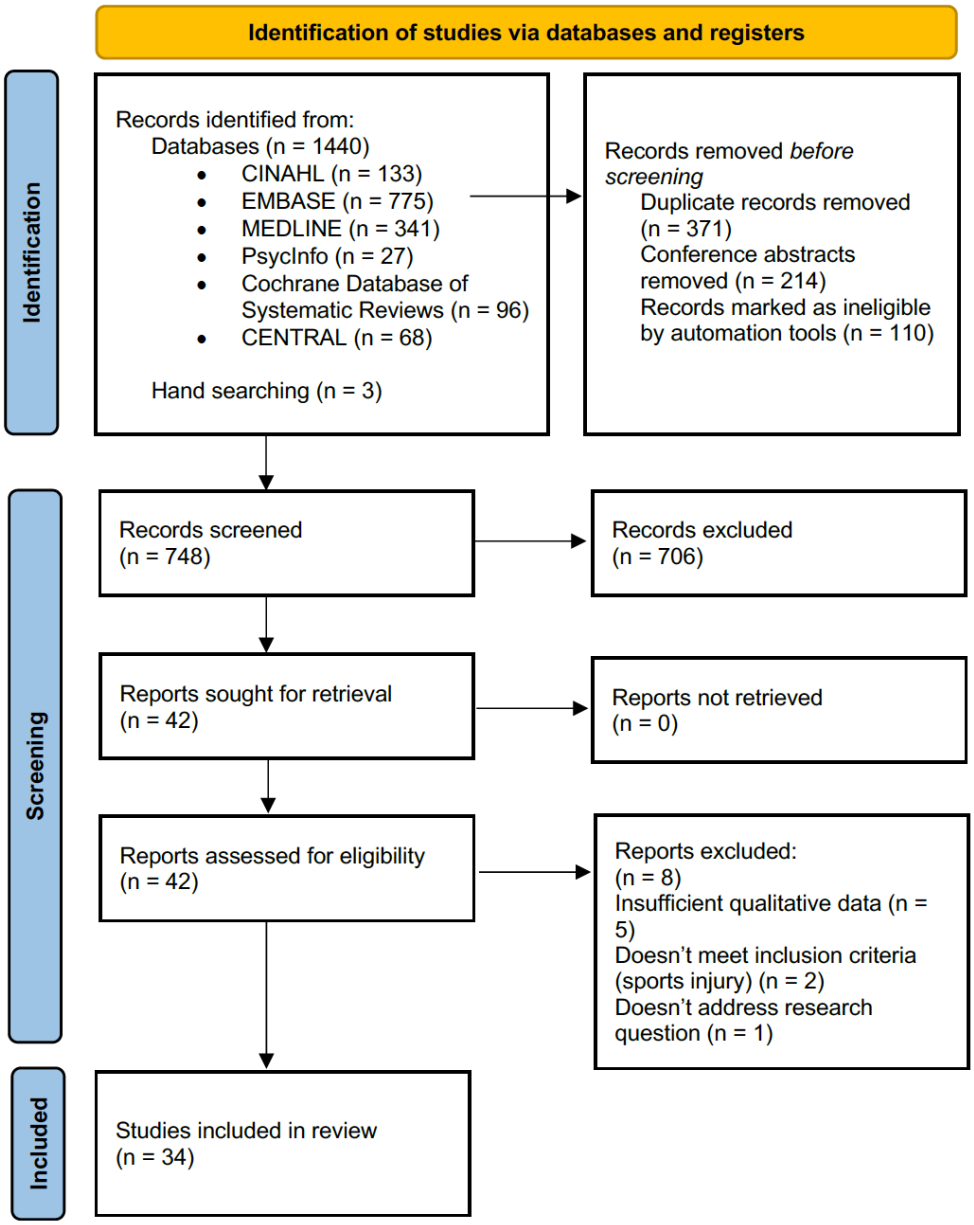


Figure 3.1: PRISMA flow diagram

An overview of the characteristics of included studies is provided in Table 3.1. The sample included a total of 503 participants. Studies were conducted in seven different countries and included participants with RA (182, 263-279), OA (280-284), gout (285-287), PsA (288), SLE (289), PTTD (290), PHP (291), AT (292), and mixed conditions. Mixed condition studies included participants with hallux rigidus or hallux valgus (293, 294), and RA or OA (295). Eighteen studies focused on participants' experiences of a specific intervention.

Table 3.1: Study characteristics

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
Backhouse 2016 (263)	Phenom-enological	UK	RA (post-surgery)	10	8 women, 2 men; age range 33-81; disease duration range 6-30 years; 1-6 months post-operative; ethnicity not reported.	Surgical (post-surgery)	Semi-structured interviews	Inductive thematic analysis	Five themes: Functional ability; participation; appearance of feet and footwear; pain; surgeons' opinion.
Bjork 2018 (264)	Critical incident technique	Sweden	RA	59	Age range 20–63 years; other demographics not reported.	N/A	Semi-structured interviews	Content analysis	Five themes: Foot hindrances in domestic life; leisure activities affected by one's feet; foot impairments influencing work; struggling to be mobile; foot impairments as an early sign.

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
Blake 2013 (265)	Case study approach	UK	RA	9	4 women, 5 men, age range 40-72 (mean 58), 1-30 years duration (mean 12.6); ethnicity not reported.	Conservative (podiatry)	Semi-structured interviews	Framework approach to content analysis	Two themes: Patient's knowledge and attitude to foot problems; patient's attitude to seeking help
Campbell 2019 (290)	IPA	UK	PTTD	5	2 women, 3 men; age range 40-80; disease duration 7-20 months; ethnicity not reported.	Conservative	Semi-structured interviews	IPA	Three themes: Adverse experience during the patient journey; treatment burden; negative self-concept.
Carter 2018 (288)	Qualitative interview	Australia, New Zealand	PsA	21	13 women, 8 men, mean age 53 (SD 13), mean disease duration 11 (9) years; 18	N/A	Semi-structured interviews	Constant comparative analysis and thematic framework approach	Three themes: foot and ankle structural and functional manifestations of PsA; the impact of foot problems on the lives of people with PsA; mediating factors that

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					Caucasian, 2 Fiji-Indian, 1 Indian.				influenced the severity of impact from foot problems on the lives of people with PsA;
Ceravolo 2020 (292)	Grounded theory	Australia	AT	11	Demographics not reported.	N/A	Focus groups and semi-structured interviews	Grounded theory	Seven themes: Adapting lifestyle; living with the condition; changes in mental well-being; conflict with identity, individual experiences; frustration; change in social well-being.
Conlin 2021 (280)	Qualitative descriptive approach	Canada	Ankle OA	10	8 men, 2 women, age range 59-90; disease duration not reported; ethnicity not reported.	Surgical	Semi-structured interviews	Content analysis	Four themes: Participants associate common characteristics with their ankle fusion and TAR; movement involving both an ankle replacement and ankle fusion

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
									necessitated learning to accommodate the characteristics of each reconstructed ankle; Individuals favor leading with one ankle reconstruction over the other for movement initiation and during transitional movements; Individuals monitor their surroundings for potential environmental challenges to facilitate safe movement.
Cotchett 2020 (295)	Qualitative descriptive design	Australia	PHP	18	12 women, 6 men; mean (SD) age 58.2 (6.6) years; mean (SD) duration of heel pain 15.9	N/A	Semi-structured interviews	Framework analysis	Seven themes: Impact of PHP; perceptions about PHP; coping with PHP; source of information; patient needs; patient

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					(16.3) months; ethnicity not reported.				unmet needs; advice to others.
Dando 2020 (295)	Qualitative	UK	RA/OA	19	8 women, 11 men; other demographics not reported.	N/A	Four focus groups	Thematic analysis	Four themes: Systems working together; finance; understanding what podiatry services have to offer; person factors of foot pain.
de Souza 2016 (293)	Qualitative	UK	RA	9	8 women, 1 man; age range 27–68 (mean 50) years; disease duration range 4–46 (mean 16.6) years; 7 White, 1 Black, 1 Mixed Race.	N/A	Two focus groups	Inductive thematic analysis within a realist paradigm	Four themes: Need for foot health information; feet ignored during routine consultations; frequency of foot examination; access to podiatry.

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
Dismore 2021 (293)	Interp-retivist paradigm with a subjective ontology	UK	Hallux valgus/ hallux rigidus	16	14 women, 2 men; age range 45-73 (mean (SD) 61 (7.23)) years; disease duration not reported; all White Caucasian.	Surgical (pre-surgery)	Semi-structured interviews	Inductive thematic analysis	Three themes: The impact of pain; decision-making process; body image, the-self and identity.
Dismore 2022 (294)	Qualitative	UK	Hallux valgus/ hallux rigidus	15	14 women, 1 man; age range 45-73 years; disease duration not reported; all White Caucasian.	Surgical (post-surgery)	Semi-structured interviews	Thematic analysis	Five themes: Physical limitations; the psychosocial impact of recovery; regaining normality; patients; expectations for recovery; an altered body-image.
Firth 2011 (182)	Qualitative interview	UK	RA	23	6 men, 17 women; age range 45-88 years; disease duration range 5-64	N/A	Individual interviews	Framework analysis	Two themes: The patient journey, patient-professional relationship.

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					years; ethnicity not reported.				
Firth 2013 (267)	Qualitative design	UK	RA	23*	6 men, 17 women; age range 45-88 (mean 69 (SD 10)) years; disease duration range 5-64 (mean 23 (SD 16)) years; ethnicity not reported.	N/A	Individual interviews	Framework analysis	Three themes: Physical impact; social consequences; impact on psychological wellbeing.
Frecklington 2019 (285)	Qualitative design	New Zealand	Gout	11	9 men, 2 women; age range 40-83; disease duration range 2-25 years; 3 Māori, 4 NZ European, 3 Pacific Island, 1 South African.	Conservative	Individual interviews	Thematic analysis	Three themes: Comfort as a priority; knowing what to buy; challenges of different environments.

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
Goodacre 2011 (268)	Qualitative symbolic interaction-ist approach	UK	RA	15	All women; age range 38-75; disease duration range 1-47 years; ethnicity not reported.	Conservative	Two individual interviews	Thematic network analysis	Three themes: Changing feet; finding the right fit; it's all part of the outfit.
Hendry 2013 (270)	IPA	Australia	RA	12	12 women; age range 44-83; disease duration 3-34 years; ethnicity not reported.	Conservative	Semi-structured interviews	Thematic analysis	Five themes: Impact of disease-related foot symptoms; footwear difficulties; medical/rheumatology encounters; foot and podiatry care access and experiences; financial hardship.
Hoque 2022 (269)	IPA	UK	RA	8	7 women, 1 man; age range 40-68; disease duration 3-	N/A	Semi-structured interviews	Thematic using	Three themes: Feet are a priority in RA; existing methods of measuring foot disease are

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					56 years; ethnicity not reported.			principles of IPA	inadequate; implementation.
Laitinen 2022 (271)	Descriptive qualitative study	UK	RA	20	18 women, 2 men; age range 24-83 (mean (SD) 64 (13); disease duration not reported; ethnicity not reported.	N/A	Semi-structured interviews	Inductive content analysis	Two themes: The hindering factors for foot health; the facilitating factors for foot health.
Liddle 2015 (286)	Qualitative design	UK	Gout	43	29 men, 14 women; age and disease duration range/mean not reported; 40 White British, 3 Asian British.	N/A	Semi-structured interviews	Thematic analysis	Six themes: patients' interpretations of symptoms; decisions about seeking medical attention; triggers and delays with the diagnostic interval; accepting or doubting the diagnosis; thoughts and feelings on receiving the diagnosis;

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
									actions on receiving the diagnosis.
Naidoo 2011 (272)	Modified IPA	UK	RA	8	All women; age range 36-84; disease duration range 3-34 years; ethnicity not reported.	Conservative	Semi-structured interviews	IPA	Six themes: The nature of foot complaints and foot deformities; aesthetic appearance and design of footwear; body image; psychosocial aspects; perceptions of footwear; the therapeutic value of retail shoes.
Pinsker 2020 (281)	Phenom-enological	Canada	Ankle OA (post-surgery)	25	12 women, 13 men; age range 25-82 years; disease duration not reported; duration since surgery range 1-10 (mean 3.2)	Surgical (post-surgery)	Semi-structured interviews	Giorgi's method for analysing interview data	Four themes: Description of vigilance; vigilance and environmental factors; vigilance and perceptions of outcome; factors influencing perceptions of vigilance.

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					years; ethnicity not reported.				
Ramos-Petersen 2021 (273)	Qualitative approach with thematic framework	Spain	RA	6	All women; age range 32-75; disease duration range 1.5-45 years (mean 17.8 years); ethnicity not reported.	Conservative	Semi-structured interviews	Inductive thematic analysis	Three themes: Improvement in physical activity; footwear … a tricky situation; social implications of RA feet.
Richardson 2015 (287)	Qualitative design	UK	Gout	14**	14 women; age and disease duration range/mean not reported; 13 White British, 1 Asian British.	N/A	Semi-structured interviews	Thematic analysis	Four themes: onset and diagnosis; understanding gout; identity; and roles and relationships.
Sanders 2017 (274)	IPA	UK	RA	5	4 women, 1 man; age range 35-78 years (mean 64);	Pharmacological	Semi-structured interviews	Thematic analysis within the	Four themes Life before biologic therapy; life with biologic therapy; sense of

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					disease duration range 6-32 years (mean 20.2); ethnicity not reported.			IPA framework	self; podiatric implications.
Tehan 2019 (275)	Qualitative approach	New Zealand	RA	20	All women; age range 27-75; disease duration range 3 months-45 years; 8 NZ European, 5 Pacific Island; 6 Asian; 1 Māori.	Conservative	Semi-structured interviews	Reflexive thematic analysis	Three themes: Comfort's number one; "I don't want to wear Nana shoes, and so many comfy shoes are Nana shoes"; "come live with my feet and you'll understand".
Thomas 2013 (282)	Qualitative interview	UK	Foot OA	11	6 women, 5 men; age range 56-80 years; disease duration not	N/A	Semi-structured interviews	IPA	Three themes: Day-to-day impact and managing symptoms; searching for

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					reported; ethnicity not reported.				explanations; consulting and meeting needs.
Williams 2007 (276)	IPA	UK	RA	13	10 women, 3 men; women age range 44-76 (mean 59) years, men age range 50-57 (mean 53) years; women disease duration range 5-26 (mean 14) years, mean disease duration range 4-12 (mean 6) years; ethnicity not reported.	Conservative	Semi-structured interviews	Hermeneutic phenom-enological analysis	Five themes: Feelings about their feet; feelings about their footwear; behaviour with the footwear; feelings about the practitioner; feelings about what would have improved their experience.
Williams 2010 (277)	IPA	UK, Spain,	RA	30	All women; UK participants mean (SD) age 57 (10.68), Spain	Conservative	Semi-structured interviews	Thematic analysis	Six themes: Feet being visibility different because of rheumatoid arthritis; the referring practitioner's

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
		Nether-lands			participants mean (SD) age 57 (SD 6.01), Netherlands participants mean (SD) age 11.93; UK participants disease duration mean (SD) 15 (6.72) years, Spain participants disease duration mean (SD) 13 (4.96) years, Netherlands participants disease duration mean (SD) 15 (5.16) years); ethnicity not reported.				approach to the patient; the dispensing practitioner's approach to the patient; the footwear being visible to others; footwear influencing social participation; the women's wishes for improving their experience.

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
Williams 2012 (278)	IPA	UK	RA	22	16 women, 6 men; women mean (SD) age 58 (11.9), men mean age (SD) 59 (6.0) years; women mean (SD) disease duration 15 (5.2) years, men mean (SD) disease duration 13 (5.0) years; ethnicity not reported	N/A	Focus groups	Thematic framework	Five themes: the significance of foot symptoms in relation to the diagnosis of rheumatoid arthritis; knowledge of and explanation about foot symptoms; Accessing foot health interventions; the effectiveness of foot health interventions; Improvements to foot health interventions.
Williams 2017 (289)	IPA	UK	SLE	12	All women; age range 42-72 years (SD 9.26); disease duration range 11-	N/A	Conversational interviews	Thematic framework approach	Six themes: Foot problems and symptoms; the impact of foot problems and symptoms on activities; disclosure

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
					35 years (SD 8.07); all White Caucasian				and diagnosis of foot problems; treatment of foot problems and symptoms; perceived barriers to professional footcare; need for professional footcare and footcare advice.
Wilson 2017 (279)	Qualitative interview	UK	RA	7	All women; age range 29-72 (mean 56); disease duration range 2-27 years; ethnicity not reported.	N/A	Semi-structured interviews	Inductive thematic analysis	Two themes: Impact of foot problems; decision to access foot care.
Yeowell 2021 (283)	Exploratory qualitative research design, using	UK	Ankle OA	9	8 men, 1 woman; age range 30-70 (mean 54); disease duration range 1-20 years (median 2	Conservative	Semi-structured interviews	Reflexive and inductive	Four themes: Signs and symptoms; impact on participation; impact on self-identity and mental

Study ID (lead author, year)	Design	Country	RMD	Sample size	Participant demographics	Intervention type	Data collection method	Data analysis method	Key findings (themes)
	hermen-eutic phenom-enology.				years); ethnicity not reported			thematic analysis	wellbeing; views on non-surgical management.
Zaidi 2013 (284)	Qualitative, interview-based approach	UK	Ankle OA	14	6 men, 8 women; age not reported; disease duration range 10-40 years; ethnicity not reported	Surgical (post-surgery)	Semi-structured interviews	Inductive thematic analysis	Three themes: Why patients opt for surgery; information sources for decision making; how patients decide the best option for them.

AT, Achilles tendinopathy; IPA, interpretative phenomenological analysis; OA, osteoarthritis; PHP, plantar heel pain; PsA psoriatic arthritis; PTTD, posterior tibial tendon disorder; RA, rheumatoid arthritis; SD, standard deviation; SLE, systemic lupus erythematosus

3.4.2 Quality appraisal

The frequency of responses ('yes' or 'no') to each signalling question in the CASP checklist is shown in Appendix C. Strengths observed in all studies included: clearly stated objectives, appropriate methodology and design, justification as to why the participants selected were the most appropriate to provide access to the type of knowledge sought by the study, clarity regarding how categories/themes were derived from the data, sufficient data presented to support the findings, explicit findings, adequate discussion of the evidence both for and against the researchers' arguments, findings discussed in relation to the original research question, discussion of contribution of the study to existing knowledge or understanding). Strengths observed in at least 31 studies included: explanation of how participants were selected, explicit and justified data collection methods, discussion of issues raised by the study and ethics approval, in-depth description of the analysis process, discussion of credibility of findings, identification of new areas where research is necessary, discussion of transferability of findings. The following limitations were identified in at least 20 studies: no discussions around recruitment, no justification of the setting for data collection, no critical examination of the researchers' own role, potential bias and influence during the formulation of the research question and data collection, insufficient details of how the research was explained to participants, lack of contradictory data taken into account, no critical examination of the researchers' own role, potential bias, and influence during analysis and selection of data for presentation.

3.4.3 Synthesis of qualitative studies

The thematic synthesis identified seven descriptive themes: pain, change in appearance, limited activities, social isolation, work disruption, financial burden, and emotional distress. Descriptive themes were further analysed to identify domains that should be considered for inclusion in a COS, to specifically address the overall research aim. Domain granularity was discussed in Chapter 2 (section 2.5.2.3); domains in this review are presented as proposed 'broad domains' and 'target domains', in line with OMERACT terminology (63).

3.4.3.1 Descriptive themes

Descriptive theme 1: Pain

Participants in all studies described experiencing foot and ankle pain. Pain occurred on first step, in the morning, or in the evening, during or after activity, or at rest, with some

participants describing spreading, burning or throbbing pain. Participants reported different severity levels of foot and ankle pain, with some describing it as unrelenting and unbearable:

> "The pain is horrific. It's just terrible, I wish someone could … you know what, I would have it cut off and a false one there if they could. It's horrendous pain, it's terrible. It's driving me round the bend. I'd go for anything to get rid of this pain, I'd try anything now." - participant with OA (283)

> "Foot pain was the most influential symptom experienced by study participants and was described as the worst aspect of the disease by the majority." – authors' interpretation (participants with RA) (270)

Pain linked closely to most other themes; it affected daily activities, influenced social participation and occupational function, and caused emotional distress:

> "If your feet don't work or they are painful it impacts on the whole quality of your life from what you do recreationally, socially, work if you are still working. It impacts on everything." – participant with RA/OA (295)

Participants discussed how foot and ankle pain caused fatigue (264, 269, 270, 273, 294) and impaired sleep (182, 285, 288, 293, 294). Foot and ankle pain also led to issues with footwear; it impacted on the ability to wear regular shoes, and was sometimes caused or made worse by footwear (182, 266, 272, 273, 275, 276, 283, 285, 288, 290, 291, 293, 294):

> "You've got a formal or a fancy event to go to, you kind of, you just sacrifice as I've said earlier you deal with the consequence tomorrow because this looks right or this is more appropriate for that activity so you just basically suck it up and consequences come tomorrow." – participant with gout (285)

> "Women expressed a desire to wear footwear which were feminine, however, in most instances were unable to wear these types of shoes due to their RA-related foot pain and deformity." – authors' interpretation (participants with RA) (272)

Descriptive theme 2: Change in appearance

Many participants reported foot and ankle deformity (263-266, 268-279, 282, 285, 286, 288, 289, 293-295):

> “My feet have undergone several surgical procedures, my toes are crossed and I cannot spread them out, my big toe grows against the other toes. I don’t like to look at them, or care for them.” – participant with RA (271)

Some participants described swelling of their foot and ankle (264, 266, 268, 269, 271, 279, 283, 286, 288, 294). Swelling occurred at rest, or during or after activity.

> “If I want to play golf I just go out in a buggy [golf buggy to avoid walking], but you do a lot of twisting. And I can see it, even though I've got my socks on, I can see it is all swollen and it's throbbing.” – participant with OA (283)

Prescence of callosities, dry skin, and nail pathologies were also reported (263, 265, 267, 269, 270, 278, 279, 288, 289):

> “They split and they can bleed because they get so dry.” – participant with SLE (289)

Some participants (265, 267, 279) discussed change in appearance in terms of different types of foot infections and the impact of foot wounds, including wound healing time and recurrence of wounds:

> ‘‘They take that long to heal. They do heal but then within three or four weeks, if I just do a bit more walking than normal, they’re back again.’’ – participant with RA (182)

A minority of participants highlighted changes in circulation (289).

Change in appearance of the feet or ankles impacted on participants’ choice of footwear (182, 263-265, 268, 270-279, 282, 285, 287-290, 293-295) and subsequent clothing options (182, 268, 270-272, 275, 276, 278, 279, 288, 293, 294):

> “I think about them all the time, because when I’m going out, I think ‘do my bunions stick out, do my shoes look alright.’ I’d love to wear strappy sandals and your flip flops and all that sort of thing, but I never would.” – participant with OA (282)

> “Another element of concern expressed about the effects of changed physical appearance was the focus on the accompanying special footwear often prescribed for patients with particular foot problems, such as deformity.” – authors’ interpretation (participants with RA) (274)

Descriptive theme 3: Activity limitations

Many participants had difficulties walking (182, 263, 264, 266, 268-271, 273-277, 279-283, 286-289, 291, 293-295). Participants emphasised the importance of being able to walk, and the negative impact of reduced mobility as a result of foot and ankle problems, particularly pain:

> "Some days I really can't walk, the pain is so bad [referring to foot pain]. I've never broken a bone in my life. But if I had to imagine what a broken bone felt like, that's what it feels like when I walk." – participant with RA (279)

> "On some level, for all participants, mobility was restricted or affected by foot pain." – authors' interpretation (participants with OA) (282)

In contrast, some participants identified that walking was affected even when pain had reduced following surgical intervention, leading to negative body image:

> "It's [the surgery] taken the worse pain away. I mean I still walk like a waddling duck." – participant with RA (263)

Foot and ankle disorders reduced step length, walking speed and distance, and led to difficulty walking on different terrains, and up or down stairs and slopes. Participants also expressed how their foot and ankle disorders meant they couldn't walk normally; they described shuffling, limping, or hobbling:

> "I'll be limping around right and people will be asking me 'what's wrong with your feet?'… It doesn't feel good." – participant with PsA (288)

A minority of participants also discussed reduction in muscle strength as a symptom (271), which was perceived as a barrier to being able to walk.

> "I would love to walk, but I can't because I have no muscle strength in my feet and my feet do not sustain my body weight" – participant with RA (271)

Some participants identified a reduction in balance, feeling unstable and fear of tripping or falling when walking or using stairs (264, 271, 272, 275, 280, 288); in some cases this was a consequence of a surgical implant or of the footwear participants had to wear due to foot and ankle pain and deformity:

> "Every time I make a step, I look to where I'm stepping … I do that subconsciously … I watch where I step. I look for potholes in the road or wherever I go." – participant with OA (280)

Many participants described modifying, limiting or ceasing specific physical and domestic activities as a result of their foot and ankle disorders (182, 263, 264, 267, 270, 271, 273, 277, 279-282, 284, 288-295):

> “Squash, golf, I can't do any of them now; the only activity I can do is swimming. I have put a lot of weight on as a result of not being able to do what I used to; I just hope to get back to doing something.” – participant with OA (284)

> “I do play bowls. I don’t play as much now, because I have to stand all of the time … it’s not so easy anymore.” – participant with PTTD (290)

> “The patients also described foot impairments during domestic activities such as assembling furniture while kneeling, climbing a ladder when painting the house, shovelling snow, and in gardening.” – authors’ interpretation (participants with RA) (264)

Some participants needed to drive instead of walk (273), whereas others explained how their foot and ankle disorders made driving difficult (264, 288). Joint stiffness, and lack of movement in the foot and ankle joints, was also discussed by participants (264, 269, 271, 273, 279, 280, 282, 283, 288). Stiffness was eased by activities in some cases, but occurred as a result of activity in others.

> “And when I get that [referring to a flare] I’m more conscious of my feet. When I’m not having a flare it’s only when I’ve walked too far or … stood for a long time. And then I become aware that my toes have become a bit stiff and my heels hurt.” – participant with RA (271)

Some participants identified the importance of joint movement for everyday activities:

> “That’s the biggest difference that I find. Just walking, any kind of movement, trying to run, or trying to catch a ball or trying to chase the [grand]kids. . . . It’s easier because I’ve got more movement in the left [replaced] ankle.” – participant with OA (280)

Descriptive theme 4: Social isolation

Foot and ankle disorders impacted on many participants’ social lives (182, 263, 264, 268-273, 275-277, 279, 281-283, 288-294). Participants experienced changes in their family roles and friendships, describing social withdrawal and isolation as a result of not being able to do their usual activities because of their foot and ankle problems:

> “I’m sat in the house, I can’t do nothing. I have no social life. My friends, they all say come and have a pint, but what’s the good in going for a pint when I’m sat there, I can’t move, I can’t go to the bar, I can’t get to the toilet.” – participant with OA (283)

> “There are things that you can do, whereas I don’t know about you guys, but when I had the Achilles, it was like being in gaol, that real restriction of your freedom for things that you enjoy doing.” – participant with Achilles tendinopathy (292)

> “The experience of pain was also an obstacle to being socially active, with some participants expressing a feeling of social isolation.” – authors’ interpretation (participants with PHP) (291)

In contrast, participants whose foot and ankle disorders had improved with treatment described positive changes in their social lives (273).

> “Some reported an improvement in their social life as a result of a reduction in symptoms of RA in their feet, connecting physical comfort with their activity levels and associating perceived improvement in their physical activity and general wellness with improved quality of life.” – authors’ interpretation (participants with RA) (273).

Limitations in footwear and clothing choices occurring as a result of foot and ankle disorders also had a negative impact on social participation, particularly for women (182, 264, 268, 270, 272, 275-277, 288):

> “I panic when I do get an invite, I think oh gosh these boots. I was invited to a wedding and just sat at home and cried.” – participant with RA (277)

> “Yes, yeah, definitely. I did, I mean, and I noticed all the other women’s footwear at the wedding and there was loads of really, really high shoes. I mean mainly young girls but there were some really classy shoes there and I thought, ‘That’s me goodbye forever.’” – participant with gout (287)

However, some participants described feelings of satisfaction when they found comfortable footwear (285).

Descriptive theme 5: Work disruption

Foot and ankle disorders negatively affected participants' work lives (275, 279, 281, 282, 284, 286, 287). Participants described difficulty in performing various job roles, restricting or changing their work activities and working hours, and taking time off work because of their foot and ankle problems:

> "Being self-employed, I was getting to the point where I was working and I couldn't work, so I was losing my self-respect as far as not been doing enough work every day right? I wasn't doing my work 100%." – participant with OA (282)

> "With my job, if we have to tie the ship up and I have to swing over the side to take the cables, I often think about how I'm going to be landing on the dock … Everything I'm doing, I've got to think." – participant with OA (281)

Other participants had to give up work entirely (269, 288, 293):

> "I actually gave up work because I just couldn't. I was a trainer and I could only train standing up so I had to give up a job I loved and it was all wrapped around pain." – participant with hallux valgus/hallux rigidus (293)

> "I quit my job...because it's mostly you have to stand... get on the ladder... physically get down on the floor... It's very physical." – participant with PsA (288)

Some participants worried they appeared lazy at work, or in unemployment, and perceived a lack of understanding among their work colleagues or within employment services because their foot and ankle disorders were hidden:

> "So that's the thing, you might come with a broken arm to work, in plaster, it's very easy to understand, but going there with small unrecognizable problems like hurting feet. . .that's tiresome and not so easy to understand." – participant with RA (264)

> "They expressed that their daily experience of living with foot pain was invisible, and something that could not possibly be understood by their peers, and this influenced their full participation in the workforce." – authors' interpretation (participants with RA) (275)

The impact of foot and ankle pain and deformity on footwear and clothing choices also affected participants' work. Participants discussed the difficulties of being unable to wear certain types of footwear at work (e.g. safety boots or smart shoes) (264, 268, 275, 279, 285, 288, 293), and felt judged by their footwear choice:

> "Footwear for work became a problem…I was conscious if I wore a dress or a suit for meetings or something like that of the shoes that I had to wear." – participant with hallux valgus/hallux rigidus (293)

Descriptive theme 6: Financial burden

Participants discussed the personal financial burden of living with foot and ankle disorders. In some cases, this related to having to pay for podiatric care to address foot and ankle pain, deformity, and skin and nail complaints (264, 269, 270, 287):

> "Many respondents were persevering with their disease-related foot problems because they couldn't afford foot care. Some participants conducted trade-offs between items that they normally included in their budget, in order to pay for foot care." – authors' interpretation (participants with RA) (269)

A minority of participants also highlighted the financial implications of needing transport to attend an increased number of appointments because of their foot and ankle disorders (26).

Footwear restrictions also led to financial burden (266, 267, 269, 272, 274, 278, 284) some participants highlighted the excess costs associated with buying multiple pairs of unsuitable shoes or needing to buy specialist shoes, whilst others were unable to afford new footwear:

> "My boots, the ones that I've worn all the way through the winter, but they haven't been comfortable. But I couldn't afford to buy another pair so I just had to make do." – participant with RA (278)

> "Many women discussed financial pressures contributed to making bad choices in relation to footwear, and this led to anxiety about purchasing decisions." – authors' interpretation (participants with RA) (274)

For some participants, the impact of foot and ankle disorders on their work resulted in loss of income (264, 268, 278, 283, 285, 288):

> "I'd worked with it for three weeks, walking on the side of my foot. It's probably the worst pain I've had, in my life. I was a piece worker so what I made I got paid for, if I didn't make it, I didn't get paid, so and, you know, we were, young family then, so eventually I just had to give in and go to the doctors and get signed off for a week or two." – participant with gout (286)

> "The ability to work was a key factor that not only induced a perceived need for surgical intervention but also resulted in patients delaying the timing of surgery. In these cases, loss of earnings during the postoperative rest period was the reason for putting off surgery. This represents a 'worker's paradox' since surgery is required to continue in employment, but the temporary loss of earnings during the recovery period is seen as being prohibitive." – authors' interpretation (participants with OA) (284)

Descriptive theme 7: Emotional distress

Foot and ankle disorders affected participants' emotions (182, 263, 264, 267, 269-273, 275-279, 281-283, 287-295). Participants described feeling anxious, angry, sad, hopeless, frustrated, ashamed, paranoid, self-conscious, embarrassed, and depressed about their foot and ankle problems:

> "I'm just useless, just because of a daft ankle. It's unbelievable that isn't it. It makes me feel as if I'm good for nothing, I might as well just turn it in, you know, just go for a couple of tablets and I'll call it a day. Just a waste of time. I'm good for nothing at the minute. I feel like crying. It's horrible. Every day of my life; it gets a bit upsetting. You just wanna give in, in the end, you get sick of it." – participant with OA (283)

> "Anxiety about the anticipated duration of ulceration was common and for some participants this was linked to specific concerns that an open wound might delay surgery or treatment." – authors' interpretation (participants with RA) (267)

Emotional distress occurred as a direct result of pain and change in appearance of the feet or ankles, and of the subsequent reduction in physical, social, and work activities:

> "I don't want to overstate the cranky and anger stuff, but there's definitely a general feeling of – it's almost depression, but not clinical depression, but you

> just don’t feel good about yourself or the world.” – participant with Achilles tendinopathy (292)

> “One of the best things about football was going for a drink with the boys afterwards. So I stopped going. Not only could I not play anymore but the loss of the social side, not seeing my mates and all that made me feel really low.” – participant with RA (279)

Change in appearance of the feet and ankles impacted on body image and self-identity (263, 268, 272, 273, 275, 276, 282, 288, 293, 294); participants identified stigma associated with disability, and described feeling visibly different to “normal” people. Participants with deformity felt embarrassed by the appearance of their feet, hated their feet, and did not want to look at their own feet, perceiving that other people did not want to look at their feet either:

> “You can see it in their faces, like I don't want to look at your feet … feet phobia.” – participant with RA (273)

Participants expressed wanting to hide their foot and ankle problems from others. Similar experiences were expressed by those who reported having deformity prior to surgical intervention, whilst those without deformity discussed how they did not want deformity to occur:

> “Well just the looks of your feet really you know, just the looks of the feet, they were awful they really.....they look a lot better than they did and you know I never liked to take my shoes off before or my socks off or anything which you know I don't mind now sort of thing.” – participant with RA (263)

> “I'd deal with it if my toes went crooked. I'd just think, well that's just part of [RA] but I don't really want them to get that unsightly. Your toes aren't the nicest things.” – participant with RA (265)

Emotions were similarly affected by dissatisfaction with footwear and clothing choices (182, 268, 272, 274-279, 287, 288, 290, 293, 295):

> “I can go in a shoe shop and within 5 or 10 minutes I can be in tears trying on shoes because I just absolutely hate my feet because of the way they have gone with the arthritis. The thing that I absolutely hate shopping for is shoes.” - participant with RA (268)

> “Powerful emotions of shame, sadness and frustration were clearly identified by these women when speaking about their feet, footwear and body image.” – authors’ interpretation (participants with RA) (272)
>
> "The shoes....as soon as I see a person I can say oh yes she's got hospital shoes on..... I compare my boots with other people and they are more feminine and pretty and that makes me feel sad.” – participant with RA (276)

In contrast, some participants who had undergone surgery for foot and ankle disorders and were satisfied with the outcome reported feeling more optimistic, with improved psychological wellbeing, mood, and quality of life (293, 294).

3.4.3.2 Analytical findings

As described above, descriptive themes were analysed further to identify and categorise domains that could be considered for inclusion in a COS. These are presented in Appendix D alongside the domains identified from the scoping review of existing studies in this area (73) and domains identified in subsequent chapters.

3.4.4 Assessment of confidence in the review findings

Moderate confidence in most of the review findings was established (Table 3.2). A detailed GRADE-CERQual Qualitative Evidence Profile is also presented in Appendix E. This was due primarily to concerns regarding methodological limitations, adequacy of data, and relevance of each contributing study to the review question (given the minority of studies including participants with SpA and CTDs, and that all data came from high-income countries), for some of the findings.

Table 3.2: GRADE-CERQual summary of findings

Summary of review finding	**CERQual assessment of confidence in the evidence**
Patients with foot and ankle disorders in RMDs experience pain.	Moderate confidence
Patients with foot and ankle disorders in RMDs experience joint deformity.	Moderate confidence
Patients with foot and ankle disorders in RMDs experience joint swelling.	Low confidence
Patients with foot and ankle disorders in RMDs experience skin and nail complaints, including wounds.	Moderate confidence
Patients with foot and ankle disorders in RMDs experience limitations in walking and changes in gait.	Moderate confidence
Patients with foot and ankle disorders experience a reduction in physical and domestic activities.	Moderate confidence
Patients with foot and ankle disorders in RMDs experience joint stiffness.	Moderate confidence
Patients with foot and ankle disorders in RMDs experience a reduction in social participation.	Moderate confidence
Patients with foot and ankle disorders in RMDs experience disruptions in work.	Low confidence
Patients with foot and ankle disorders experience dissatisfaction with footwear.	Moderate confidence
Patients with foot and ankle disorders experience financial impact.	Moderate confidence
Patients with foot and ankle disorders experience emotional distress.	Moderate confidence

3.5 Discussion

This qualitative synthesis achieved its aim to identify domains that are important to patients with RMDs who have experienced foot and/or ankle disorders. It has established the symptoms that patients experience, explored the impact of these symptoms, and interpreted data in the context of development of a COS, whilst considering the quality of individual studies and confidence in the review findings. Additionally, this review has identified gaps in the existing literature base to inform future qualitative research.

3.5.1 Main findings

Findings from this review indicate that foot and ankle disorders are debilitating and affect multiple aspects of life, causing considerable disruption for patients with RMDs. Both pain and change in appearance of the feet and ankles directly led to other issues, including activity limitations, social isolation, work disruption, financial burden, issues with footwear, and emotional distress, therefore addressing these specific symptoms or manifestations may lead to improvements in others. Patients' experiences of foot and ankle symptoms and their impact appear to be similar regardless of RMD, supporting the scope of the proposed COS and suitability of a single core domain set for foot and ankle disorders across multiple RMDs, although not all RMDs in the scope of the COS have been explored in qualitative research.

In congruence with a previous scoping review of domains in existing clinical trials and observational studies (73), pain was a commonly reported symptom in this thematic synthesis. Functional limitations was also a prominent finding, aligning with what has been measured previously. In contrast, whilst the scoping review revealed other objective domains of potential importance to researchers, such as joint range of motion (ROM), disease activity and gait biomarkers, these findings were less prominent in the qualitative literature. Other key findings from the thematic synthesis were that foot and ankle disorders impacted on social and occupational function, footwear, and emotional status. These were rarely specified by researchers as domains to measure in existing clinical trials and observational studies, although they may have been captured in broader measures of function. Previous clinical trials and observational studies have measured patient satisfaction with different interventions. In this qualitative synthesis, experiences of satisfaction or dissatisfaction with treatments was predominantly related to footwear.

The current study identified additional domains of importance to patients that were not reported in previous clinical trials and observational studies. These included balance issues/feeling unstable, nail pathologies, circulation, ulceration, sleep, body-image/self-identity, fear of falling, and clothing options. Additionally, the thematic synthesis increased understanding of the scope of certain domains. For example, whilst presence of callosities has been measured in previous clinical trials, the qualitative synthesis identified additional skin pathologies that are meaningful to patients. Similarly, additional target domains were identified within the broad domains of pain and stiffness. For example, whilst the broad domain of pain has been commonly measured in existing studies, the target domains of pain on first step and pain in the morning have not been specified in previous studies. This may reflect the inclusion of a wider range of RMDs within the qualitative synthesis, such as MSK conditions (e.g. plantar heel pain), where these particular pain domains are common symptoms (51).

In previous studies, presence of deformity reflected a measurement of structural change, with most researchers using imaging (73), whereas in qualitative studies, patients were concerned about the appearance of foot and ankle deformity and how it was viewed by others. Similarly, gait mattered to patients in terms of appearance (e.g. limping), extending beyond purely objective assessments. Financial burden was also an important domain to patients in this qualitative synthesis, improving understanding of both healthcare utilisation and direct cost domains.

3.5.2 Strengths and limitations

This study was strengthened by robust methodology, utilising a well-established, systematic and transparent synthesis method (134). The study benefited from a comprehensive database search strategy co-produced with information scientists and experts in the field (including PPI contributors), and from two independent reviewers during study screening, quality appraisal and judging confidence in the findings. Peer debriefing and collaborative analysis with members of the wider research team, including PPI contributors, reduced the potential for interpretation bias (296).

Findings from this study must also be considered in light of several limitations. Firstly, the inclusion criteria were limited to qualitative studies using interview or focus group data collection methods, thus relevant qualitative data ascertained through other methods such as surveys with open-ended questions or consensus studies may have been missed. However, the intention was to conduct an in-depth exploration of the symptoms and impact of foot and ankle disorders on the lives of people with RMDs,

and it is considered unlikely that studies using other methods would have fully addressed the research aim.

Secondly, only ten of the studies included in this review reported the race/ethnicities of participants, and in studies where race/ethnicity was reported, most participants were of White background. This reflects a broader issue in health research; methods of capturing and analysing race and ethnicity are inconsistent (297), and these constructs are poorly reported in clinical trials (298). Although the review search had no restrictions on setting or country, only studies published in English were included, and all studies included in the review were from high-income countries. The findings may therefore fail to represent the views of participants from other ethnic backgrounds, LMICs, or non-English speaking participants. Additionally, the majority of studies in this review included participants with foot and ankle disorders in RA or OA, and whilst the findings appear relevant to the other included RMDs, they may not fully reflect the experiences of patients with conditions that were under-represented in this review. There were no included studies involving participants with SSc, or additional MSK disorders affecting the foot and ankle in the absence of systemic disease, such as lesser toe deformities or peroneal tendinopathy.

Thirdly, this study utilised the CASP tool to assess quality, and GRADE-CERQual to assess confidence in the study's findings. Whilst these can be considered strengths, as they enabled a systematic assessment of the methodological quality of included studies, and a transparent framework for judging confidence in each synthesised finding, neither had a formal scoring system. Both processes therefore introduced a degree of subjectivity. This was mitigated by having two independent reviewers, and by providing detailed explanations alongside ratings in the evidence profiles (Appendix E).

Finally, limitations in the overarching methodology (qualitative synthesis) must be acknowledged. None of the studies included in this review were designed to address the specific aim of the study, which was to identify domains of potential importance to patients in the context of developing a COS. Consequently, the interpretation of existing qualitative data in relation to important domains to measure may differ from findings from primary qualitative research with this specific focus.

3.5.3 Implications

3.5.3.1 Implications for clinical practice

Findings from this qualitative synthesis have implications for clinical practice. It is important for clinicians to ascertain what patients with foot and ankle disorders want treatments to achieve, so that clinical appointments can be focussed accordingly. This review specifically focused on the experiences of patients rather than exploring potential domains of importance to researchers or HCPs. Comparable to previous studies involving patients with RMDs (257, 299), the psychological impact of living with foot and ankle disorders was highlighted in this review and should not be underestimated by HCPs.

3.5.3.2 Implications for research

This study provides insight into the breadth and depth of domains that are important to patients, to inform a future consensus study and ultimately the development of a COS for foot and ankle disorders in RMDs. The review has highlighted the paucity of existing qualitative research involving participants with foot and ankle problems in RMDs other than RA and OA, languages other than English, and settings from LMICs. Additionally, race/ethnicity was under-reported in this review, reflecting a broader issue in health research. In studies where race/ethnicity was reported, participants were mostly of a White background. These gaps in the literature guide future qualitative studies.

3.6 Conclusion

This qualitative review has identified that the lives of patients with foot and ankle disorders in RMDs can be affected by pain, change in appearance of the feet/ankles, activity limitations, social isolation, work disruption, issues with footwear, financial burden, emotional distress, and treatment experiences (including satisfaction). Findings provide insight into the breadth and depth of potential domains of importance to patients inform a Delphi consensus study with all key contributors (Chapter 6), leading to an internationally agreed, standardised core domain set for foot and ankle disorders in RMDs. Patients' experiences of foot and ankle symptoms, and their impact, are similar regardless of the RMD, supporting the suitability of a single core domain set for foot and ankle disorders across multiple RMDs, although gaps in the literature have been identified. The perspectives of patients with foot problems in one under-represented RMD, SSc, will be explored in the next chapter.

Chapter 4 Secondary analysis of focus groups

4.1 Introduction

The qualitative synthesis reported in Chapter 3 identified domains of potential importance to patients living with foot and ankle disorders in different RMDs. It also exposed gaps in the existing qualitative literature, whereby the perspectives and experiences of patients with some RMDs included in the scope of the intended COS for COMFORT have never been explored. The following chapter builds on the patient perspectives described in the existing qualitative literature by reporting experiences derived from new data. It reports a secondary analysis of an existing qualitative dataset of transcripts from focus groups with patients who have systemic sclerosis-related Raynaud's phenomenon (SSc-RP). Findings have been published as:

> **Chapman LS**, Alcacer-Pitarch B, Pauling JD, Flurey CA, Redmond AC, Richards P, Herrick AL, Merkel PA, Proudman S, Menz HB, Helliwell PS, Hannan MT, Domsic RT, Saketkoo LA, Shea B, Siddle HJ. Patients' perspectives on systemic sclerosis-related Raynaud's phenomenon in the feet: A qualitative study from the OMERACT Foot and Ankle Working Group. Semin Arthritis Rheum. 2024 Apr;65:152372.
> doi: https://10.0.3.248/j.semarthrit.2024.152372

4.2 Background

As discussed in Chapter 1, one RMD within the proposed scope of the COMFORT COS for foot and ankle disorders is systemic sclerosis (SSc), a complex disease with an estimated global prevalence of 18.87 per 100,000 (300). SSc has three key features: vasculopathy, immune activation with production of autoantibodies and alteration in immune cells, and fibroblast dysfunction with excessive matrix deposition and subsequent fibrosis of the skin and internal organs (301). There are two main types of SSc: limited cutaneous SSc, and diffuse cutaneous SSc, which are differentiated by the extent of skin involvement (302).

Raynaud's phenomenon (RP), characterised by episodic vasospasm causing discolouration and pain in the digits, is a hallmark feature of SSc, affecting ~96% of patients (43). RP is typically triggered by exposure to changes in ambient temperature or emotional stress (303), and is part of SSc-related vasculopathy, resulting in

digital ischaemia. When severe, this ischaemia can lead to ulceration, necrosis, and digit autoamputation (304).

Foot problems in SSc are well recognised, with up to 67% of patients with SSc reporting having had foot involvement during the course of their disease (46, 305, 306). However, clinical trials investigating treatments for patients with SSc-related RP (SSc-RP) tend to focus on measuring RP domains in the hands (307, 308), thus overlooking RP-related foot involvement (306, 309).

This lack of focus on RP in the feet could be because the incidence of SSc-RP affecting the foot is underreported, or that the problem is underappreciated by HCPs and researchers. The consequences of RP in the feet have not been explored in depth. Qualitative research involving patients with SSc-RP in the hands demonstrated impact, and highlighted potential domains of importance for this aspect of SSc (159). It is possible however, that SSc-RP impacts patients differently based on the specific location of the RP. For example, given the weightbearing activity of the foot and issues associated with wearing footwear, the impact of SSc-RP in the feet is likely different from the impact of SSc-RP in the hands; these issues are well established among patients with other RMDs (254).

The OMERACT Scleroderma Vascular Disease Working Group are developing a COS for SSc-related RP, to address the lack of valid, reliable, and feasible outcome measures for this condition, and ultimately the lack of effective treatments (310). Thus far, the group has developed a core domain set following a scoping review of existing trials investigating pharmacological treatments, qualitative focus groups with patients, and an international Delphi consensus study. The resulting core domain set included the following domains: pain due to RP, severity of attacks of RP, duration of attacks of RP, frequency of attacks of RP, impact of RP on function, health-related quality of life, impact of RP on daily life/adaptation, hospitalisation or need for urgent intervention for RP, and death (311). Coldness, temperature, emotional wellbeing, and sensory changes (including numbness) were included as domains important for future consideration.

As SSc-related RP affects the feet, there is overlap between the OMERACT Scleroderma Vascular Disease Working Group COS and the COS being developed in the COMFORT programme of work. However, the Scleroderma Vascular Disease Working Group acknowledged that they did not include foot problems in their qualitative analysis, and did not ask Delphi participants to consider domains in relation to foot problems. It is therefore not known which domains are important to patients with SSc-RP in the feet and should be measured in future interventional trials in this area.

4.3 Aim and objectives

This study aimed to identify, from the perspectives of patients, foot-related domains that are potentially important in SSc-RP. The objectives were to:

- Explore the nature of foot-related symptoms in patients with SSc-RP and the impact of these symptoms;
- Compare these domains to those identified in the qualitative synthesis (Chapter 3), the previous scoping review of foot and ankle RMD studies (Chapter 1, section 1.5.1), and to the core domain set for SSc-RP;
- Review the scope of the COS for foot and ankle disorders in RMDs in light of any new findings.

The rationale for this study was based on the candidate's clinical and research experience. From a clinical perspective, foot problems are debilitating for patients with SSc, but patients often report that most HCPs focus on their hands. From a research perspective, the thematic synthesis in Chapter 3 identified that no previous qualitative studies have focused on the perspectives of patients living with SSc, meaning the lived experience of foot problems in SSc remains poorly understood. Additionally, in the context of developing a core domain set for foot and ankle disorders across multiple RMDs, this gap in the literature means it is unclear whether domains of importance to people with SSc align with those identified in other RMD populations, and therefore whether SSc can plausibly be included in the scope of the COMFORT core domain set. Whilst Chapter 3 identified other RMDs that have been under-represented in the qualitative literature, this chapter focuses specifically on SSc-RP due to the opportunity to conduct a secondary analysis of existing data, where foot and ankle problems had previously been overlooked.

4.4 Methods

A secondary analysis of an existing qualitative dataset was conducted. As discussed in Chapter 2, qualitative research is an important step in identifying potential domains for inclusion in a COS (63, 65). COMET also recognises the potential of secondary analysis of qualitative data as part of the foundations for developing a COS (65). The original qualitative study aimed to explore the domains that form the patient perspective of SSc-RP, through understanding of the patient experience, to inform PROMs assessing the severity and impact of SSc-RP (159). The full dataset (six focus group transcripts) from the original study was re-analysed for the current study.

4.4.1 Participants and data collection

Forty participants (34 women, mean (SD) age 56.6 (13.4) years, mean (SD) disease duration 10.5 (9.1) years, with SSc-RP took part in one of six face-to-face focus groups: two in Bath, UK, one in Pittsburgh, USA, and three in New Orleans, USA. Participants were purposively sampled to ensure diversity in disease type (limited cutaneous SSc and diffuse cutaneous SSc), disease duration, and ethnicity, with White, African American, and Hispanic representation. The number of participants per focus group ranged from six to nine, and each focus group lasted approximately 60 minutes. The topic guide included questions on SSc-RP impact, self-management, and expectations from treatments, and was reported upon previously (159).

4.4.2 Data analysis

Qualitative data analysis methods were summarised in Chapter 2 (Table 2.5). Qualitative content analysis was employed for this study to identify key concepts in the data relating to foot-specific symptoms and impact, aligning to the research aim of identifying potential domains of importance. There are three main types of qualitative content analysis: conventional, summative, and directive. A directed approach to content analysis was selected for this study. This method of analysis was selected as it generates descriptive knowledge and understanding of the phenomenon under study, using relevant research findings as guidance (312). In this case, findings from the thematic synthesis in Chapter 3 were used to guide initial coding. Directed content analysis was more suitable than conventional content analysis, which uses an inductive approach rather than predetermined categories and is most appropriate when existing literature in the research area of interest is limited. Summative content analysis, which utilises a combined quantitative and qualitative approach to explore the frequency and meaning of specific content, was also unsuitable (312).

To conduct the analysis, each anonymised transcript was read multiple times by the candidate, who then extracted data relating specifically to feet into Microsoft Excel 2022, coded this data, and grouped similar codes into themes. Two predetermined themes, based on previous literature (73, 254), were constructed prior to the analysis: pain and function. New themes were constructed for any codes that did not fit into the predetermined themes. All codes were independently checked by one of the candidate's supervisors (CAF). Finally, an expert in foot problems in people with SSc (BAP) independently read 50% of the transcripts to ensure all foot-related data were captured by the secondary analysis. The analysis team (the candidate, CAF, PR, BAP, and HJS), which included a PPI contributor, discussed the names of the themes until consensus was reached. Each theme directly related to foot-specific domains of

potential importance, and is supported by verbatim quotes from the transcripts. Pseudonyms from the primary qualitative study have been maintained (159).

In the original study (159), PPI contributors were part of the study management team. The authors reported that their involvement in decision-making ensured the study was "appropriate and suitable for achieving the project's stated aims" (159: p.7). Specifically, PPI contributors were involved in co-designing the focus group topic guide and interpreting findings. In this secondary analysis, one PPI contributor (PR) was involved in discussing findings from the perspective of a patient living with foot problems in RMDs, and naming themes.

4.5 Results

Twenty-eight participants (70%) self-reported RP in their feet. Five themes were identified from the focus group data: temperature changes, pain, cramping and stiffness, numbness, and colour changes. A thematic map, portraying the relationship between themes (domains of potential importance) and their impact, is presented in Figure 4.1.

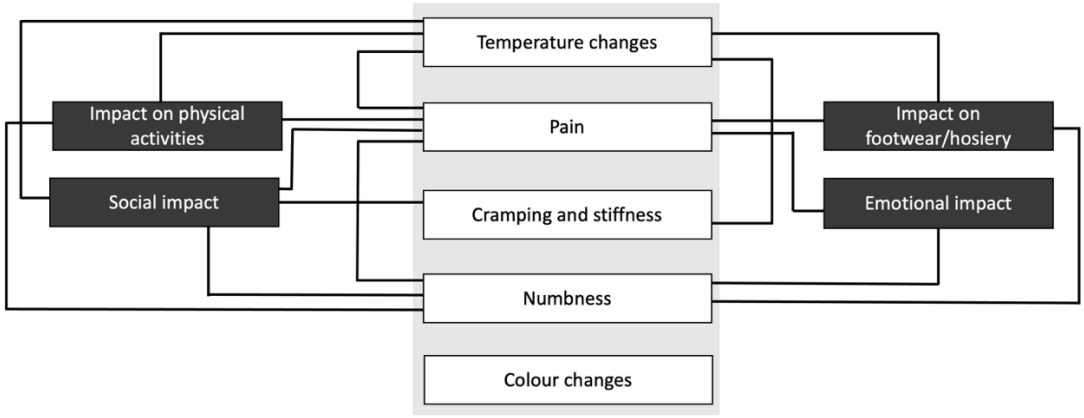


Figure 4.1: Thematic map

Figure reproduced without modification from Figure 1 in Chapman et al. (313).

Theme 1 - temperature changes

Most participants with SSc-RP experienced extreme temperature changes in their feet, particularly when in bed:

> “About a year ago it started with the feet, and so my feet are frozen my hands are cold … I love cold weather, but it just attacked my hands and my feet, and when I'm laying down in the bed, my temperature and my body is warm, extremely hot, but my hands and my feet are like ice.” – Grace, focus group (FG) 4

Participants described the life impact of these temperature changes, and how medication had helped:

> “Basically, what Raynaud's means to me is … it's debilitating at times. Temperature plays a big part in it. If I go outside, man, before I was given … prescribed a certain medication … 50 to 60° [Fahrenheit] was my cut off. If I would go outside and the temperature … it would affect my hands and my feet very, very quickly, very rapidly to where it was almost debilitating. It's not as bad as it was; medication has helped immensely, but yes, it can be downright debilitating.” – Frank, FG4

In contrast, one participant expressed that temperature changes did not improve with medication.

Participants discussed various strategies to deal with temperature changes, including foot warmers and hot water bottles. Temperature changes also impacted on hosiery and footwear choice:

> “If I go into a restaurant, I have to put on my Ugg boots whether it's in the summer or not because [of] my feet.” – Jan, FG5

> “I've been buying boots and they're lined with like a fur and they're great. If on a bad day I can't get them on I just keep my slippers on and stay home but if I can go out, I put these on and it does help, yes.” – Pam, FG3.
> “And I pre-empt it, so I don't go out of the house if it's particularly damp without a pair of gloves, or without thick socks on to keep my body temperature up.” Clive, FG1

Theme 2 - pain

Most participants with SSc-RP in their feet experienced pain. In some cases, this theme linked to Theme 1; pain was directly related to change in temperature during an episode of RP:

> “My feet get cold. I can't … it's like needles in my toes. You can't walk.” Jan, FG5

In other cases, participants described residual pain following RP episodes:

> “And my feet, and now, it's gotten so bad ‘til I wear two to three pair of socks now, because I can't deal with the cold. It used to happen mostly at night, but now it happens all during the day while I'm working and everything, so I finally just started putting extra socks on, and I get comfort. Now at night, when it's time to get into bed, I somehow can't sleep with socks, so I just pretty much wiggle my feet until I go to sleep, but the pain is just there.” – Grace, FG4

Residual pain after recurrent RP attacks also influenced patients’ choices in hosiery and footwear:

> “I'll put socks on, slippers on indoors and then it starts so I think I just want to walk up and down the hallway. I can't go out in the garden because it hurts too much to put shoes on … I bought silk socks, they're not very thick but I tell you what when I wear them, and I put a pair of ordinary socks over the top I don't get so much pain in my toes. I have to put them on first, but they were really expensive for a pair of socks … I think when my feet were pushing in and got holes in them, and I didn't dare stitch them because the edging would cause more pain. Any pressure on anything makes it worse.” – Pam, FG3

> “I can go to bed, and I can put socks on like [another participant] used to say, and I can wake up in the middle of the night or most of the night and my feet are cold and my hands. Or I can go the other way that my feet are hot but I'm tight and I get a warm feeling coming up, and I've got to take the socks off. And I've got to throw the duvet off my leg because I can't stand the pressure on it.” – Julie, FG1

One participant feared getting ulcers on her toes after having them on her fingers:

> “That's my fear of getting ulcers back, and on my toes as well, yes.” Pam FG3

Foot pain impacted on participants’ daily lives, particularly affecting the ability to walk:

> “Sometimes I can't get out the chair to walk because it's, you know, it's painful.” – Emily, FG1

> “My feet, I can hardly walk about indoors with my feet they hurt so much.” – Pam, FG3

Foot pain could also be related to skin lesions, potentially as a result of the overall systemic condition rather than RP specifically:

> “On my heels, and I've got a deep one [a crack] at the moment, my right heel has split open … it cracks and it bleeds … I know my scleroderma has been flared up lately and I have had Raynaud's and I do get numbness here. Even if they haven't gone white, they've been hurting on the sides. But that one, I got a deep cut in my foot, it’s split open, and it's painful.” – Julia, FG1

Theme 3 - cramping and stiffness

Some participants with SSc-RP in the feet described cramping and stiffness:

> “Cramp doesn't do my thumbs either, it's just my fingers and my toes.” – Emily, FG1

> “Funnily enough I don't get it during the winter, mine starts when it starts to get warmer. That's when my muscles start going into spasm and I get the cramp.” – Laura, FG3

Cramping and stiffness in the feet were often directly linked to temperature change:

> “When my toes get really cold, if I'm lying in a bed or something, my toes will stay in this surface place, like it'll get stuck.” – Betty, FG5

> “My toes do that. If I'm very cold they will lock up, like, stiff.” – Jan, FG5

Consistent with Theme 2, cramping also led to foot pain. In some cases, cramping could occur when the feet were warm, independently of a RP episode:

> “Cramp in my toes at night usually … [when my feet] are quite warm … you have to get out of bed and walk about to get the pain to go away.” – Ali, FG5

Theme 4 - numbness

Many participants with SSc-RP in their feet described numbness, which caused discomfort and irritation:

> “They get very numb. It's not exactly pain, it's like that really, really deep numbness, like it almost hurts, but to me it's not exactly the same as pain. And then as it comes back it is uncomfortable. But like joyfully so, because you know it's getting better.” – Michelle, FG2

> “I'm constantly feeling tingling and numbness in my feet … It's there right now, you know? I'll be wiggling my toes, you know, because the numbness, it's just really aggravating so I kind of try to tune it out … I just kind of deal with it and pray on it and, and just ignore it ‘cause it's I been tried so many things and nothing really helps the, the toes, you know. Maybe a little less but it always be, you know, really… And then when I try to …when I have a tight sock on, I always have to loosen it up or take it off or something like that.” – Dawn, FG5

> “The left foot stays numb most of the time. It's really bad at night … the foot, it's like, it goes to sleep as soon as you lay down and … it don't seem to be getting better.” – Carl, FG4

Numbness in the feet also affected daily tasks such as walking and driving, causing fear for some participants:

> “I think the pain would be bad enough, but my feet do go numb and it's frightening sometimes if you're trying to stand up and there's no feeling in the bottom of your foot. I've had that quite a few times, even in the night when I've got up and you go to put your foot down and you can't even feel your foot.” – Julia, FG1

> “I was driving, and my foot literally slipped off the pedal driving and [I] didn't know. That made me say, you know what, something is going on. I need to go and get this checked but it was fingers and toes completely numb and I felt like my entire body was going numb.” – Jan, FG5

> “It got to where I couldn't … I couldn't walk. I could only move maybe 20, 30 feet.” – Carl, FG4

Numbness and temperature changes in the feet could lead to a stinging sensation:

> “My foot when it gets cold and numb, it stings a whole lot. So I just rub it until … trying to relieve it, you know.” – Ella, FG4

One participant described needing to use crutches due to numbness in the feet:

> “At night my feet can get absolutely numb, and it is that recovery when I get into bed or put a thing over them, it's that I don't like, is the feeling of coming back. And then I do struggle to put my weight on them. Whereas when they're completely numb I use my crutches though.” – Ann, FG1

Theme 5 - colour changes

Participants described colour changes in their feet. Presentations in the toes included white, purple, and black, although the phases were not always clearly defined:

> “They start out purple and white. I've noticed my toes more white. I didn't think my toes did that before, but I've noticed that more now.” – Heidi, FG2

> “I find with myself if I get cold, I get purple lips, my hands, my toes and even my knees go purple.” – Mag, FG3

> “Mine go black …very, very dark.” – Liz, FG1

4.6 Discussion

This secondary analysis of qualitative data explored foot symptoms, and the impact of these symptoms, in patients with SSc-RP, to identify which foot-related domains are potentially important to these patients.

4.6.1 Main findings

Consistent with previous research exploring hand involvement in SSc-RP (65), pain, numbness and temperature changes in the feet had significant impact on the daily lives of participants. Whilst colour changes in the feet were described by participants in this

study, this issue appears to have limited impact. In contrast, colour changes, and digital ulceration in the fingers can cause unwanted attention and embarrassment, particularly at work and in public, and in some cases lead to dissatisfaction with body image (159). Change in appearance of the feet and ankles is an important symptom to patients with other RMDs (254). As discolouration of the toes due to RP is temporary, and feet are often hidden by socks and shoes, change in appearance of the feet may not be as important to patients with SSc-RP as it is to patients with other RMDs (e.g. RA and OA), where foot deformities can be extensive, permanent, and may lead to drastic changes in choices of footwear and clothing (272, 314).

The experiences and impact of pain in the feet as a result of SSc-RP, a key finding from this study, is congruent with the experiences of patients with foot and ankle problems in other RMDs, including RA, PsA, OA, gout and SLE, as identified in the qualitative synthesis in Chapter 3. Foot pain is the dominant symptom across these conditions and an important domain to measure in future research. Similar to other RMDs, the target domains within the broad domain of pain in SSc-RP included pain with weightbearing, at rest, and at night. Findings also indicate that pain can be transient, occurring during an episode of RP, or can be residual. Residual pain after recurrent RP episodes is associated with nerve sensitivity and chronic allodynia (315), such as not being able to tolerate the pressure of the duvet at night. Findings also highlighted the impact of cramping and stiffness in the feet on patients with SSc-RP, and the effect of SSc-RP on footwear and clothing choices. This suggests that, in contrast to SSc-RP affecting the hands, some domains of importance to patients are more foot-specific than disease-specific.

Findings from the current study were integrated into the list of candidate domains (Appendix D), to compare domains across this study, the scoping review, and the qualitative synthesis. Foot pain, stiffness, and subsequent functional limitations, footwear impact, and emotional impact have been identified as potentially important domains to patients with foot and ankle problems in other RMDs, in the qualitative synthesis (Chapter 3) and measured in the previous scoping review (73). This indicates that SSc-RP is in line with the proposed scope of COMFORT. In contrast, temperature changes, numbness, colour changes, and cramping were not previously identified in qualitative research or measured in existing studies, although these domains did overlap with pain among participants with SSc-RP. These domains will be explored further in primary qualitative research with patients with different RMDs (Chapter 5), to identify whether they are SSc-RP specific domains or important in multiple conditions.

Overall, the current study indicates that there are some common domains of potential importance in SSc-RP and other RMDs, suggesting that the scope of a COS for foot

and ankle disorders in RMDs should give some consideration to SSc-RP. There is a clear overlap between some foot-related domains and the domains included in the core domain set for SSc-RP, e.g. pain due to RP, impact of RP on function, and impact of daily life/adaptation. The current study also identified a foot-specific domain, footwear impact, that is potentially important to patients with SSc-RP but is not covered within the SSc-RP core domain set. It is possible, therefore, that future studies that are specifically investigating foot and ankle interventions should use the foot and ankle core domain set, but give consideration to any additional domains in the SSc-RP core domain set. Conversely, future studies that are investigating the effectiveness of treatments for SSc-RP in general (e.g. drug trials), should use the SSc-RP core domain set, but give consideration to any additional domains in the foot and ankle core domain set. As the current study identified that 70% of participants in the original qualitative focus groups for this COS had foot symptoms that were overlooked, the core domains severity of attacks of RP, duration of attacks of RP, and frequency of attacks of RP should also give consideration to the foot. Domains within a COS are intended to be measured in all relevant studies (63, 65). However, use of the COS in future research would not preclude measurement of additional domains of interest to researchers. The use of additional domains is expected and will depend on the specific RMD, study question, or intervention being tested. Further work is needed to understand how best to harmonise both core domain sets, and overlapping COS will be further considered in Chapter 7.

4.6.2 Strengths and limitations

This study utilised an existing dataset to address a gap in existing foot and ankle RMD research. It was strengthened by independent coders, including an expert in foot and ankle problems in people with SSc who facilitated interpretation of the data based on clinical experience. Additionally, the study benefited from PPI input into discussions during analysis, ensuring that findings were grounded in the patient perspective.

The study also has several limitations. Firstly, this was a secondary analysis of an existing dataset addressing a research aim relating to the development of OMIs for SSc-RP; therefore, none of the questions in the focus group topic guide related specifically to experiences of foot problems. Previous research has shown that patients with foot problems may not be forthcoming in talking about their feet, sometimes viewing them as a separate issue to the rest of their condition and reporting that they are ignored during routine consultations (266, 279). Consequently, not all aspects of living with foot problems in SSc-RP may have been discussed in depth. Although

COMET highlights that secondary analysis of a qualitative dataset can be carried out during the COS development process (65), the broader methodological limitations of conducting a secondary analysis must also be considered. As the candidate did not conduct the focus groups, contextual data were limited. It was not possible to build rapport, observe non-verbal cues (such as emotions, or the body language participants used when talking about foot problems), or explore any points of interest further. This limited the depth and richness of the analysis. Despite this, novel perspectives were identified in the data and these initial findings can be built upon in future primary qualitative studies. The potential for researcher bias during data analysis in this study must also be acknowledged, as three podiatrists were involved in interpreting the data. Offsetting this, the analysis team also included a PPI contributor and an experienced qualitative researcher with limited experience of foot and ankle disorders, who independently checked all coding decisions.

Finally, in the context of developing a COS for foot and ankle disorders in RMDs, this study was limited to RP, which is only one aspect of disease in SSc. Some symptoms described by participants may have been due to impaired circulation, skin, musculoskeletal or peripheral nerve involvement in SSc rather than specifically due to RP (e.g. cramping, stiffness, and pain).

4.6.3 Implications

4.6.3.1 Implications for clinical practice

This study has implications for clinical practice; findings indicate that RP is common in the feet, similar to other studies reporting prevalence (46, 305, 306), and that RP in the feet can have differing manifestations and impacts multiple areas of patients' lives. Current national guidelines for SSc management were developed with podiatry input and detail pharmacological treatment recommendations for SSc-RP, but do not provide any foot-specific recommendations (309). Similarly, the 2023 EULAR guidelines for SSc do not consider foot problems (316). This study indicates that foot involvement should be addressed within the overall rheumatological management of patients with SSc, with footwear and hosiery advice provided as necessary. The importance of establishing patients' priorities and subsequently the outcomes to target with treatments is an important aspect of medical care.

4.6.3.2 Implications for research

By exploring patients' experiences of foot problems in a RMD that was not represented in the previous literature, this study has identified domains of potential importance that were not previously captured, including temperature changes, numbness, colour changes, and cramping. Further exploration of these domains among patients with other RMDs is needed to establish their relevance within the context of a COS. Further work is also needed to understand the full breadth and depth of foot-specific domains that are important to patients with SSc, to address gaps in existing qualitative research through exploration of patients' experiences of living with this condition.

4.7 Conclusion

This study achieved its aim of identifying foot-related domains that are potentially important to patients with SSc-RP, by exploring the experiences of patients with this condition. Key domains of importance were identified, including pain and footwear impact. SSc-RP commonly involves the feet, presents in varied patterns, and affects multiple aspects of patients' lives. Findings highlight the need for foot-specific interventions, and indicate where future foot-specific interventions for RP could be targeted. Findings from this study improve understanding of the domains that are important to patients with SSc-RP affecting the feet, contributing to the development of the COS for foot and ankle disorders in RMDs. The next chapter will explore the experiences of patients with foot problems relating to a broad range of RMDs, including SSc and other under-represented RMDs in more depth, through primary qualitative interviews.

Chapter 5 Qualitative interview study

5.1 Introduction

Following a thorough exploration of the literature in Chapters 3 and 4, the following chapter reports an international qualitative interview study to better understand the priorities of patients. This study builds on the synthesis of existing qualitative studies (Chapter 3), addressing gaps in the literature. It also explores the new domains identified in the secondary analysis of focus groups with patients with SSc-RP (Chapter 4) among patients with a wider range of RMDs. Findings have been published as:

> **Chapman LS**, Flurey CA, Richards P, Redmond AC, Soliman E, Moshrif A, Malone L, Joyce C, Arnold JB, Golightly YM, Hofstetter C, Helliwell PS, Menz HB, Hannan MT, Rahman MN, Shea BJ, Smith TO, Siddle HJ. What outcomes are important to people with foot and ankle disorders in rheumatic and musculoskeletal diseases? An OMERACT qualitative interview study across four continents. Semin Arthritis Rheum. 2025 Jun;72:152671. doi: https://10.1016/j.semarthrit.2025.152671

Chapter 3 established that existing qualitative studies have predominantly focused on the impact of foot and ankle disorders in patients with RA and OA in high-income countries, of White or unreported race/ethnicity, and whose first language is English (254). As outlined in Chapter 2, to facilitate synthesis of all research in a specific area of health, a COS should be applicable and adopted across relevant settings and disciplines. Therefore, it is important to understand and compare domain priorities among patients in different geographic locations, with a range of RMDs, first languages and races/ethnicities.

5.2 Aim and objectives

The aim of this study was to build on the literature and establish, through primary research, which domains are important to people living with RMDs who had sought advice and treatment for foot and ankle disorders, to inform a Delphi consensus study (Chapter 6).

The objectives were to:

- Examine the range and scope of domains experienced by patients with a variety of RMDs, through an exploration of their experiences;
- Understand and compare the importance of the identified domains across a range of RMDs, countries, and races/ethnicities;
- Explore patients' choice of words when discussing domains.

5.3 Epistemology

Key epistemological approaches are discussed in Chapter 2 (section 2.5.2). To recap, key approaches include positivism, interpretivism, constructivism, and pragmatism. This qualitative study was conducted from a pragmatic epistemological standpoint. Pragmatism focuses on using the approaches that are most suitable to address the research question, prioritising knowledge that can be applied to real world situations (122), in this case, the development of a COS.

5.4 Methods

Ethical approval for this study was obtained from the Health Research Authority (HRA) (reference 22/NE/0226) and University of Leeds School of Medicine Research Ethics Committee (reference MEC-22-071). Local ethical approval was also obtained at each clinical recruitment site. The HRA ethical approval document is presented in Appendix F. This study was reported in-line with the COREQ framework (168). Reporting guidelines for qualitative studies, including their advantages and limitations, are discussed in Chapter 2 (section 2.5.2.1). Despite its limitations, the COREQ framework was deemed appropriate for this study as it primarily focuses on research involving interviews and focus groups, as opposed to other, broader qualitative reporting guidelines such as the SRQR (167).

The benefits of primary qualitative research as part of the COS development process have already been discussed. In summary, primary qualitative research permits more relevant and specific data collection relating to meaningful domains, with less reliance on interpretation compared to a synthesis of existing qualitative studies or a secondary analysis. The initial intention in this study was to conduct semi-structured interviews and focus groups. The advantages and disadvantages of both were discussed in depth in Chapter 2 (section 2.5.2.2). In summary, interviews allow for the collection of in-depth responses, whilst focus groups allow for the collection of multiple perspectives in a short amount of time (121, 150). All participants who were invited to attend a focus

group were offered an individual interview if they were unwilling or unable to attend a focus group. Due to multiple focus group participant cancellations, a decision was made to conduct individual, semi-structured interviews only.

5.4.1 Rigour

An overview of rigour in qualitative research was presented in Chapter 2. Credibility, transferability, dependability, and confirmability were considered when designing this study. Examples of how rigour was ensured in the current study are provided throughout the Methods section. The supplementary material of the associated publication also contains a completed COREQ checklist (317: supplementary file 1).

5.4.2 Participants

Adults (≥ age 18) with RMDs (IA, OA, crystal arthropathies, CTDs, and MSK conditions in the absence of systemic disease) who were receiving or had received treatment for foot and ankle disorders within the last 12 months were eligible to participate, in accordance with the predetermined scope of the intended COS (as defined in Chapter 2, section 2.9.1). Inclusion and exclusion criteria are presented in Table 5.1.

Table 5.1: Participant inclusion and exclusion criteria

Inclusion criteria
Aged 18 or over
Diagnosis of a RMD and have received treatment (conservative, pharmacological, or surgical) for a foot and/or ankle disorder within the last 12 months
Able to give informed consent
Exclusion criteria
Acute trauma or injury to the foot/ankle (e.g. fracture, rupture, sprain), or a sports injury
Comorbidities affecting the foot/ankle (e.g. diabetes, primary neurological conditions, or peripheral arterial disease)
Lacking capacity to give informed consent

5.4.3 Sampling

Various sampling techniques can be adopted in qualitative research; an overview of the most common techniques were presented in Chapter 2 and include purposive sampling, convenience sampling, snowball sampling, and theoretical sampling. Selection of a sampling technique depends on the context, aims, and research questions being explored in a qualitative study (158). In the current study, gaining perspectives from a diverse range of participants was important to ensure domains within the subsequent Delphi consensus study and final COS were relevant and meaningful to patients. As purposive sampling involves selecting participants with different characteristics, who provide diverse perspectives, it is a compatible sampling technique with the qualitative stage of COS development (65). Purposive sampling can also improve transferability (171). In contrast, convenience sampling is the least rigorous sampling technique and may result in low credibility and lack of transferability (172). The purposive sampling criteria in this study were based on condition (RMD type, duration of foot/ankle symptoms), geographic location (by continent), and demographic characteristics (age, sex, first language, race/ethnicity).

5.4.4 Recruitment

Patients were recruited through clinical departments via local facilitators in the UK (clinics at four NHS Trusts: Leeds Teaching Hospitals NHS Trust, Midlands Partnership Foundation Trust, Homerton Healthcare NHS Foundation Trust, Harrogate and District NHS Foundation Trust), Australia (private podiatry and university podiatry clinics) and Egypt (hospital clinics), and through OMERACT electronic mailing lists reaching multiple countries. The NHS Trusts included as recruiting sites represented acute and community podiatry and rheumatology services. Potential participants recruited through clinical departments were informed about the study, screened against the eligibility criteria by the local facilitator, and asked to complete a consent to contact form (Appendix G). Those who consented to be contacted about the study were given a recruitment pack consisting of an invitation letter, PIS and consent form (Appendix H). Following identification of an eligible English-speaking participant from a clinical site who consented to be contacted, the candidate made contact by telephone or email to discuss the study. Potential participants recruited through OMERACT mailing lists were provided with an email invitation and invited to contact the candidate directly for more information. These potential participants were then sent a recruitment pack via email and invited to contact the candidate again if they would like to proceed with the study. Further information was collected from all potential participants who expressed an interest in participating, including type of RMD, duration of foot/ankle symptoms,

geographic location, first language, race/ethnicity, gender, and age, to ensure the purposive sampling criteria were addressed. The invitation letter, consent form, and participant information sheets (PIS) used in the study were developed with the research team, including PPI contributors.

Potential participants in Egypt who did not speak English were approached by one of two members of the wider OMERACT Foot and Ankle Working Group based in Egypt (ES, AM), and provided with the same information about the study. All written patient-facing documents for the study (invitation letter, consent form, PIS), in addition to the interview schedule, were translated into Arabic; the wording was discussed with ES and AM prior to written translation, to ensure cultural sensitivity. Documents were translated by a university-approved third party translation service (dictate2us) by translators with bilingual competence and medical expertise, and independently back-translated to check for consistency and accuracy. All translated documents were then checked and approved by ES and AM. Potential participants in Egypt followed the same recruitment process with ES and AM (there was no direct contact with the candidate).

5.4.5 Sample size

Sample sizes in qualitative research studies are highly variable (318). They are usually less than those in quantitative research due to differing purposes; qualitative research typically aims to collect rich, descriptive data to understand experiences and perspectives*,* rather than to test a hypothesis or measure/quantify variables. Whilst quantitative sample sizes are guided by statistical power and generalisability, this is not the case in qualitative research. Although COREQ includes the need for data saturation, this is now outdated (181, 319). Adequate qualitative sample sizes are now guided by information power, which focuses on the extent to which the data contributes to answering a specific research question or advancing the understanding of a phenomenon, rather than reaching an endpoint of knowledge (157). The size of a sample with sufficient information power can depend on the aim of the study, sample specificity, use of established theory, the quality of dialogue, and the analysis method (157). As outlined in Chapter 2, OMERACT recommend a minimum of 30 participants for the qualitative stage of COS development (63). Sample sizes in OMERACT qualitative studies varied considerably; for example, the development of the hip and knee OA COS included international focus groups with 35 patients (320), whilst the updated COS for PsA involved a total of 24 focus groups with 130 patients (233). In comparison, interviews with 28 patients were carried out as part of both the patient

outcomes in longitudinal studies for RA COS (321) and the COS development for calcium pyrophosphate deposition (322). The broad purposive sampling criteria for this study was outlined in section 5.4.3. As the study sought to explore the perspectives of participants across multiple characteristics, a pragmatic approach to recruitment was undertaken to explore as many of the purposive sampling criteria as possible within the timeframe of the study. The intention, based on this timeframe and on resources, was to include 50-60 participants (87).

5.4.6 Data collection

The medium of data collection can affect the quality of qualitative data. A key objective of this overall project was international agreement on a COS, thus inclusion of geographically diverse participants was important. The option for online and telephone interviews can facilitate geographical diversity, whilst also aiding inclusion of participants with poor mobility and those in remote areas, where face-to-face interviews may be inconvenient (323). Disadvantages of online qualitative research have been recognised, however; some participants may be unfamiliar with videoconferencing tools, and may not have a stable internet connection or appropriate device. This can exclude certain demographics, particularly older populations or those in areas with limited digital infrastructure and low levels of digital literacy (324, 325). Additionally, video conferencing platforms can have security issues (326). However, online and telephone interviews can lead to participants feeling more at ease in their own environment and able to share more, compared to feeling uncomfortable and intimidated in a face-to-face setting, resulting in more open and richer responses to interview questions, especially when discussing sensitive topics (153). This may be augmented further in telephone interviews, where there is added anonymity. In contrast, it is sometimes easier to build rapport and trust in face-to-face interviews, which can lead to richer data (153). Distractions in the surroundings are a greater risk in online and telephone interviews, where the researcher has less control over the participant's environment. Additionally, face-to-face interviews can reveal non-verbal cues such as body language and facial expressions, which can enhance the depth of a participant's response, e.g. through an understanding of their emotional state, with greater potential to interpret emotions, attitudes, and nuances of responses. The interviewer can therefore respond to verbal and non-verbal cues, offering more empathetic or supportive reactions, which can encourage participants to be more open and reflective. A major drawback of interviews and focus groups that are not face-to-face, particularly telephone interviews, is that the researcher might not be able to respond to signs of distress (153). Interviews in the current study were held online

(Microsoft Teams or Zoom), by telephone, or face-to-face (UK only, in a quiet room within a hospital setting), depending on each participant's preference. This decision recognised that there are specific benefits and drawbacks to all three modes of data collection and aimed to be as inclusive as possible whilst ensuring geographical diversity in the sample. Several strategies were employed to address the potential drawbacks of telephone and online interviews; these are presented in Table 5.2.

Table 5.2: Strategies for online interviews

Risk	**Mitigation**
Lack of rapport	The first 5-10 minutes of each interview were used to talk about topics unrelated to the interview and were not recorded, in an attempt to build rapport.
Technology issues	Participants' confidence levels with technology were discussed, and participants were offered a practise call prior to the interview.
Security issues	Only university-approved platforms (Microsoft Teams and Zoom) were used.
Lack of non-verbal cues	Participants were reassured that they could decline to answer any questions they did not feel comfortable answering, and could take a break or stop the interview at any time.
Confidentiality	Participants were asked to find a quiet space for their interviews, and the candidate conducted the interviews in private office.

Interviews were conducted between October 2022 and March 2024. With the exception of those carried out with participants based in Egypt, all interviews were conducted in English by the candidate. In interviews conducted in English, attendance of an independent interpreter was offered to participants whose first language was not English, to aid participation. Interviews with participants in Egypt were conducted in Arabic by one of two researchers (ES, female, or AM, male, both rheumatologists). The candidate and the two interviewers met before interviews in Egypt commenced, to discuss the interview schedule, ensuring it was culturally appropriate, and to practise an open-ended approach to questioning.

In all cases, participants were unknown to the interviewer. This can improve credibility in qualitative research, as it reduces biased questioning and interpretation of answers

based on preconceptions or expectations (164). In addition, it reduces the potential for power imbalance, and participants may feel more comfortable and willing to share their true experiences when they are unfamiliar with the interviewer rather than feeling pressured to please them, particularly when talking about their experiences of healthcare and treatments (327).

An interview schedule (Appendix I) was developed based on existing literature, input from PPI contributors, and discussions with multidisciplinary HCPs and researchers. The interview schedule was piloted with a PPI contributor (PR) during a full length practise interview and amended based on feedback. Examples of changes to the interview schedule as a result of PPI input included the addition of questions about the impact of foot and ankle disorders on voluntary work and, if applicable, work carried out before retirement. Both of these aspects of occupational function would have otherwise been overlooked by the candidate. Additionally, the following questions were added into the interview schedule based following piloting: '*is there anything you wish your team had considered more when treating you for your foot/ankle problems?*'; '*is there anything you feel that* you *should have considered more?'; 'what has been successful about your treatments?'; 'is there a reason why they weren't successful?'* These questions were co-produced with the PPI contributor with the aim of gaining richer data by helping participants to focus on their own experiences of healthcare, in addition to answering more abstract questions relating to what treatments for foot and ankle disorders should achieve.

The interview schedule was discussed with members of the research team based in Egypt (ES, AM) during online meetings and by email before and after written translation into Arabic, to ensure cultural sensitivity. The interview schedule was iteratively modified throughout data collection, as any new concepts raised by participants were identified. The interview duration ranged from 16-75 minutes. Interviews were audio recorded on a digital voice recorder.

Field notes were made during and after each interview to supplement the audio recordings and transcripts. Field notes facilitate the collection of rich contextual information, such as the participant's environment, body language, and facial expressions that may not be apparent through audio recordings and transcripts (328: p.71). They allow the interviewer to note their thoughts and feelings, including any biases, emotional responses, or dynamics, which can enhance credibility by encouraging reflexivity They also allow the interviewer to document any patterns in the data or issues with questions for modification before further interviews.

5.4.7 Data analysis

Commonly used qualitative analysis methods were presented in Chapter 2 (Table 2.5). Interview data were organised using NVivo (Lumivero, USA) and analysed using primarily a deductive (bottom-up, theory-driven) approach to the framework method (329). The deductive approach allowed code and category names to be refined based on participants' language. However, an inductive (top-down, data-driven) approach was used to identify any new codes. This allowed new codes that did not appear in the deductive coding framework to be identified. The framework method was developed in the 1980s and is widely used in health research (329). Its purpose is to identify, describe, and interpret key patterns within and across cases relating to the phenomenon of interest (329). Alternative methods of qualitative analysis were considered, including IPA, grounded theory and thematic analysis. As this study sought to understand a broad range of patient perspectives rather than an in-depth lived experience, IPA was not appropriate. Grounded theory aims to develop concepts or theories based on what exists within the study data, which was not the aim of the current study, and the qualitative interviews built on existing knowledge from a qualitative synthesis of existing studies, so grounded theory's fully inductive approach was also unsuitable (330). A focus of the current study was to identify domains that should be included in a future Delphi consensus study, whilst taking into consideration domains that had already been identified in previous studies. Thematic analysis was initially planned for the primary qualitative research within the COMFORT study, as stated in the published protocol (87). However, whilst both framework method and thematic analysis allow for deductive coding, framework method gives a clearer matrix-based structure to allow comparison across participants and themes (329). Framework method was therefore selected for this study as it allowed domains identified in the data to be systematically compared and categorised. This method aligns with the epistemological approach of pragmatism as it is not bound by theoretical frameworks and focuses on providing practical solutions (331). Framework method consists of seven key steps. These steps, and details relating to how they were followed in the current study, will now be outlined.

Transcription

To facilitate immersion in the data, two audio recordings were transcribed by the candidate. All other English audio recordings were transcribed verbatim by a third-party transcription service (1st Class Secretarial Services). The accuracy of each transcript was verified by the researcher and corrections were made where necessary. Arabic audio recordings were simultaneously translated and transcribed verbatim by third

party translators (dictate2us) with bilingual competence and medical expertise, independently back-translated to check for consistency and accuracy, and checked and approved by ES and AM. One approach that can enhance credibility in qualitative research is member checking, also known as participant validation, which involves asking participants to give feedback on the accuracy and adequacy of transcripts and/or the analysis undertaken (328, p.282). Member checking was not considered appropriate for this study due to the extra time and resources required. However, following completion of an interview, two participants asked to view their interview transcript. There was no expectation for participants to provide feedback, but both confirmed transcript accuracy.

Familiarisation

Familiarity with each interview was achieved by reading and re-reading the transcript alongside field notes and reflective logs from the interview. Corresponding audio recordings were re-listened to where necessary. Field notes, thoughts and impressions were recorded on each transcript.

Coding

Coding was carried out to classify all data so that it could be compared systematically with other parts of the dataset. For the deductive stage of coding, each transcript was read line by line, and predefined codes were applied. Predefined codes were the domains identified previously (Appendix D). Codes were categorised as broad domains and target domains, according to OMERACT definitions (63). The broad domain of activity and participation was categorised into three target domains (obligatory, committed, discretionary), as outlined in Table 5.3 (332).

Table 5.3: Activities and participation domain definitions

Type of activity	Definition
Obligatory	Activities required for survival and self-sufficiency, including hygiene and self-care, walking, and driving.
Committed	Activities associated with principal productive social roles, including paid work, household responsibilities, and caring responsibilities.
Discretionary	Participating in religious and spiritual activities, socialising, exercise, engaging in leisure time activities and pastimes, and pursuing hobbies.

Table modified from Verbrugge and Jette (332).

If none of the predefined codes captured what the candidate interpreted as important, a new code was applied. Field notes were referred to throughout the coding process. For confirmability, two other members of the research team (CAF, senior qualitative rheumatology researcher; PR, PPI contributor) independently coded three transcripts. These additional perspectives helped to ensure that findings reflected participants' true experiences, rather than researcher bias. Input from PPI at this stage also challenged the candidate's assumptions and promoted reflection on how participants' experiences might be understood differently, compared to from the perspective of a podiatrist. For further confirmability, alternative perspectives were also sought from two HCPs (LM and CJ, both podiatrists) with limited knowledge of COS development and OMERACT, who coded six transcripts each as part of an independent research exercise.

Developing a working analytical framework

All independent coding decisions were compared to the candidates own decisions and differences in interpretation were discussed until consensus was reached. The final set of codes was then applied to all subsequent transcripts. Codes were grouped together into clearly defined categories, forming a working analytical framework. Predefined codes were arranged by core areas according to the OMERACT Filter: manifestations (signs, symptoms, biomarkers), life impact, and societal/resource use (333).

Applying the analytical framework

The working analytical framework was then applied by indexing subsequent transcripts using the existing categories and codes. NVivo was utilised to store and organise the data.

Charting data into the framework matrix

A matrix was generated in Microsoft Word and the data were charted into this matrix. Charting involved summarising the data by category from each transcript. The chart included references to relevant verbatim quotations.

Interpreting the data

Impressions, ideas and early interpretations of the data were noted throughout the analysis. These were discussed with the wider research team and PPI contributors during regular analysis meetings. This process of peer debriefing was employed to ensure credibility. Characteristics of and differences between the data were identified, including connections between categories and theoretical concepts (predefined domains, new domains and other new concepts within the data). Domains and concepts identified during this study were added to the list of domains for potential inclusion in a Delphi study, as presented in Appendix D.

5.4.8 Ethical considerations

The risks involved in this study were discussed with the research team and deemed to be minimal. Participating in an interview can involve emotional issues for the participant, and/or the interviewer may detect ongoing issues related to the participant's medical condition or emotional wellbeing. The intention was to discuss any significant issues raised during the interviews with the participant outside of the interview, and, if appropriate, refer the participant to an appropriate professional, such as their GP. No significant issues were raised during the study. The inconvenience related to participant time was considered in the methods and minimised where possible. For example, participants were offered reimbursement for travel expenses of up to 50 GBP to attend a face-to-face interview, and payment for childcare/carer cover or personal assistants was offered to ensure inclusion of participants who would otherwise not be

able to take part. A rate of 30 GBP per participant per hour was funded for childcare/carer/personal assistant costs in this study. Interviews were offered and conducted at times most convenient for participants, to minimise burden. Additionally, gift cards were offered for study participation.

Before being enrolled in the study, informed consent was gained for all participants. Potential participants were given at least 24 hours to consider information about the study before being contacted to discuss study participation. If the potential participant did not respond within one week, the researcher attempted to make contact by telephone and email again. There were no further contact attempts after this. The nature and objectives of the study were discussed, including how the participant's information would be used, and participants were given the opportunity to ask any questions. In addition, potential participants were encouraged to spend as much time as they need discussing the study with disinterested third parties (e.g. their family). Potential participants were given a maximum of one month to consider participation. Prior to the interviews, each participant was asked to provide informed consent electronically (for online or telephone interviews), or to complete a paper consent form (face-to-face interviews). Before online or telephone interviews, an identical electronic consent form hosted by Online Surveys, an established and secure system used in other studies at the University of Leeds, was administered via email. A copy of consent was emailed to participants. The translation and transcription services utilised in this research were approved by the University of Leeds and bound to a confidentiality agreement. Prior to commencing each interview, verbal consent was also obtained. A data management plan for this study, adhering to the standardised University of Leeds data management plan template, is presented in Appendix J.

5.4.9 Reflexivity

Reflexivity is an important consideration in qualitative research, and key to enhancing rigour (334). It can be defined as a set of continuous, collaborative, and multifaceted practices through which researchers self-consciously critique, appraise, and evaluate how their subjectivity and context influence the research processes (335). The candidate conducted the interviews during her Health Education England/NIHR Clinical Doctoral Research Fellowship. The candidate is a podiatrist with undergraduate and postgraduate degrees in podiatry and psychology, who had previous experience of conducting qualitative interviews. The candidate discussed the study via telephone or email with all participants prior to the interview. Five participants were patients at the hospital where the candidate was based in her clinical role, whilst eight participants

were patients at the hospital where the candidate was based in her fellowship role. No participants received clinical care from the candidate. All participants were aware that the candidate was a podiatrist undertaking a PhD with the ultimate aim of improving the quality of evidence for treatments for foot and ankle problems in RMDs.

Prior to the interview, the candidate explained her clinical and research interests and reasons for conducting the study. The researcher's dual position as a podiatrist and researcher was an important consideration. This dual position may enhance trust with research participants and facilitate appropriate probing (336). Interviews, which took the form of one-to-one conversations, utilised the candidate's strengths of a podiatrist; seeking information about a patients' symptoms, experiences, and priorities is common practice in a podiatrist's role. However, the risks associated with the researcher being a podiatrist must not be overlooked; interviews led by HCPs can resemble clinical consultations, leading to a potential power imbalance. To address this potential issue, the candidate used prompt questions based on the interview schedule to help ensure that all relevant areas were fully explored in this study.

Being a podiatrist may result in the candidate having more preconceptions about the phenomenon being explored. The candidate is also likely to have developed preconceptions through undertaking PPI activities, planning the COMFORT project, and conducting a previous scoping review of domains (73), qualitative synthesis (Chapter 3) and secondary analysis of focus group data (Chapter 4). To help ensure that preconceptions did not compromise the candidate's interpretation of the data, regular discussions and reviews of interview technique and coding took place with supervisory team members (CAF, PR). Both have a wealth of experience in conducting qualitative research, and one (PR) is also a PPI contributor with lived experience of foot and ankle problems. Additionally, the candidate kept reflective logs (including field notes) throughout the data collection and analysis phases to support reflections on how these and additional factors influenced the study. Reflective logs included personal reflections on what was done and why, and what was felt. These were discussed with an experienced qualitative researcher (CAF) at regular points during the study and are explored more in section 5.5.2.

5.5 Results

Fifty-six patients from eight countries (across four continents) participated in an interview. It is not possible to confirm how many patients were approached about this study due to the broad recruitment strategy, which included OMERACT promotion. Four participants contacted the candidate about the study via email but were not invited

to participate in the study as they had RA, and purposive sampling criteria for RA had already been reached at this point. One participant (Black, 20-29 age bracket, UK) consented to participate and then dropped out prior to the interview due to family issues. There were no other dropouts or withdrawals and no participants declined to answer any of the interview questions. Participant characteristics are summarised in Table 5.4; these characteristics have been aggregated rather than reported for individual patients to ensure patients are not identifiable. Patients who had more than one RMD affecting the foot/ankle were asked about each RMD in turn. Forty-four interviews were held fully in English, two were held in English and aided by an interpreter, and ten were conducted in Arabic. The domain categorisation matrix is presented in Table 5.5.

Table 5.4: Participant characteristics

Gender	Women (37 participants), Men (19)
Age (years)	Median: 56.5; IQR: 22.5
Geographic location	Europe: UK (27), Republic of Ireland (1), Serbia (1), Malta (1) Africa: Egypt (10) Australia: Australia (8) North America: USA (6), Canada (2)
Race/ethnicity	White British (18), Egyptian (10), White American (8), White Australian (8), Chinese (2), White Canadian (2), White Irish (1), Maltese (1), Arab (1), Bangladeshi (1), Indian (1), Black African (1), Other African (1), Mixed (1)
First language	English (38), Arabic (11), Mandarin (2), Malayalam (1), Farsi (1), Portuguese (1), Turkish (1), Maltese (1)
RMD[a]	RA (18); SpA: AS (4), PsA (5), enteropathic arthritis (3), reactive arthritis (1); JIA: juvenile-onset RA (2), juvenile-onset PsA (1); OA (8); CTDs: SSc (5), SLE (2) mixed/undifferentiated CTD (3), dermatomyositis (1), secondary Sjogren's syndrome (2), secondary RP (12); Crystal arthropathies: gout (1), pseudogout (1); MSK disorders with no systemic condition: plantar heel pain (2), Achilles tendinopathy (3), peroneal subluxation (1), toe deformities (1), PTTD (1)
Self-reported duration of foot/ankle symptoms (years)	Median: 12; IQR: 20

AS, ankylosing spondylitis; CTD, connective tissue disease; IQR, interquartile range; JIA, juvenile idiopathic arthritis; MSK, musculoskeletal; OA, osteoarthritis; PsA, psoriatic arthritis; RA, rheumatoid arthritis; RMD, rheumatic and musculoskeletal disease; RP, Raynaud's phenomenon; SLE systemic lupus erythematosus; SSc systemic sclerosis; SpA, spondyloarthropathy.

[a]Some participants had multiple RMDs.

Table 5.5: Categorisation of domains

OMERACT core area	Broad domain	Continent representation	RMD representation	Target domains
Manifestations (signs, symptoms, biomarkers)	Pain	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Pain severity
				Pain during weightbearing
				Pain after weightbearing
				Pain at rest
				Pain on movement
				Morning pain
				Pain on pressure
				Pain at rest
				Pain at night
	Cramping	Europe; North America	RA; SpA; OA; CTD	Cramping during weightbearing
				Cramping at rest
				Cramping at night
	Numbness	Europe; Africa; North America; Australia	SpA; OA; CTD; MSK disorder	Numbness

OMERACT core area	Broad domain	Continent representation	RMD representation	Target domains
	Poor circulation	Europe; North America	RA; SpA; CTD	Poor circulation
	Fatigue	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Overall fatigue
				Foot/ankle-specific fatigue
				Cognitive fatigue
	Physical function	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Reduced range of movement
				Stiffness
				Muscle weakness, instability, and balance
				Altered gait/walking
	Deformity	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Deformity
	Swelling	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Swelling at rest Swelling after weightbearing
	Temperature changes	Europe; Africa; North America	RA; SpA; CTD; MSK disorder	Temperature changes

OMERACT core area	Broad domain	Continent representation	RMD representation	Target domains
	Skin and nail health	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Skin condition
				Nail condition
				Ulceration
				Colour changes
				Infection
Life impact	Activities and participation	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Obligatory activities/participation
				Committed activities/participation
				Discretionary activities/participation
	Psychological impact	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Changes in mood
				Fear of falling
				Perception of appearance
				Self-confidence
	Sleep	Europe; Africa; North America; Australia	RA; JIA; SpA; CTD; MSK disorder	Sleep

OMERACT core area	Broad domain	Continent representation	RMD representation	Target domains
	Footwear impact	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Footwear comfort
				Footwear aesthetics
				Footwear choice
				Impact on clothing
	Treatment experience	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Treatment experience
Resource/societal use	Healthcare utilisation	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Appointments/consultations
				Hospitalisation
				Medication use
	Personal expenses	Europe; Africa; North America; Australia	RA; JIA; SpA; OA; crystal arthropathy; CTD; MSK disorder	Appointments/consultations
				Medication use
				Work impact
				Transport costs
				Device costs

5.5.1 Framework analysis

Deductive and inductive findings are presented simultaneously according to broad domains for each OMERACT core area, with a narrative overview of participants' descriptions and experiences of each domain. Verbatim quotes supporting each domain are presented throughout each section. To maintain patient confidentiality, no personal identifiers are included in the results.

Manifestations (signs, symptoms and biomarkers)

Pain

All participants reported foot and/or ankle pain. Severity of pain varied, with severe pain described as "unbearable … the most excruciating pain ever" (P46) and "like torture" (P19). Participants described different types of pain, including "a throbbing, sharp, stabbing pain" (P13), "like walking on knives" (P55), "shooting pain" (P20), "like a bruising" (P21), "a burning hot pain" (P44), "like somebody's sticking pins in the ends of your toes" (P10), and "like [a] toothache in your foot" (P4, P5, P40). Pain was experienced at different times, including during or after weightbearing, at rest, on movement of the foot or ankle, on pressure, in the morning, and at night. For most participants, resolution or reduction of foot/ankle pain mattered most, compared to other treatment outcomes. Most participants with systemic diseases deemed their foot or ankle the area of their body most affected by pain or their biggest problem, and it was often the first area to have been affected.

Physical function

Most participants described limitations in physical function, regardless of RMD, race/ethnicity or geographic location. Participants conveyed limitations in movement in their foot/ankle, describing a loss of flexibility. Some participants perceived that range of movement was as important or more important than pain, contributing to "feeling disabled" (P26), although one participant described it as "something that creeps up slowly, so you get used to it" (P7). Stiffness in the foot/ankle was also frequently reported by participants, described as locking and gelling of the joints, and the feeling that "a little rod of steel is running through my big toe" (P47). Limited range of movement and stiffness were closely linked, with some participants considering these to be the same concept, although others felt that stiffness was less impactful in comparison:

> “Ideally I would like to be able to move my ankle better. I feel like at this point I can put up with the pain and I would rather just not be limited in functionality more so.” (P36)

> “If you’re driving you need your ankles to move, so driving becomes difficult, climbing stairs, bending about, anything that you do in a regular world becomes more difficult.” (P7)

> “Stiffness for me is a heavy feeling, with mild pain. My foot feels tired, I walk slowly, like I’ve already been walking for hours and hours. It takes more energy to move.” (P48)

> “If I had a choice between a tablet for pain and a tablet for stiffness, I'd take pain any day.” (P20)

Participants also experienced weakness and joint instability in their foot/ankle, reporting experiences of “a sudden twist” (P23), “giving way” (P21), and “feeling like my ankle could turn very easily” (P52). Some participants perceived that muscle weakness led to balance issues; participants reported being more conscious of their feet, and being more conscious of the ground when walking, as a result:

> “The whole foot feels very unstable because quite a lot of the time, if I’m just walking normally, even in flat shoes across the car park at work, if I’m on the slope I don’t feel steady.” (P51)

> “I have this weakness in my feet so I don’t feel as steady, I’m afraid. I feel my balance is not that steady anymore.” (P49)

Changes in gait were reported by participants, including a reduction in walking speed and stride length:

> “So, I’ve almost got to the point where I’m making little short paces. Just like, going, tick, tick, tick, from side to side because I can’t stride out normally.” (P51)

Participants described “limping” (P20, P55, P51, P52, P45), “waddling” (P24), “hobbling” (P38, P41), “hopping about” (P6), “walking on the outside of my feet” (P56, P30) and being “unable to walk in a straight line” (P10).

> “I'm aware that I can’t walk as much as I used to, I used to love walking a long way. I also can’t walk so fast.” (P25)

Some participants felt their foot/ankle disorders impacted on other joints, particularly the knees and hips, as a result of altered gait patterns:

> “If you limp, then your knees will go bad, then your hips will go on you.” (P33)

Participants also recognised the negative impact of reducing weightbearing activity on other joints:

> “I did (sic) rest the foot, which goes against my personal mantra, because then everything else just seized up. So my foot felt a lot better, but then didn't leave me overall better off.” (P47)

Fatigue

Participants reported experiencing physical fatigue as a result of their foot/ankle problems (Q4), describing “getting tired so easily” (P2), feeling “absolutely washed out” (P19), “exhausted” (P47) and “weary” (P43), “like the whole body doesn’t want to move.” (P49)

> “The more pain I’m in, the more fatigue I have. The pain will go away, but the fatigue will still be there, because I’m just completely strung out by the amount of pain that I’ve pushed through.” (P46)

Foot and ankle problems could result in both overall fatigue and fatigue localised to the foot and ankle (Q5), described as aching and tired feet. Participants with systemic RMDs considered the fatigue resulting from their foot and ankle problems as “extra fatigue” (P39, P41) and that their overall fatigue was worse when their feet or ankles were worse. Foot/ankle pain was perceived to be the primary cause of fatigue for most participants, although muscle weakness was also implicated. Foot/ankle-specific fatigue often occurred as a result of activities and participation, whilst some participants perceived that being unable to exercise due to foot/ankle problems was the cause of overall fatigue. Participants also reported experiencing cognitive fatigue, described as “a heavy weight or burden that I’m carrying around” (P47) and feeling “weary and worn out” (P43) “mentally drained” (P36, P48) as a result of foot/ankle problems, particularly pain and associated worry. Cognitive fatigue affected concentration:

> “The discomfort becomes a distraction to everything, it overrides everything … it affects your concentration because you just … you’re, like, oh I can’t be bothered, I just want to sit down.” (P43)

Fatigue was reported by participants with all RMD types, but was more frequently reported and more impactful in those with systemic conditions.

Deformity

Participants discussed a range of deformities in their foot/ankle, including hammer and retracted toes, overlapping toes, bunions, and flatfoot. They described “crooked and squished” (P55), “bent” and “distended” (P56) toes, feet “growing sausagey (*sic*)” (P47) and “going over to one side” (P16), and ankles “falling inward” (P40). Appearance of deformity was less important to some participants than pain:

> “I know people say, oh, I want my feet doing because they don't look nice. I don't care what they look like. I'd have all my toes off. I say sometimes, oh, I just wish I could have my feet taken off, have a transplant if they did foot transplants. But it's not about how they look, it's how they feel … it’s the pain.” (P15)

Deformity contributed to stiffness and reduced range of movement, skin and nail issues, and altered gait. Deformity was reported among participants with all RMD types, but was deemed less severe among those with non-systemic, localised MSK disorders compared to other RMDs. For example, non-systemic, localised MSK disorders did not tend to cause an impact on footwear choice or emotional distress.

Skin and nail health

Participants reported skin and nail problems, including callosities, blisters, dry skin, cracking skin, ingrowing, breaking, and thickened nails. Skin and nail issues were linked to deformity and attributed to discomfort with footwear and foot orthoses, and were problematic irrespective of RMD, race/ethnicity or geographic location. These issues often resulted in pain:

> "I don't like having that dry skin … because it's painful, to be honest, because actually it cracks and then goes deeper inside." (P21)

Most participants with skin and nail problems were unconcerned by their appearance, and many participants perceived they were more easily treated. Some participants also experienced colour changes in their feet and described their experiences of foot ulceration and infection; in CTDs, wound healing was often a priority due to the severity of pain that digital ulcers caused:

> "The ulcers are so painful, so painful. I wouldn't wish it on nobody, even my worst enemy." (P7)

> "The thing that concerns me most is that the skin between my toes dries up and breaks, it bleeds and there is pus." (P32)

Swelling

Of participants who reported swelling in their foot/ankle, some described a feeling of tightness (P11) or imminent explosion:

> "Even when I wear the flat shoes, the ones I wear mostly out of everything, my foot throbs because it feels like it's going to explode with the swelling." (P16)

Swelling could occur at rest or after weightbearing. Swelling was reported in all RMD types, but was often deemed more important when it caused pain, particularly in crystal arthropathies:

> "There's been occasions where my feet, my ankles especially, [they've] really swollen up … it sort of bursts your footwear off. You tie your laces and it's extremely painful, and when you take your shoe off it looks like – I can only describe it like an elephant's foot." (P17)

Temperature changes

A minority of participants described temperature changes in their foot/ankle, which could be painful, particularly in RP.

Numbness

A minority of participants experienced episodes of numbness, which could occur at the same time as pain:

> "It feels like, on the bottom of both feet, right in the pad of your foot, it feels like a wasp sting that's maybe an hour old. So it's numb and achy and hurts all at the same time." (P38)

Poor circulation

Some participants, with RA, SpA, and CTDs, described poor circulation, with one participant with a CTD stating that circulation was the biggest problem affecting their feet.

Cramping

A minority of participants reported cramping in their feet, resulting in pain. Cramping was described as “locking”, “a horrible feeling” (P4) and “unpredictable” (P56), and could occur during weightbearing, at rest, and at night.

Life impact

Activities and participation

All participants acknowledged the impact of foot/ankle disorders on activities and participation. This impact varied; some participants were severely restricted, whilst others felt they could do their usual activities but for less time, or by making adaptations, pushing through pain or accepting they would be in pain afterwards.

Obligatory activities

Most participants reported that foot/ankle disorders impacted on walking:

> “I used to like walking a lot, but now if I walk halfway to the bus station, my foot swells.” (P28)

For many, being able to walk was a priority. Participants described reducing the quantity and distance of walks due to pain when weightbearing, and difficulties walking up and downhill up and down stairs and on uneven surfaces, linking to reduced range of movement, muscle weakness, instability, and fear of falling:

> “It’s the fear of falling down, this is the biggest, for me, because sometimes I feel like, okay, I forgot all about my achiness or something, but for me, this is my fear. When I fall, it stops all my happiness going out.” (P21)

Swelling, temperature changes, numbness, and skin and nail problems also contributed to reduced walking ability. Some participants required walking aids or to hold onto something or someone to be able to walk. Personal care was also affected:

> “Your feet are the most important thing and if you can’t stand on them, you can’t walk, you can’t even do the fundamentals of going to the toilet or using a bath or shower.” (P5)

Foot/ankle disorders affected driving; participants described changing to an automatic car and limiting driving duration as a result of pain, cramping, or numbness:

> “For me, it’s about maintaining the ability to walk independently and being able to drive as well. They are what I would call my priorities. I want to be able to maintain movement in the joints – I don’t like the idea of not being able to bend them, but for me it’s about maintaining independence and mobility.” (P36)

> “I’ve had times where I’ve not known where to put my foot on the accelerator pedal and I’ve had to use my left foot, which is obviously not good, that’s dangerous. The same with the left foot, struggling to use the clutch.” (P5)

Committed activities

Participants described giving up work, modifying their roles at work or changing jobs, making adjustments (e.g. needing to take breaks at work, change shoes, and stand less), reducing their hours, and taking time off work:

> “I had to give up teaching actually because of my feet and that was my life.” (P45)

> “At times, my feet are so bad I can’t move, so I can’t go to work.” (P31)

Some participants perceived that their professional image was affected by their foot and ankle disorders, due to restricted work activities, footwear impact, and subsequent clothing restrictions. Another participant described feeling they were not as visible at work due to restricted work activities, and perceived that this had affected their career progression (P11):

In most cases, participants in jobs that involved weightbearing were most affected, although pain at rest and cognitive fatigue could also affect non-weightbearing work. Work disruption caused psychological distress and had financial implications. Foot/ankle disorders also impacted on participants' family roles, leading to reliance on others and loss of independence, and affected household responsibilities, including shopping, cooking, cleaning, and ironing:

> "I can't take them out, the little ones. I can't take them out for a walk 'cause I can't walk anywhere. It makes you miserable." (P43)

Discretionary activities

Participation in religious activities was also affected by foot/ankle disorders. Participants of, for example, an Islamic cultural background, in Egypt and the UK, described modification of Muslim prayer movements, such as sitting in a chair to pray, and praying at home rather than at a mosque, due to pain when weightbearing and limited range of movement in the foot/ankle:

> "I like praying and performing soujoud and kneeling, but I can't now. I pray sitting on a chair. At times I cannot even lift my foot to do the ablution, so I just wet it by hand, but I do wash it properly for the dawn prayer. But for the rest of the prayers, I just pass my wet hand over it because it is very painful." (P28)

These restrictions led to psychological impact for some participants. Many participants reported a decline in socialising due to their foot/ankle disorders, avoiding leaving the house, and cancelling or rescheduling plans, which affected relationships with friends and family:

> "Everybody is going out for nice long walks … it was very much a case of, like, I actually physically can't walk that far … it's definitely impacted relationships and quality time." (P24)

> "If I'm meeting friends, I have to think about if we're going to be walking, how far are we going to be walking, can I walk that distance or am I going to be holding everybody up and [sic] embarrassing… If necessary, do I need to take a wheelchair, do I need to take my scooter? And so, yes, I've got to think about these things all the time, whereas most people just get there and do it, you know." (P8)

> "We go up and see [family member], but the corns and everything, the hard skin on the heels … it does tend to put you off doing anything apart from tootling about the house." (P2)

A reduction in social activities were usually a direct result of pain, but was also influenced by temperature changes, instability, balance and fear of falling, or foot ulceration. One participant acknowledged the social consequences of not wanting to show the bottom of her feet:

> "I'm most comfortable with my feet up so I always try to find a place where I can do that. And if you go out to eat or you go to a friend's, you feel a little funny showing the bottom of your feet so I don't do it as much, so that constrains some of my social activities that I probably would otherwise do." (P38)

Multiple sports, hobbies and activities for relaxation and pleasure were affected, and led to social isolation in some cases:

> "I used to have a hobby. I liked fishing … last time I went was a few years ago, I stayed there a short time and my foot became so swollen, almost like a football. My children asked me, why am I doing this to myself? … I have fishing gear at home and I look at it every so often and smile." (P28)

> "I used to play football regularly. I don't go anymore. Now the people who go, they're not texting me anymore because every time they text me, I say no. Sometimes they say, oh, come on, just sit with us. And I said, 'it makes me more upset. I'd rather sit at home. Not to be there, to watch you all because I want to join, I want jumping and playing, you know.' But what can I do?" (P1)

Other participants highlighted the impact of foot and ankle disorders on the ability to exercise, expressing that foot health is central to staying active:

> “The feet are your foundation. How do you walk around and live a normal life if you’ve got sore feet all the time, and seriously sore feet? And how can you encourage people to exercise and go out and walk when it’s their feet that are the biggest problem that they have?” (P56)

> “I do think that it would be great if researchers could find ways to address issues that relate to the exercise, fatigue, sore feet syndrome because I do think that people can get into a pain cycle and they can’t break out of it. One way to break out of it, I think, is exercise, and if you can’t do that because your feet hurt, then it affects so much more.” (P40)

Participants also discussed the impact of their foot/ankle disorders on holidays and special occasions, which were particularly affected by activity limitations and footwear impact.

Footwear impact

Most participants described limitations in footwear choice, including the inability to wear off-the-shelf or “normal” (P28, P32) shoes, being limited to trainers and giving up wearing “decent shoes” (P55).

> “I can’t wear what I want. Impossible.” (P22)

Footwear limitations could impact on participants’ social participation and hobbies. Whilst pain and deformity were major factors affecting choice of footwear, limitations also occurred as a result of swelling, foot/ankle instability, and reduced balance. Circulation issues and temperature changes could also limit footwear choice, particularly in winter and in colder locations. Many participants highlighted issues with footwear aesthetics, particularly women and/or those with RA, SpA, and JIA, who wore orthopaedic shoes due to severe foot deformity. These were described as looking weird, clumpy, frumpy, ugly, and heavy. Footwear limitations impacted on clothing choices; some participants (all women) felt they were unable to dress up:

> “Trying to get shoes is a nightmare … everything I have now is not pretty, they do not make shoes to be pretty for our health issues.” (P50)

> “If you put a dress on, well then you’ve got to have shoes to suit your dress. But if you’ve got a dress and you’ve got clumpy shoes on, it doesn’t look too appealing.” (P52)

Participants tried to hide their shoes and noting it was “hard to find something that looks somewhat attractive” (P41). Some participants highlighted footwear as a priority:

> “I know this is really vain, but I want [my feet] to feel okay, so I can wear my shoes.” (P4)

Some participants acknowledged that they cared what their feet and shoes looked like when they were younger, but no longer did, highlighting the influence of age on priorities. Footwear limitations impacted on activities and participation, gait, and psychological health, leading some participants to identify footwear impact as a priority, irrespective of race/ethnicity or geographic location.

However, many participants prioritised comfort, as uncomfortable shoes led to pain when weightbearing:

> “I’m just not able to walk as well in certain types of shoes, so I definitely prioritise the comfort aspect.” (P36)

Participants who had found comfortable shoes explained how “they allow me to do what I want to do” (P46). linking to gait and perception of body image/self-identity:

> “I can walk and not limp. I can walk normally and people don’t see immediately that I have a health problem. Nobody can tell.” (P55)

Psychological impact

Foot/ankle disorders affected mood; participants described multiple emotions, including anxious, depressed and frustrated, feeling "fed up" (P52), "miserable" (P43), "down" (P21, P22, P24, P15, P43, P3) and "wanting to be alone" (P19):

> "It's depressing, and very anxiety inducing too, because … is it going to be like this forever? Is it going to get worse? Just the uncertainty." (P54)

Some participants described an inability to cope. Foot/ankle pain was the primary symptom that affected participants' moods, but mood changes were also attributed to overall, foot/ankle-specific, and cognitive fatigue. Participants expressed worries about swelling, numbness and discoloration, paranoia about their feet changing shape and looking different, and anxiety regarding causing foot or ankle damage by doing certain activities. Other participants expressed frustration and misery in having to restrict activities and participation, which affected their independence, morale, and self-confidence. The benefits of exercise for mental health, pain and fatigue, and subsequent negative psychological impact of not being able to exercise due to foot/ankle problems, were also emphasised by participants.

> "I used to jog myself to get that rush feeling, you know, the good feeling … you know … the heartbeats. That's what I miss the most … I'd go off to gym, so I'd become more addicted to it. You know what I mean? One day if I don't go to gym, it was like something [was] missing in my life. Now it seems like I've lost it completely … I need to clear my head, so I'm going to walk on my hand?" (P1)

For some participants, particularly those with RA, SpA and adult JIA, appearance of deformity and the subsequent footwear impact also had a negative psychological impact, whilst other participants were unconcerned by deformity, prioritising pain, activities, and participation. Some participants described a fear of falling, linking to muscle weakness, instability, and balance issues, and also expressed fears relating to the future, ending up in a wheelchair. Emotions were particularly affected by uncertainties around prognosis and treatment, especially when participants perceived that treatments had failed, there was no plan in place, or treatment options had been exhausted. Other participants recalled being affected emotionally when they first experienced foot/ankle problems, but described getting used to these over time,

highlighting the potential influence of RMD duration on this domain. Psychological impact was important to participants regardless of race/ethnicity or geographic location.

Sleep

Foot/ankle disorders affected participants' sleep. Sleep was primarily affected by pain at rest or at night:

> "A lot, all the time, I cry because of pain when I try to stand on them. I cannot sleep without the treatment, I cannot walk or go anywhere." (P26)

However, other contributors included temperature changes and stiffness:

> "If it's not the pain, it's the stiffness. I have to get up and move for a while. Go for a walk around the house, then go back to bed. I can't lie still for more than three or four hours." (P20)

Some participants reported that they were unable to sleep without painkillers for their foot/ankle pain, and lack of sleep was perceived to cause fatigue.

Treatment experience

Participants discussed their experiences of seeking and receiving treatments for foot and ankle disorders, with many reporting unresolved symptoms:

> "I have taken five different kinds of tablets and injections, and I still have foot pain." (P27)

Some participants perceived that health professionals were uninterested in or unconcerned by foot and ankle problems, whilst others felt they had exhausted all treatment options, having no treatment plan in place and feeling given up on, with "lack of answers almost as hard as the pain itself" (P41):

> “I don’t have to walk on my hands, they can get a break and calm down and stuff. It’s very hard to get treatment, it’s difficult and I’ve just been putting up with it the best way that I can. Yeah, I feel like the doctors have given up on it with me.” (P47)

> “I've complained considerably to my rheumatologist, and he seems to think there's nothing we can do about that. There's an attitude among the medical profession referring to my rheumatologist and my GP, everyone gets it [foot OA], anyone who's older gets it, and there's nothing we can do so shut up and go home.” (P20)

> “Where is the support? You can’t find evidence, you can’t find encouragement, you can’t find professionals who know the ins and outs of footcare and how to remain functional, how to maintain the function of your foot through a disease that’s going to last you a whole lifetime.” (P55)

Dissatisfaction with a range of foot and ankle treatments, including footwear, steroid injections and surgical procedures was reported, in addition to overall dissatisfaction with the lack of treatments offered or available. One participant described foot and ankle treatments as crude compared to treatments for other areas of the body. As explored above, experiences of treatments had psychological impact, and were influenced by participants’ expectations of treatments, education and knowledge about their condition and treatments, self-efficacy, self-advocacy, and healthcare providers:

> “I’m a very strong advocate for my own care and if I need somebody to look at something and they won’t do it, then I’ll go and find somebody that will.” (P46)

Societal/resource use

Healthcare utilisation and personal expenses

Healthcare utilisation and personal expenses were influenced by geographic location and subsequent healthcare systems, and overlapped in some cases. Participants described needing multiple consultations with different HCPs and undergoing surgery. Participants with foot ulceration required ongoing wound management and dressings.

In some cases, foot/ankle infection had resulted in, or lengthened, time in hospital. Participants identified extra personal expenses because of their foot/ankle disorders, including those for footwear and orthoses, highlighting that only certain brands were comfortable and that shoes wore down quickly due to deformity:

> "I'd hate to think how much money I've spent on orthotics over the years. Shoes, mainly sports shoes, mainly those that give me that cushion effect and rock forward, yeah I'm still in those sort of shoes all the time these days. Probably every four to five months it's another $220." (P45)

Participants, particularly those outside of Europe, emphasised the cost of medications and devices for their foot/ankle disorders, either directly or through insurance plans:

> "My doctor prescribed me an injection, but, the price increased, it used to sell for 50 Egyptian pounds, so I stopped buying it. I bought the generic, but it made me worse, so I stopped it completely. I only take painkillers now." (P32)

Multiple consultations with different HCPs, particularly podiatrists and physiotherapists, also impacted on personal expenses. In many cases, participants felt that costly consultations and treatments had not been successful. Giving up work, reducing hours, or changing job role had financial implications for participants, as did the need for public and private transport as a result of reduced work. RMD type had no influence on healthcare utilisation or personal expenses.

> "When I had problems with my foot, I couldn't stand anymore. My job required me to stand all day … I gave up. I have been unemployed for nine or ten years now … I am in receipt of social security payments. It affected my psyche." (P28)

5.5.2 Reflexivity summary

As discussed previously, the candidate made field notes and kept reflective logs throughout the interviews. Field notes included descriptions of behaviour and non-verbal observations, e.g. one participant sounded angry when asking why nothing could be done about his foot and ankle problems. Another participant discussed how difficult it was not being able to spend quality time with her son because of her foot pain; the field notes in this instance detailed the sadness in her voice. These field notes

aided the analysis process; a challenge of analysing the translated and transcribed interviews from participants in Egypt was the lack of context, compared to the richness of data obtained from the interviews that the candidate carried out.

The candidate considered her own interview technique after each interview and did not do more than one interview per day, so that adequate time could be taken to reflect. During early interviews the reflective logs expressed some concerns about technique; for example, feeling awkward during silences and not giving participants time to answer before asking another question. Some of the early interviews were short in duration, such as when participants did not elaborate on their answers. Upon listening to the audio recordings, the candidate recognised that additional prompts could have been used. The reflective logs often contained an action plan, made with the candidate's supervisor (CAF), for future interviews. For instance, one of the action plans was to stay focused on the conversation taking place rather than worrying that a participant is only giving short answers, and to work on putting participants at ease and building rapport. The candidate continued to reflect over the course of the qualitative study, and was able to recognise improvements in technique, with silences becoming more comfortable and more probing occurring when appropriate.

5.6 Discussion

The qualitative interview study reported in this chapter achieved its aim of building on existing literature and establishing which domains are important to patients living with RMDs who had sought advice and treatment for foot and ankle disorders.

5.6.1 Main findings

This is the largest qualitative study involving patients with foot and ankle disorders in RMDs. It is also the first to consider the perspectives of patients with foot and ankle disorders in less common RMDs, including adult JIA, AS, enteropathic arthritis, mixed CTDs, and pseudogout, where interventional studies are particularly lacking, and has identified domains that are important to patients with foot and ankle disorders across a range of RMDs and geographic locations, of diverse races/ethnicities. Findings provide rich insight into the experiences underpinning priorities to inform the development of a standardised COS for future foot and ankle studies.

Whilst many domains identified from the qualitative interviews were common across all conditions, the importance of some domains differed depending on the RMD. Notably, deformity was an important domain among most participants with RA, SpA and JIA, but was less of an issue for most participants who had MSK disorders affecting the foot

and ankle in the absence of any systemic condition. Additionally, skin health (including callosities, foot ulceration, and infection) was particularly important to participants with CTDs. In addition to domains, this study has identified contextual factors, defined as variables that are not outcomes of a study but need to be recognised and measured to understand the study's results (240). Contextual factors can be categorised as personal, disease-related, and environmental factors (160); findings from the qualitative interviews indicate that age, sex, race/ethnicity, disease duration, geographic location, treatment experiences, and healthcare systems could influence the results of foot and ankle studies and should be considered in future work. For example, healthcare systems represented in the current study included publicly funded systems (e.g. UK, Canada), predominantly private insurance/out of pocket payment systems (e.g. USA), and mixed public-private systems (e.g. Australia, Egypt), which may impact on patients' experiences of and satisfaction with treatments, and their personal expenses.

Findings from this study corroborate the qualitative synthesis of patient perspectives, which emphasised the extensive impact of foot and ankle pain, deformity, and physical function domains on activities and participation, psychological health, footwear needs, and personal finances (73, 254). This study also builds upon findings from the secondary analysis of focus group data (Chapter 4), increasing understanding of the domains cramping, numbness, temperature changes, and colour changes. These domains were described by patients with multiple RMDs, despite not being measured in the previous scoping review (73) or identified in the systematic review of qualitative studies (Chapter 3).

The current study explored the full scope of physical function in more depth, indicating that stiffness, limited range of movement, joint instability, and altered gait patterns could also affect participants across multiple life impact domains. The qualitative synthesis identified that treatment satisfaction was a meaningful domain to patients, predominantly relating to dissatisfaction with footwear choice. In the current study, treatment experience dissatisfaction was more broadly identified, covering experiences of healthcare, but encompassed satisfaction level with different foot and ankle treatments (or lack thereof). In some cases, patients' perspectives may have been influenced by the paradigm shift in the treatment of RMDs over the last two decades (337). For example, the advent of biologic therapies has revolutionised the treatment and outcomes in many rheumatic diseases, changing patients' expectations about disease control. There have also been advances in the treatment of hip and knee OA, with definitive surgeries available (e.g. total hip and total knee replacements) (338, 339). As discussed in Chapter 1 (section 1.5), there is a lack of high-quality evidence for treatments of foot and ankle disorders in RMDs, and progress in this area has not

kept pace with other areas of the body or with broader disease. Therefore, there is a potentially a gap between patients' treatment expectations and available treatment options, contributing to their dissatisfaction.

The contribution of foot and ankle pain to the overall broad domain of physical fatigue (extreme tiredness resulting from physical exertion) is well established (254). The current study highlights the influence of foot and ankle disorders on the target domain cognitive fatigue, with participants identifying the specific impact of their foot and ankle problems on concentration. Comparably, in a qualitative synthesis undertaken to inform the development of the OMERACT core domain set for shoulder disorders, experiences of cognitive dysfunction (problems concentrating and with memory) were identified (257). These experiences transpired as a direct result of shoulder pain or due to the sleep deprivation that occurred because of this pain.

In the current study, a new target domain, foot and ankle-specific fatigue, was also identified, describing an overwhelming tiredness localised to the foot or ankle. This concept was not identified in other core domain sets focusing on specific areas of the body in RMDs (149, 320, 340), and may reflect the functional importance of the foot and ankle in all weightbearing activities. The impact of foot and ankle disorders on other joints (as a result of altered gait) was also identified by participants. Finally, whilst activity and participation limitations were prominent in the preceding qualitative work, the increased participant diversity in this interview study led to the identification of the impact of foot and ankle disorders in RMDs on religious activities.

5.6.2 Strengths and limitations

The large overall sample size and the inclusion of ethnically and culturally diverse participants, from multiple countries representing different healthcare systems are strengths of this study (121). PPI input into the interview schedule, its pilot, the coding framework, and the analysis were also key strengths, ensuring questions the participants were asked were relevant and accessible, and that the interpretation of findings were grounded in the patient perspective.

The small number of participants within some RMD categories could be considered a limitation, and some RMD subtypes (e.g. limited cutaneous SSc versus diffuse cutaneous SSc) are very rare diseases that were not part of the purposive sampling criteria. Similarly, RMD/foot and ankle disorder severity was not recorded in this study, and duration of foot and ankle problems was self-reported by participants. Therefore, findings from this study may not represent the priorities or experiences of other patients with foot and ankle disorders in RMDs. Notwithstanding, findings relating to the impact

of foot and ankle disorders in RMDs provide a grounding for future studies in under-researched areas.

A further potential limitation is that the interviews were conducted by HCPs, therefore clinician-patient power dynamics may have influenced recruitment and data collection (336, 341). For example, patients might participate in research studies because they feel that declining could affect their care, and might not be fully open about their experiences if they are talking to a HCP, especially those relating to experiences of healthcare. To mitigate this, potential participants were reassured that participation was voluntary, their care would not be affected, and their data would be anonymised. This information was also provided in the PIS. The candidate was transparent about her dual role as a researcher and a podiatrist, and reiterated her motivations for conducting the research (e.g. to understand what is important to patients, in order to inform future research).

Finally, interviews conducted in English with participants whose first language was not English, and translation of interviews conducted in Arabic, may have led to the loss of nuances and cultural references. For example, certain concepts can lack direct language equivalents, resulting in ambiguity and hindering understanding of the context and depth of responses (342). Nevertheless, all participants provided valuable accounts of their experiences of living with foot and ankle disorders and the potential for interpretation bias was offset by robust translation and transcription processes. The process of involving international collaborators and participants from LMICs in the development of the COS will be reflected upon in more depth in Chapter 7.

5.6.3 Implications

5.6.3.1 Implications for clinical practice

By providing a comprehensive overview of the priorities of patients with foot and ankle disorders in multiple RMDs, findings from this study are directly relevant to clinical practice, helping to guide the content of clinical appointments, particularly in the management of patients with RMDs where foot and ankle involvement has previously been overlooked.

5.6.3.2 Implications for research

This study also has implications for future research. The variation in experiences of seeking and receiving foot and ankle advice and treatments demonstrated in this study

highlights a case for measuring health resource use in future research. Experiences of treatment uncertainty and unresolved symptoms reflect the critical need for future and definitive RCTs to determine their clinical effectiveness. Findings from this research inform a Delphi consensus study aiming to prioritise domains for inclusion in a COS, which will be presented in the next chapter. Qualitative findings will ensure the domains presented in the first round of the Delphi reflect what is important to a diverse range of patients. Interview data will also inform plain language definitions of domains, aiming to improve study accessibility, which may enhance patient recruitment and retention. Participants' choice of words when describing domains can also be used to inform the measurement of domains in future work.

5.7 Conclusion

The qualitative interview study reported in this chapter achieved its aims and objectives by examining the range and scope of domains that are important to patients with foot and ankle disorders in RMDs, exploring why these domains were important, and establishing participants' choice of words. Findings provide a valuable original contribution to the literature by including the experiences and opinions of a diverse range of patient participants, with a range of RMDs, first languages, and races/ethnicities, and from geographic regions that had previously been under-studied. The next chapter presents an international Delphi consensus study, where the domains identified in this qualitative work are prioritised for inclusion in a core domain set.

Chapter 6 Consensus on a core domain set

6.1 Introduction

The following chapter addresses the COMFORT project's second objective, to achieve multidisciplinary, multi-contributor, and expert international consensus on a core domain set for foot and ankle disorders in RMDs. This work incorporates findings from a scoping review of existing studies (73), as discussed in Chapter 1 (section 1.5.1), a synthesis of existing qualitative studies (Chapter 3), a secondary analysis of focus group data (Chapter 4), and primary qualitative interviews with patients (Chapter 5).

6.2 Aim and objectives

The aim of this study was to gain international, expert consensus on a core set of domains to measure in future clinical trials and observational studies investigating the effectiveness of treatments for foot and ankle disorders in RMDs. The objectives were to:

- Determine a list of domains and their respective definitions, identified from existing clinical trials and observational studies, and from qualitative research;
- Provide participants with this list and a mode to confirm how important each domain is in the context of measurement in future foot and ankle RMD research;
- Analyse the level of agreement between patients and other key contributors (HCPs and researchers);
- Provide feedback to participants regarding their own and other group responses, and allow participants to re-rate their level of agreement;
- Achieve final consensus on a core domain set.

6.3 Methods

The COMFORT consensus process involved an online, four-round, modified Delphi consensus study, informed primarily by OMERACT's standardised COS methodology (63). Consensus methods were discussed in detail in Chapter 2 (section 2.5.2.3); to recapitulate, the Delphi method is frequently used to achieve consensus during COS development as it allows for wide geographic dispersion and for large numbers of key contributor groups to participate (174). Additionally, responses are anonymous,

avoiding the effect of dominant individuals (174). The consensus process was developed and reported in line with COS-STAD and COS-STAR, respectively (109, 110) (Chapter 2, section 2.5). The COS-STAD checklist for COMFORT is presented in Table 6.1.

Table 6.1: COS-STAD checklist

Domain	**Standard number**	**Methodology**	**Thesis location**	**Notes**
Scope specification	1	The research or practice setting(s) in which the COS is to be applied	Chapter 2, section 2.9.1, p.57.	-
	2	The health condition(s) covered by the COS	Chapter 2, section 2.9.1, p.57.	-
	3	The population(s) covered by the COS	Chapter 2, section 2.9.1, p.57.	-
	4	The intervention(s) covered by the COS	Chapter 2, section 2.9.1, p.52.	-
Stakeholders involved	5	Those who will use the COS in research	Chapter 1, section 1.5.1, p.9-11	Involvement in preceding scoping review.
	6	Healthcare professionals with experience of patients with the condition	Chapter 6	Involvement throughout Delphi.
	7	Patients with the condition or their representatives	Chapters 3-6	Involvement throughout qualitative studies and Delphi.
Consensus process	8	The initial list of outcomes considered both healthcare professionals' and patients' views	Chapter 6, section 6.3.1, p.154-155.	-

	9	A scoping process and consensus definition were described a priori	Chapter 6, section 6.3.8, p.173.	-
	10	Criteria for including/ dropping/adding outcomes were described a priori	-	Decisions to add, combine or drop domains were made through discussions (Chapter 6, section 6.3.7.3, p.172) and not based on criteria.
	11	Care was taken to avoid ambiguity of language used in the list of outcomes	Chapter 6, 6.3.7, p.166	-

6.3.1 Survey development

To develop Round 1 of the Delphi survey, all domains identified in the scoping review and qualitative research were discussed with the PAG in an online meeting on 17th October 2024. As a higher number of domains in a Delphi study can negatively affect response rate (213), the PAG made decisions about combining and dropping domains. The OMERACT 'binning and winnowing' process was followed: domains were either 'binned' (involving grouping related domains or concepts into a broader domain) or 'winnowed' (involving dropping domains because they were too narrow, too broad, or not relevant for the COS) (343). Decisions were influenced by domain granularity; some broad domains overlapped with target domains, which can be confusing for Delphi participants (214).

An example of the process that was undertaken to reduce the number of domains in the first round of the Delphi study is presented in Appendix K. All new domains identified in the scoping review and qualitative studies were recorded in a standardised OMERACT template, as they were identified. Each domain was then discussed with the PAG to determine whether it overlapped and could be combined with any other domains, was too broad or too narrow, or lacked importance or relevance. For example, swelling was identified in the qualitative synthesis (Chapter 3) as a potentially important domain, but this was deemed too narrow as it overlapped with the broader domain of inflammation (identified in scoping review). Therefore, swelling was 'binned'

and inflammation was included in the first round of the Delphi study. An initial list of 137 domains was reduced to 32 Delphi items during this process.

Names of domains, with working definitions, were constructed by the candidate after each phase of the core domain set development process and discussed with the PAG until agreement on each definition was reached. Patient-reported domain definitions were based on the data (e.g. language used by patients in the qualitative work). The final domains and definitions taken into the first round of the Delphi study, organised by broad and target domains within the OMERACT core areas of pathophysiological manifestations, life impact, and societal/resource use, are displayed in Table 6.2, Table 6.3, and Table 6.4, respectively. As adverse events (including death) is a mandatory domain in all OMERACT core domain sets, it was not presented in the surveys.

Table 6.2: Pathophysiological manifestation domains

Broad domain	Target domain	Definition
Pain	Foot/ankle pain intensity	How severe pain in the foot/ankle is.
	Foot/ankle pain during weightbearing	Foot/ankle pain experienced whilst weightbearing (e.g. walking, standing).
	Foot/ankle pain during non-weightbearing/at rest	Foot/ankle pain when sitting or lying down.
Cramping	Cramping	A cramping sensation in the foot or ankle.
Fatigue	Fatigue	Range of symptoms from mild subjective feelings of tiredness to an overwhelming debilitating, and sustained sense of exhaustion that likely decreases one's ability to execute daily activities and function normally in family or social roles.
	Foot/ankle-specific fatigue	Fatigue localised to the foot/ankle – a feeling of tiredness and aching in the foot/ankle; needing to rest.

Numbness	Numbness	A loss of feeling in the foot/ankle.
Disease activity	Foot/ankle inflammation	Swelling, tenderness, redness and/or warmth to an area of the foot/ankle.
	Joint damage/deformity	How much the shape of a foot/ankle joint has changed with the condition.
	Colour changes	Colour changes in the foot/ankle.
	Patient global change in foot symptoms	How symptoms in the foot/ankle have changed between different timepoints.
	Physician global change in foot symptoms	How well the physician/clinician feels that symptoms in the foot/ankle have changed between different timepoints.
Joint ROM	Joint range of motion	The current amount of movement in a foot/ankle joint.
Stiffness	Foot/ankle stiffness	How stiff the foot/ankle feels.
	Foot/ankle morning stiffness	How difficult it is to move the foot/ankle when first getting out of bed in the morning.
Joint instability	Joint instability	How weak or unstable the foot/ankle joints are; how much balance is affected.
Nail abnormalities	Nail abnormalities	Ingrowing, thickened or otherwise abnormal toenails.
Presence of callosities	Presence of callosities	Hard pieces of skin on the feet (e.g. callus, corns).
Ulceration	Foot/ankle ulceration	An open wound on the foot/ankle.
Circulation	Sufficiency of blood supply	Sufficiency of blood supply to the foot.
Gait	Normality of gait	Any differences from expected ‘normal’ walking.

Table 6.3: Life impact domains

Broad domain	Target domain	Definition
Physical function	Physical function: Necessary activities and participation	Impact of foot and ankle problems on walking (including ability to walk, walking distance), driving, personal care, household tasks, caring responsibilities and paid work.
	Physical function: Discretionary activities and participation	Impact of foot and ankle disorders on being able to participate and enjoy life events, e.g. sports, exercise, social and recreational activities, religious activities, hobbies and voluntary work.
Emotional wellbeing	Emotional wellbeing	Emotional condition as a result of foot/ankle problems, e.g. low mood, anxiety or ability to cope.
	Body image	A person's emotional attitudes and perceptions of their body as a result of foot/ankle problems.
Sleep	Sleep	Ability to sleep (including falling asleep and staying asleep) as a result of foot/ankle problems.
Footwear	Footwear comfort	How comfortable footwear feels.
	Footwear acceptability	Acceptability of footwear, including choice and appearance of footwear, and the extent to which footwear choice impacts on clothing choices.
Treatment satisfaction	Treatment satisfaction	How good the patient feels the treatment received has been.

Table 6.4: Societal/resource use domains

Broad domain	Target domain	Definition
Health service use	Direct costs	Cost to health services as a result of foot and ankle problems, including medication use/co-interventions, consultations with healthcare providers, primary care contacts, outpatient appointments, A&E/ED visits, hospitalisation or re-admission, length of hospital stay, investigative procedures (tests to find out what the problem is), foot and ankle surgery, and complications following surgery (e.g. re-operation rate).
	Cost-effectiveness	The costs of healthcare treatment and the benefits to patients are balanced to determine which are the most effective, whilst being good value for money.
Personal expenses	Personal expenses	How much treatment and the delivery of treatment for foot and ankle disorders costs a patient (including consultations, insurance, medication costs, the purchase of devices such as orthotics and footwear, extra costs for transport due to limitations in walking, and reduction in salary due to the impact of foot and ankle disorders on work).

6.3.2 Ethics

Ethical approval was obtained from the University of Leeds Research Ethics Committee for School of Medicine (reference 2029) (Appendix L). The time taken to complete each round of the survey was deemed the only inconvenience to participants. At the end of the second round, patient participants were informed about an incentive for completing all rounds (prize draw to win one of 50 gift cards). As online surveys do

not involve direct contact with participants and their risks are considered to be minimal, consent was implied by completion and submission of the survey. The introductory text on the Delphi landing page contained the same information as a consent form (e.g. voluntary participation, risks, confidentiality/anonymity, and right to withdraw). Links to the PIS (Appendix M) and the University of Leeds Privacy Notice (344) were also provided. The data management plan for this study is presented in Appendix N.

6.3.3 Participants and recruitment

Recruitment materials are presented in Appendix O. Recruitment was targeted towards three key contributor groups: patients, HCPs, and researchers. Methods of identification and approach for each contributor group are provided below.

Patients

Patients were invited to participate through OMERACT, patient organisations, patient networks and social media campaigns. Patients who participated in previous phases of this research, and who consented to receive updates from the OMERACT Foot and Ankle Working Group, were also invited to participate by email. Additionally, PPI contributors within the PAG and OMERACT Foot and Ankle Working Group shared information about the Delphi through their patient networks.

HCPs

HCPs from different disciplines (e.g. medicine, podiatry, physiotherapy, prosthetics and orthotics, occupational therapy, nursing, orthopaedic surgery) with clinical experience of managing patients with RMDs who present with foot and ankle disorders, were invited to participate in the Delphi study. Members of the PAG/OMERACT Working Group who work clinically nominated suitable HCP representatives. Potential participants were invited to participate through direct, personalised emails, professional organisations, and social media campaigns.

Researchers

Clinical researchers, known by the PAG/OMERACT Working Group to have expertise in foot and ankle disorders in RMDs, were invited to participate. Researchers were identified and recruited by direct personalised contact, by PAG/OMERACT Working Group members based in research institutions, through professional organisation promotion (including through social media), and at relevant foot and ankle sessions/meetings affiliated with professional organisations. Multiple organisations and groups were approached during the recruitment phase; those confirming support and

methods of recruitment are presented in Table 6.5 (HCPs/researchers) and Table 6.6 (patients).

Table 6.5: Recruiting organisations (HCPs/researchers)

Organisation/group	**Country**	**Profession targeted**	**Recruitment methods**
OARSI International Foot and Ankle OA Consortium	International	HCPs/researchers of multiple professions with a special interest in foot and ankle OA	Emails sent to all members
OMERACT (professional network)	International	HCPs/researchers of multiple professions, industry representatives	Emails sent to all members; promotion in OMERACT newsletters
Royal College of Podiatry	UK	Podiatrists	Emails sent to all members
British Association of Prosthetists and Orthotists (BAPO)	UK	Orthotists	Emails sent to all members.
British Society for Rheumatology (BSR) Foot Health in IA Guideline Working Group	UK	HCPs/researchers of multiple professions	Emails sent to all members
Indian Rheumatology Society	India	Rheumatologists	Emails sent to all members
Japan College of Rheumatology	Japan	Rheumatologists	Emails sent to all members

Table 6.6: Recruiting organisations (patients)

Patient organisation/group	Country	Condition targeted	Recruitment methods
OMERACT PRP Network	International	Multiple RMDs	Emails sent to all members; promotion in OMERACT PRP newsletters
Scleroderma and Raynaud's UK	UK	SSc and RP	Social media posts
Enteropathic arthritis patient support group (Facebook)	International	Enteropathic arthritis	Social media posts
SpA patient support group	International	SpA	Social media posts
Foot arthritis patient support group	International	Foot OA	Social media posts

6.3.4 Eligibility criteria

Potential participants self-assessed their eligibility based on the following criteria, which was presented on the landing page of the Delphi.

Participants were eligible to participate if:

1) They were aged 18 or over;
2) They could read and understand English;
3) They were able to give informed consent;
4) They had access to relevant technology (e.g. digital device capable of loading and completing the online surveys, as well as email).

No limitations were applied regarding gender, ethnicity, socioeconomic grouping or geographic location. Specific eligibility criteria were also applicable depending on the contributor group. Patients were eligible to participate if they had sought treatment for a foot or ankle problem in the last 12 months and had a diagnosis of an RMD (as per the scope outlined in Chapter 2, section 2.9.1). Patients were ineligible to participate if they had diabetes, peripheral neuropathy, peripheral arterial disease, a primary neurological condition affecting the foot or ankle (e.g. multiple sclerosis), current acute trauma to the

foot or ankle (e.g. a fracture, rupture or sprain), or had experienced any of these issues in the last 12 months. The justification for these exclusion criteria was also discussed in Chapter 2 (section 2.9.1).

HCPs were eligible to participate if they worked in a clinical setting (where part of their role is employed by a clinical organisation, including private clinics) managing patients with foot and ankle problems in RMDs, and they had at least one year of experience treating adults with foot and ankle disorders in RMDs.

Researchers were eligible to participate if they had at least one year of experience working in an area relevant to foot and ankle disorders in RMDs.

6.3.5 Sampling

6.3.5.1 Sample size

As discussed in Chapter 2 (section 2.5.2.3), there is no standardised method for determining the sample size for a Delphi study, but higher numbers of Delphi participants have been shown to increase reliability of data (210, 345). The recruitment target for Round 1 was therefore 200; this was a pragmatic decision based on recruitment in previous Delphi studies for COS in RMDs and in previous foot and ankle surveys (320, 346).

6.3.5.2 Sampling technique

A convenience sampling approach was used, allowing any participants self-assessing as eligible to take part. Participants were asked to confirm their eligibility before starting Round 1 of the survey.

6.3.6 Distribution and approach

Recruitment materials directed potential participants to the Delphi landing page. Participants were asked to proceed to the next page of the survey if they agreed to take part. On the next page of the survey, participants were asked to provide their name and email address for further rounds of the survey, and invited to complete a short demographic questionnaire (gender, ethnicity, geographic location, type of RMD, and duration of foot and ankle problems (patients), professional background and years of relevant experience (other contributor groups)). On the following two pages of the survey, the domains were presented alongside their definitions and rating scales.

Upon the launch of Round 2 and Round 3 of the survey, each participant was sent an email with a new survey link and a summary of their own scores from the previous round, in Portable Document (PDF) format, as an attachment. Participants were asked to review their own scores alongside summaries of responses from both groups for each domain, which were presented within the survey itself, as they completed the survey round. Upon the launch of Round 4 of the survey, participants were sent an email with the new survey link, but were not asked to review any previous results.

6.3.7 Attrition

Attrition was discussed in depth in Chapter 2 (section 2.5.2.3). Multiple methods aiming to improve participant recruitment and retention, thus reduce attrition, were employed in the Delphi study (Table 6.7).

Table 6.7: Recruitment and retention methods

Potential issue affecting recruitment/ retention	**Description/evidence**	**COMFORT methods aiming to address this**
Survey length	Longer surveys (e.g. those with more Delphi domains) can affect recruitment and retention (213).	• The number of domains was reduced from an initial number of 137 domains to reduce the length of the survey and subsequent participant burden. Round 1 included 32 target domains, reducing to 31 domains in Rounds 2, 25 in Round 3, and ten domains in Round 4. • Any text deemed optional to read to understand the survey, rather than crucial, was attached via a PDF link to keep the main body of text as simple as possible. • Participants were informed of the anticipated time commitment in the PIS, and in each participation invitation.
Poor accessibility	Inaccessible surveys (e.g. surveys that are difficult to complete, confusing or ambiguous questions, dense text, certain fonts, small fonts	Accessibility guidelines (347, 348) were followed, for example: • A dyslexia-friendly (sans serif) font was used throughout the survey.

	and colours), and complicated login processes can deter potential participants (193).	• Graph colours were accessible to those with colour blindness. • All graph text was horizontal and labels were legible. • No shaded backgrounds, unnecessary borders, shadows, patterns, 3D shapes or thick/dark gridlines were used. • The survey did not involve a login process.
Lack of understanding regarding the overall purpose of the study, and what commitment involves	If participants do not understand the overall aim of the study, what taking part involves, and the study's relevance and benefits, they may not be motivated to participate/continue participating. Evidence suggests that Delphi participants can struggle to understand the concept of a COS (185).	• Following piloting, amendments were made to how the purpose of the study was presented, to improve clarity. • Round 1, and all invitations for subsequent rounds, gave an estimate of how long the survey would take to complete. • Reminder emails reiterated that the overall purpose of the study was to improve the evidence for foot and ankle problems, and that for the participant's views to count, it was important to complete all rounds. • Reminder emails sent to patients were co-produced with PPI and included a quote from a PPI contributor, detailing their motivation for being involved with the study (Figure 6.1).

		• Round 4 included examples of each domain within the context of a clinical trial. • A progress bar was included within each survey.
Poor readability/too much technical jargon	Text that is too difficult for participants to read and understand, particularly when it contains medical jargon, can lead to them abandoning participation (193). Many adults have limited health literacy levels and struggle to read and understand health information, particularly if it includes both text and numbers (221).	• The PIS and survey text was co-produced with PPI, written in plain language and piloted with patients. • Survey background text and instructions were entered into a validated readability checker, the Simplified Measure of Gobbledygoop (SMOG) (349) to ensure it was not classed as difficult or very difficult to read.
Generic initial recruitment and reminder emails	Previous Delphi studies have reported issues with spam filters in institutional emails when using Delphi-specific software to generate recruitment/reminder emails, leading to high dropout rates (350). Generic emails, such as those generated from a system, may also be easier to ignore.	• All invitation emails were personalised and sent individually, directly from the candidate's email address to avoid spam filters. • The written text in reminder emails was tailored according to RMD (patients) and profession (HCPs) (examples in Appendix P). • Members of the PAG who knew non-responding HCPs/researchers personally made direct contact.

	Personalised communication has been shown to improve recruitment and retention in Delphi studies conducted as part of the COS development process (193). Showing endorsement from influential individuals (e.g. distinguished researchers in the field) is recommended method of recruitment (65).	• Participants were thanked for their time in every invitation/reminder email, and at the end of each survey round. The invitation to participate in the final round included a personal thank you from the candidate, for their contribution to her PhD.
Lack of incentive	Whilst altruism is a motivating factor to take part in research for some participants, a Delphi study involves multiple rounds of seemingly repetitive surveys. This can become boring and burdensome for participants (351), and they may lose interest without an incentive.	• During the invitations to complete the third and fourth round of the survey, patient participants were informed that by completing all rounds of the study, they would be entered into a prize draw to win one of 50 gift cards (with the option to opt out). • Similarly, during the invitations to complete the second and third round of the survey, HCP/researcher participants were informed that by completing all rounds of the study, they would be acknowledged by name in the resulting publication (with the option to opt out).

	Acknowledging participants who complete all Delphi rounds in any publications arising from the study is recommended (65).	
Rigid survey closure dates	Inflexible closure dates may result in dropout of participants who would have completed the survey when it was overdue. Extending the closure date of the Delphi is recommended (65).	• The deadline for each round was extended by up to 48 hours for participants who were unable to complete the survey by the deadline to accommodate participants' schedules (193)

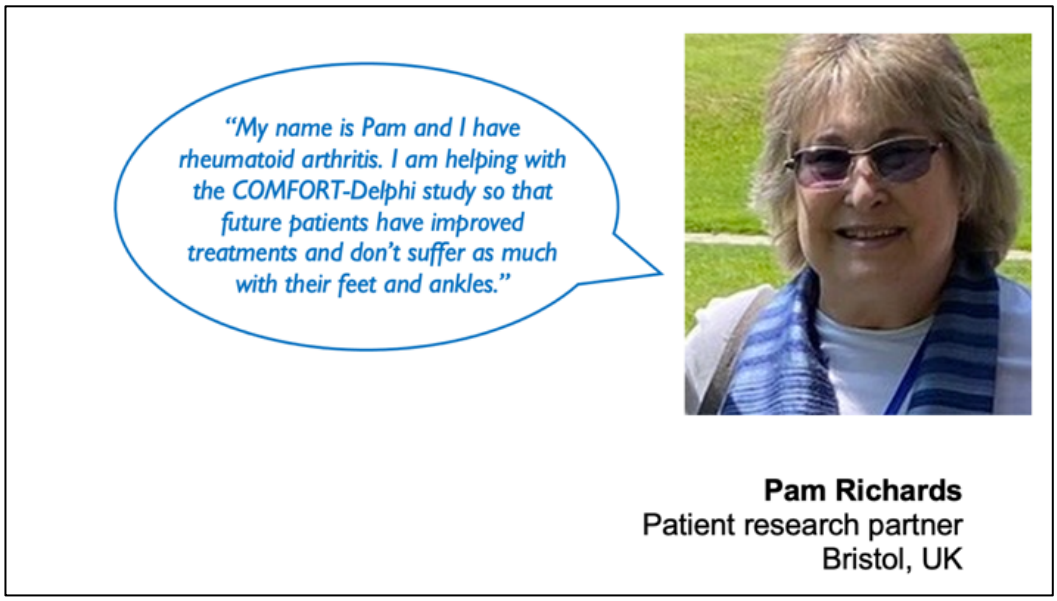


Figure 6.1: Example reminder email to patient participants

Consent was obtained from PR for the inclusion of this photograph and quote.

6.3.7.1 Data collection

Variations in the number of rounds and different scales for Delphi studies were discussed in Chapter 2 (section 2.5.2.3). In the current study, data were collected using four rounds of online surveys, through secure survey software approved by the University of Leeds (Jisc OnlineSurveys, version 3). Other software was considered for this study, including DelphiManager, the web based system specifically designed by COMET to build and manage Delphi surveys (204). However, OnlineSurveys was selected for this Delphi so that the candidate could lead the design of the surveys and survey feedback, and explore participants' experiences of these aspects. An overview of the first round of the survey is presented in Appendix Q, including the landing page, introduction to domain ratings, the first domain rating scale (pain intensity), and the free-text feedback question.

In Rounds 1-3, participants were asked to rate, on a scale of 1-9, the importance of each domain to be included in the core domain set. Scores of 1-3 corresponded to not important (with 1 denoting the least important), 4-6 to important but not a priority, and 7-9 to critically important and a priority for inclusion in a core domain set (with 9 denoting the most important). The type of scale selected, and the number of rounds, followed established COS methodology (63). Domains were presented according to the core areas of the OMERACT Filter (228), as outlined in Chapter 2 (section 2.6). As changes were made to domains and definitions in Round 2, conducting a third round

permitted all domains to go through at least two rounds, so that participants could consider their own and others' responses and re-rate accordingly.

In Round 1 of data collection, the survey included a free-text box after each domain, where participants were able to comment and suggest domain amalgamations, additions and modifications to wording/definitions, and make any other comments. Free-text responses were analysed by the candidate, and presented and discussed in a consensus workshop in preparation for Round 2 of data collection. All statements, with relevant amendments/additions if deemed appropriate by the PAG were taken into Round 2 of data collection. In Round 2, there were no free-text response boxes available.

After completing Round 1, participants were provided with a summary of results and asked to re-rate the importance of domains using the same numerical scale. The same process was then repeated for Round 3. As multiple domains (> 7) met the consensus threshold for critical importance following Round 3, a fourth Delphi round took place to prioritise this list of domains. In Round 4, participants were presented with domains, their definitions, and examples of each domain in the context of a hypothetical clinical trial (Appendix R), and asked to select each domain as *in* or *out* (with regards to inclusion in the final core domain set), and then rank domains in order of importance.

Participants were also able to comment on the reasons for their inclusion and exclusion decisions. The order of domains within the ranking exercise was randomised to reduce the potential for survey bias (190). At the end of the survey, participants were invited to provide anonymous feedback on their experience of taking part in the Delphi study, to inform future researchers who are conducting Delphi studies as part of COS development.

Each round of the survey was open for two weeks (with an additional extension of up to 48 hours) to allow for wide promotion. Round 1 opened on 9th April 2025 with a closing date of 25th April. Round 2 opened on 29th May 2025 with a closing date of 14th June. Round 3 opened on 3rd July 2025 with a closing date of 19th July. Finally, Round 4 opened on 20th August 2025 with a closing date of 5th September 2025. In each round, a personalised email reminder was sent after ten days to each non-responder. Only one formal reminder was sent per participant, to avoid potential harassment (352). Non-responders were excluded from subsequent rounds. There was a period of approximately four weeks between each round to allow for analysis of responses and survey design amendments.

6.3.7.2 Piloting

The first round of the Delphi was piloted by the PAG, two HCPs with limited research experience, and four members of the Leeds NIHR BRC PPI group. Piloting involved reviewing the presentation and functioning of the survey, including the PIS and domain definitions. Key changes made as a result of pilot feedback are presented in Table 6.8.

Table 6.8: Changes made following piloting

Feedback (verbatim comment)	**Change made**
"It is difficult to see if I am eligible or not."	Changes made to the layout of the survey to make the eligibility section clearer.
"The purpose of the study is not clear."	Survey wording relating to the aim of the study amended for clarity; changes made to the layout of the survey to make the purpose section clearer.
"It is not clear what a domain actually is, or consensus."	Changes made to the wording of the survey to clarify the meanings of 'domain' and 'consensus'.
"All of the domains are important and could be core."	Wording of the survey changed to clarify that the end result of the study will be 5-7 core domains that are most important.

Round 2 and Round 4 of the survey were also piloted with the PAG. No significant changes were made to the survey between Rounds 2 and 3, so Round 3 was not piloted. The aim of piloting Round 2 was to gain feedback on the presentation of the summaries of results from Round 1, whilst the Round 4 pilot sought to ensure the survey had a logical flow, examples for each domain were understandable, and that the ranking exercise worked. As a result of piloting, the labels on graphs in Round 2 were made clearer, and the wording of the examples in Round 4 was amended to clarify meaning.

6.3.7.3 Consensus workshop

A 90-minute consensus workshop was held with the wider OMERACT community on 16th May 2025, between Round 1 and Round 2 of the Delphi study. In this workshop,

the candidate presented findings from previous qualitative work and from a content analysis of participants' free-text responses from Round 1. Discussions around the suggested domain additions, combinations, and reductions were held to inform the development of the survey for Round 2. The quantitative results from Round 1 were not presented or discussed in the consensus workshop. This aimed to minimise bias by ensuring decisions were based directly on participants' suggestions and not influenced by overall scores.

6.3.7.4 Final PAG meeting

A final meeting with the PAG was held online on 8th September 2025. The results from Round 4 were the focus of the meeting, but results from other rounds and from the previous qualitative studies and scoping review were also drawn upon to inform discussions. The final core domain set was agreed in this meeting.

6.3.8 Data analysis

6.3.8.1 Quantitative data

Domain ratings (Rounds 1-3)

Participants' responses were analysed descriptively using Microsoft Excel (version 16.66.1), using frequency distributions (n, %). Each participant's own score and collective responses were fed back to participants after each round, allowing them to consider the views of others before re-rating each domain. Following established methodology (63), data were analysed and presented to participants by two groups, patients and non-patients, using bar graphs. Consensus definitions were discussed in Chapter 2 (section 2.5.2.3). Consensus for Rounds 1-3 in this study was defined according to the criteria in Table 6.9. These criteria are recommended by OMERACT (63).

Table 6.9: Consensus criteria for Rounds 1-3

Definition	Criteria
Consensus that a domain is critically important for a core domain set	≥ 70% of participants in both groups (patients and HCP/researchers) scored the item as critically important (score 7 to 9); after two rounds, these domains were acknowledged as having met criteria for importance to a core domain set and were held for final rating.
Consensus that a domain will not be included	≥ 70% of participants in both groups scored the item as "not important" (score 1 to 3); these domains were not taken to the subsequent round.
Dissensus but important to one group.	≥ 70% participants in one group scored items as critically important (score 7 to 9), these domains continued to the next round as having no consensus yet; if the domain did not reach consensus level at end of the Delphi, but was still important to one group, it became a potential "circumstance-specific core domain" or a "domain for future consideration."
No consensus	For all other results, the domain continued to the next round as having no consensus yet. If the domain did not achieve consensus by the end of the third round, and no groups had supported it ≥ 70%, then the domain was not endorsed for the core domain set.

Final domain ratings (Round 4)

Responses in the final round were analysed by percentage of participants (patients, other contributors, and overall) selecting each domain as 'in'. Domains where ≥ 70% of both groups combined rated the domain as 'in' achieved consensus for inclusion as core domains. The ranking exercise was analysed by the number of participants ranking each domain in each place (1 - most important to 10 - least important). A rating of 1 corresponded to a score of 10, 2 to 9, 3 to 8, etc. The mean score for each domain was then calculated. The ranking exercise did not have a consensus threshold; results were used to aid decision-making for the final core domain set.

6.3.8.2 Qualitative data

Free-text responses (Round 1)

An overview of qualitative analysis methods was provided in Chapter 2 (Table 2.5). Directed content analysis was used to organise free-text responses in the first round of the Delphi to elicit meaning from this data. Thematic analysis was also considered as a potential method for analysing free-text responses, but content analysis was deemed most appropriate as it can include quantification of qualitative data (312). The directed content analysis of free-text responses aimed to identify any domains that could be combined or dropped, and additional domains that should be included in the second round. Additionally, this analysis aimed to identify any wording issues in the survey, so that amendments could be made to improve readability in subsequent rounds. All free-text responses were imported and analysed in Microsoft Excel. A pre-existing framework with four main categories (domain combinations, domain additions, domains to drop, wording) guided the analysis. The analysis involved identifying meaning units (MUs) (words, sentences or paragraphs containing aspects related to each other through their content and context) (353).

The candidate undertook all steps of each content analysis. Firstly, each free-text comment was read and given an inductively generated code. Codes that did not relate to the aims of the analysis (e.g. irrelevant comments) were discarded. Similar codes were grouped together and assigned to one of the predefined categories. Codes that did not fit within the pre-specified categories were grouped into an ‘other’ category. To facilitate trustworthiness, all coding and categorisation decisions were verified by at least one of the candidate’s supervisors (HJS, CAF). Verbatim quotations from the free-text responses were provided to support analysis decisions where appropriate.

Free-text responses (Round 4)

The first free-text content analysis in Round 4 aimed to inform final consensus on the core domain set. A pre-existing framework with each of Round 4 domains as categories guided the analysis. The content analysis methods employed are described fully in section 6.3.8.2. The second free-text analysis in Round 4 aimed to explore positive and negative aspects of participants’ experiences of completing the Delphi surveys, to inform the design of future COS studies. In the second Round 4 content analysis, categories were developed inductively.

Results

In total, 206 participants (84 patients) completed Round 1 of the Delphi survey, with 149 (57 patients), 130 (50 patients), and 126 (49 patients) completing Rounds 2, 3 and 4, respectively. The flow of participants through the study is presented in (Figure 6.2). Table 6.10 presents the characteristics of those who participated in each round of the Delphi survey.

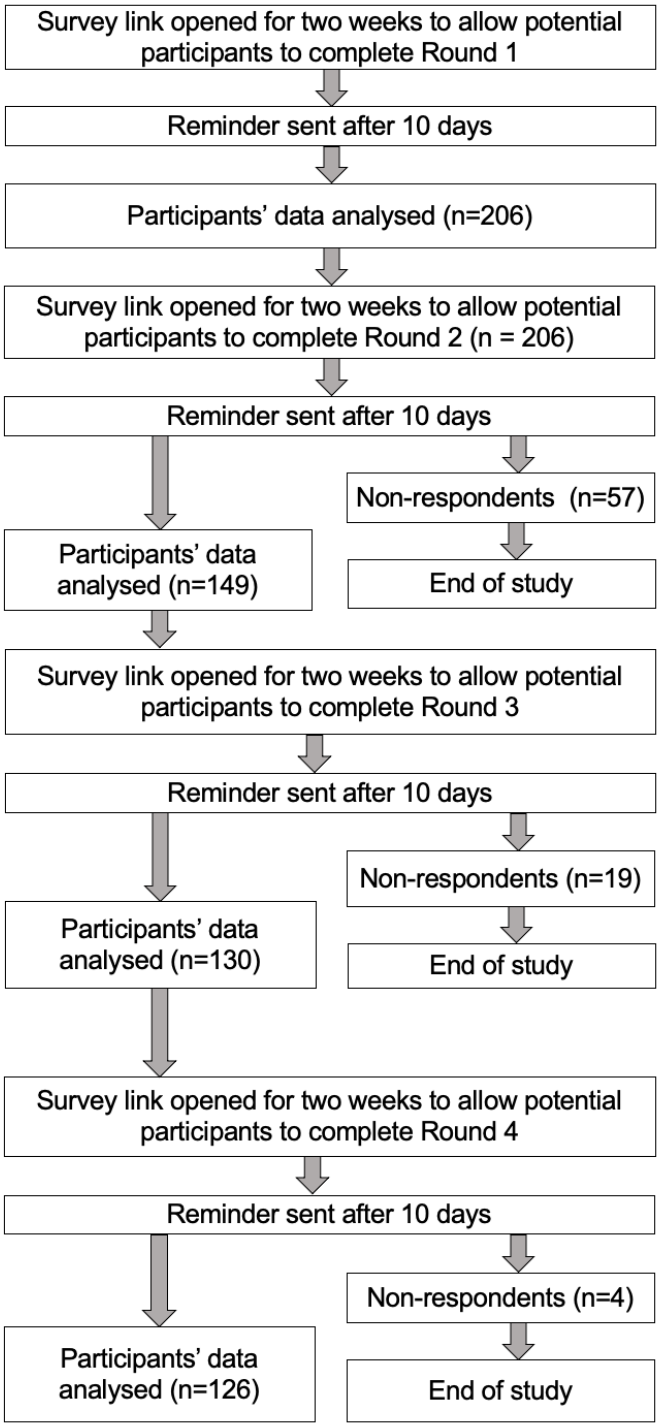


Figure 6.2: Flow of participants

Table 6.10: Participant characteristics

Variables	**Round 1, n (%)**		**Round 2, n (%)**		**Round 3, n (%)**		**Round 4, n (%)**	
	Patients (n = 84)	**HCPs/ researchers** (n = 122)	**Patients** (n = 57)	**HCPs/ researchers** (n = 92)	**Patients** (n = 50)	**HCPs/ researchers** (n = 80)	**Patients** (n = 49)	**HCPs/ researchers** (n = 77)
Gender[a]								
Woman	64	51	41	40	36	34	35	33
Man	20	71	15	52	14	46	14	44
Transgender	1	0	1	0	0	0	0	0
Non-binary	1	0	1	0	0	0	0	0
Age category								
18-29	3 (4%)	4 (3.3%)	1 (1.8%)	3 (3.3%)	1 (2.0%)	3 (3.8%)	1 (2.0%)	3 (3.9%)
30-39	10 (12%)	33 (27.0%)	7 (12.3%)	22 (23.9%)	5 (10.0%)	20 (25.0%)	5 (10.2%)	18 (23.4%)
40-49	13 (15%)	51 (41.8%)	12 (21.1%)	38 (41.3%)	10 (20.0%)	30 (37.5%)	9 (18.4%)	30 (39.0%)
50-59	28 (33%)	23 (18.9%)	16 (28.1%)	20 (21.7%)	16 (32.0%)	18 (22.5%)	16 (32.7%)	17 (22.1%)
60-69	16 (19%)	8 (6.6%)	8 (14.0%)	6 (6.5%)	8 (16.0%)	6 (7.5%)	8 (16.3%)	6 (7.8%)
70-79	13 (15%)	3 (2.5%)	12 (21.1%)	3 (3.3%)	9 (18.0%)	3 (3.8%)	9 (18.4%)	3 (3.9%)
80+	1 (1%)	0 (0.0%)	1 (1.8%)	0 (0.0%)	1 (2.0%)	0 (0.0%)	1 (2.0%)	0 (0.0%)

Ethnicity								
East Asian	1 (1%)	2 (1.6%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
South Asian	1 (1%)	30 (24.6%)	1 (1.8%)	16 (17.4%)	1 (2.0%)	11 (13.8%)	1 (2.0%)	10 (13.0%)
Southeast Asian	0 (0%)	2 (1.6%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Latin American	2 (2%)	2 (1.6%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
North African	15 (18%)	9 (7.4%)	13 (22.8%)	6 (6.5%)	13 (26.0%)	6 (7.5%)	14 (28.6%)	6 (7.8%)
South Pacific	4 (5%)	9 (7.4%)	1 (1.8%)	9 (9.8%)	1 (2.0%)	9 (11.3%)	1 (2.0%)	8 (10.4%)
White	60 (71%)	63 (53.3%)	41 (71.9%)	55 (59.8%)	35 (70.0%)	48 (60.0%)	33 (67.3%)	47 (61.0%)
Multiracial	0 (0%)	2 (1.6%)	0 (0.0%)	2 (2.2%)	0 (0.0%)	2 (2.5%)	0 (0.0%)	2 (2.6%)
Preferred not to answer	1 (1%)	1 (0.8%)	1 (1.8%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
First language								
English	62 (73.8%)	74 (60.7%)	43 (75.4%)	65 (70.7%)	36 (72.0%)	57 (71.3%)	35 (71.4%)	55 (71.4%)
Arabic	13 (15.5%)	5 (4.1%)	11 (19.3%)	4 (4.3%)	11 (22.0%)	4 (5.0%)	11 (22.4%)	4 (5.2%)
Bengali	0 (0.0%)	15 (12.3%)	0 (0.0%)	6 (6.5%)	0 (0.0%)	3 (3.8%)	0 (0.0%)	2 (2.6%)
Hindi	0 (0.0%)	3 (2.5%)	0 (0.0%)	3 (3.3%)	0 (0.0%)	3 (3.8%)	0 (0.0%)	3 (3.9%)
Japanese	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Spanish	1 (1.2%)	5 (4.1%)	0 (0.0%)	4 (4.3%)	0 (0.0%)	4 (5.0%)	0 (0.0%)	4 (5.2%)

Ukrainian	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Dutch	3 (3.6%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Greek	0 (0.0%)	2 (1.6%)	0 (0.0%)	2 (2.2%)	0 (0.0%)	2 (2.5%)	0 (0.0%)	2 (2.6%)
Telugu	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Malayalam	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Italian	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Multiple[b]	5 (6.0%)	12 (9.8%)	3 (5.3%)	4 (4.3%)	3 (6.0%)	4 (5.0%)	3 (0.0%)	4 (5.2%)
Country of response								
Europe								
UK	36 (42.9%)	48 (39.3%)	26 (45.6%)	43 (46.7%)	23 (46.0%)	36 (45.0%)	23 (46.9%)	35 (45.5%)
Republic of Ireland	2 (2.4%)	0 (0.0%)	1 (1.8%)	0 (0.0%)	1 (2.0%)	0 (0.0%)	1 (2.0%)	0 (0.0%)
Serbia	1 (1.2%)	0 (0.0%)	1 (1.8%)	0 (0.0%)	1 (2.0%)	0 (0.0%)	1 (2.0%)	0 (0.0%)
The Netherlands	3 (3.6%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Spain	0 (0.0%)	3 (2.5%)	0 (0.0%)	3 (3.3%)	0 (0.0%)	3 (3.8%)	0 (0.0%)	3 (3.9%)
Italy	0 (0.0%)	3 (2.5%)	0 (0.0%)	3 (3.3%)	0 (0.0%)	2 (2.5%)	0 (0.0%)	2 (2.6%)
Germany	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Ukraine	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Cyprus	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
North America								
USA	16 (19.0%)	3 (2.5%)	9 (15.8%)	3 (3.3%)	6 (12.0%)	3 (3.8%)	6 (12.2%)	3 (3.9%)
Canada	5 (6.0%)	2 (1.6%)	5 (8.8%)	2 (2.2%)	5 (10.0%)	1 (1.3%)	4 (8.2%)	1 (1.3%)
Puerto Rico	1 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Mexico	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Australasia								
Australia	5 (6.0%)	17 (13.9%)	2 (3.5%)	14 (15.2%)	1 (2.0%)	14 (17.5%)	1 (2.0%)	13 (16.9%)
New Zealand	0 (0.0%)	2 (1.6%)	0 (0.0%)	2 (2.2%)	0 (0.0%)	2 (2.5%)	0 (0.0%)	2 (2.6%)
Africa								
Egypt	14 (16.7%)	8 (6.6%)	12 (21.0%)	5 (5.4%)	12 (24.0%)	5 (6.3%)	12 (24.5%)	5 (6.5%)
Morocco	0 (0.0%)	1 (0.8%)	0 (0.0%)	1 (1.1%)	0 (0.0%)	1 (1.3%)	0 (0.0%)	1 (1.3%)
Asia								
Japan	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
India	0 (0.0%)	12 (9.8%)	0 (0.0%)	7 (7.6%)	0 (0.0%)	7 (8.8%)	0 (0.0%)	7 (9.1%)
Bangladesh	0 (0.0%)	16 (13.1%)	0 (0.0%)	6 (6.5%)	0 (0.0%)	3 (3.8%)	0 (0.0%)	2 (2.6%)
Saudi Arabia	1 (1.2%)	0 (0.0%)	1 (1.8%)	0 (0.0%)	1 (2.0%)	0 (0.0%)	1 (2.0%)	0 (0.0%)

Philippines	0 (0.0%)	1 (0.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Patient condition								
Condition duration (years)								
Mean (SD), median, range	25.4 (14.1) 10.0, 1-75		16.3 (14.9) 10.0, 1-75		15.6 (13.1) 10.5, 1-52		15.8 (13.1) 11.0, 1-52	
Multiple RMDs*	23 (27.4%)		13 (22.8%)		11 (22.0%)		10 (20.4%)	
Inflammatory arthritis								
RA	20 (23.8%)		15 (26.3%)		15 (30.0%)		15 (30.6%)	
AS	3 (3.6%)		1 (1.8%)		1 (2.0%)		1 (2.0%)	
PsA	1 (1.2%)		1 (1.8%)		1 (2.0%)		1 (2.0%)	
Enteropathic arthritis	5 (6.0%)		4 (7.0%)		2 (4.0%)		2 (4.1%)	
JIA	3 (3.6%)		2 (3.5%)		1 (2.0%)		1 (2.0%)	
OA								
Foot or ankle OA	6 (7.1%)		4 (7.0%)		3 (6.0%)		3 (6.1%)	
Crystal arthropathies								
CPPD	1 (1.2%)		1 (1.8%)		1 (2.0%)		1 (2.0%)	

Systemic autoimmune conditions								
SLE	3 (3.6%)		2 (3.5%)		2 (4.0%)		2 (4.1%)	
SSc	7 (8.3%)		4 (7.0%)		3 (6.0%)		3 (6.1%)	
Sjogren's	1 (1.2%)		0 (0.0%)		0 (0.0%)		0 (0.0%)	
MSK disorders								
Achilles tendinopathy	7 (8.3%)		6 (10.5%)		6 (12.0%)		6 (12.2%)	
PHP	3 (3.6%)		3 (5.3%)		3 (6%)		3 (6.1%)	
Hallux valgus	1 (1.2%)		1 (1.8%)		1 (2.0%)		1 (2.0%)	
Professional background								
HCP		93 (76.2%)		64 (69.6%)		53 (66.3%)		51 (66.2%)
Researcher		27 (22.1%)		26 (28.3%)		25 (31.3%)		24 (31.2%)
Joint clinical academic		2 (1.6%)		2 (2.2%)		2 (2.5%)		2 (2.6%)
HCP/researcher profession								
Rheumatologist		55 (45.1%)		34 (37.0%)		26 (32.5%)		25 (32.5%)
Podiatrist		44 (36.1%)		36 (39.1%)		35 (43.8%)		35 (45.5%)
Physiotherapist		10 (8.2%)		10 (10.9%)		10 (12.5%)		9 (11.7%)

Orthotist		5 (4.1%)		4 (4.3%)		3 (3.8%)		2 (2.6%)
GP		1 (0.8%)		1 (1.1%)		1 (1.3%)		1 (1.3%)
Physiatrist		1 (0.8%)		1 (1.1%)		0 (0.0%)		0 (0.0%)
Athletic trainer		1 (0.8%)		1 (1.1%)		1 (1.3%)		1 (1.3%)
Consultant in Sport and Exercise Medicine		1 (0.8%)		1 (1.1%)		0 (0.0%)		0 (0.0%)
Diagnostic radiographer		1 (0.8%)		1 (1.1%)		1 (1.3%)		1 (1.3%)
Unspecified/not applicable		3 (2.5%)		3 (3.3%)		3 (3.8%)		3 (3.9%)

AS, ankylosing spondylitis; CPPD, crystal pyrophosphate deposition disease; GP, general practitioner; HCP, healthcare professional; MSK, musculoskeletal; OA, osteoarthritis; PHP, plantar heel pain; PsA psoriatic arthritis; PTTD, posterior tibial tendon dysfunction; RA, rheumatoid arthritis; RP, Raynaud's phenomenon; SD, standard deviation; SLE, systemic lupus erythematosus; SSc, systemic sclerosis; UK, United Kingdom; USA, United States of America.

[a]Participants were able to select multiple genders.

[b]Multiple conditions are detailed in Appendix S.

6.3.9 Domain ratings (Round 1)

The consensus levels for each of the 32 domains, by group (patients, HCPs/researchers) are presented in Table 6.11. Domains achieving ≥ 70% consensus as critically important are highlighted green. Two domains were rated as critically important by ≥ 70% consensus in both groups: physical function (necessary activities and participation) and physical function (discretionary activities and participation). No other domains were rated as critically important by patients, but six other domains were rated as critically important by HCPs/researchers: pain intensity, pain when weightbearing, foot and ankle inflammation, joint damage/deformity, treatment satisfaction, and cost-effectiveness.

Table 6.11: Round 1 domain ratings

Domain	Not important		Important but not critical		Critically important[a]	
	Patients	HCPs/ researchers	Patients	HCPs/ researchers	Patients	HCPs/ researchers
Pathophysiological domains						
Pain intensity	6 (7%)	1 (1%)	28 (33%)	15 (12%)	50 (60%)	106 (87%)
Pain during weightbearing	4 (5%)	1 (1%)	24 (29%)	16 (13%)	56 (67%)	105 (86%)
Pain during non-weightbearing	21 (25%)	7 (6%)	25 (30%)	42 (34%)	38 (45%)	73 (60%)
Cramping	23 (27%)	27 (22%)	31 (37%)	67 (55%)	30 (36%)	28 (23%)
Overall fatigue	13 (15%)	13 (11%)	29 (35%)	63 (52%)	42 (50%)	46 (38%)
Foot/ankle-specific fatigue	11 (13%)	10 (8%)	31 (37%)	61 (50%)	42 (50%)	51 (42%)
Numbness	31 (37%)	15 (12%)	22 (26%)	37 (30%)	31 (37%)	70 (57%)
Foot/ankle inflammation	15 (18%)	2 (2%)	29 (35%)	15 (12%)	40 (48%)	105 (86%)
Joint damage/ deformity	18 (21%)	2 (2%)	25 (30%)	27 (22%)	41 (49%)	93 (76%)
Colour changes	29 (35%)	15 (12%)	32 (38%)	51 (42%)	23 (27%)	56 (46%)

Patient global	7 (8%)	3 (2%)	30 (36%)	36 (30%)	47 (56%)	83 (68%)
Clinician global	17 (20%)	9 (7%)	39 (46%)	51 (42%)	28 (33%)	62 (51%)
Joint ROM	13 (15%)	4 (3%)	27 (32%)	35 (29%)	44 (52%)	83 (68%)
Foot/ankle stiffness	14 (17%)	6 (5%)	35 (42%)	50 (41%)	35 (42%)	66 (54%)
Foot/ankle morning stiffness	13 (15%)	9 (7%)	28 (33%)	49 (40%)	43 (51%)	64 (52%)
Joint instability	15 (18%)	7 (6%)	28 (33%)	41 (34%)	41 (49%)	74 (61%)
Nail abnormalities	35 (42%)	28 (23%)	31 (37%)	60 (49%)	18 (21%)	34 (28%)
Presence of callosities	34 (40%)	17 (14%)	31 (37%)	77 (63%)	19 (23%)	28 (23%)
Foot/ankle ulceration	45 (54%)	8 (7%)	11 (13%)	27 (22%)	28 (33%)	87 (71%)
Sufficiency of blood supply	30 (36%)	10 (8%)	15 (18%)	31 (25%)	39 (46%)	81 (66%)
Normality of gait	15 (18%)	4 (3%)	27 (32%)	39 (32%)	42 (50%)	79 (65%)
Life impact domains						
Necessary activities and participation	6 (7%)	0 (0%)	13 (15%)	13 (11%)	65 (77%)	109 (89%)
Discretionary activities and participation	4 (5%)	1 (1%)	20 (24%)	24 (20%)	60 (71%)	97 (80%)

Emotional wellbeing	10 (12%)	4 (3%)	26 (31%)	36 (30%)	48 (57%)	82 (67%)
Body image	18 (21%)	11 (9%)	36 (43%)	60 (49%)	30 (36%)	51 (42%)
Sleep	19 (23%)	5 (4%)	16 (19%)	37 (30%)	49 (58%)	80 (66%)
Footwear comfort	5 (6%)	5 (4%)	25 (30%)	51 (42%)	54 (64%)	66 (54%)
Footwear acceptability	9 (11%)	8 (7%)	31 (37%)	61 (50%)	44 (52%)	53 (43%)
Treatment satisfaction	11 (13%)	3 (2%)	20 (24%)	28 (23%)	53 (63%)	91 (75%)
Societal/resource use domains						
Direct costs	11 (13%)	3 (2%)	28 (33%)	42 (34%)	45 (54%)	77 (63%)
Cost-effectiveness	12 (14%)	3 (2%)	35 (42%)	30 (25%)	37 (44%)	89 (73%)
Personal expenses	12 (14%)	3 (2%)	18 (21%)	46 (38%)	54 (64%)	73 (60%)

HCP, healthcare professional; ROM, range of motion.

[a] Cells highlighted green correspond to domains that have met or exceeded the 70% threshold for consensus.

6.3.10 Round 1 content analysis and consensus workshop

A total of 32 patients and 37 HCPs/researchers provided free-text responses relating to pathophysiological domains, whilst 24 patients and 22 HCPs/researchers provided free-text responses relating to life impact domains. Twenty-one participants (13 researchers, five patients, two HCPs, and one industry representative), including six members of the PAG, attended the workshop. Any suggestions that were not discussed in the consensus workshop due to time constraints were discussed by the PAG. Following the consensus workshop, the following domains were combined: physical function (necessary activities and participation) with physical function (discretionary activities and participation), and foot/ankle stiffness with foot/ankle morning stiffness. Further, one domain was added (falls), and no domains were dropped. Five domain definitions were amended as a result of discussions, and amendments to the survey instructions were made on the basis of feedback. An overview of the categories, sub-categories, and meaning units generated in the content analysis is presented in Table 6.12, along with details of any changes made to the design of the Round 2 survey following discussions in the consensus workshop. To maintain confidentiality, no personal identifiers are included in the results.

Table 6.12: Round 1 content analysis

Category	Sub-category	MU (patients)	MU (HCPs/ researchers)	Details of changes made
Domain combinations	Combine overall fatigue and foot/ankle-specific fatigue	2	3	Patients indicated that overall fatigue and foot/ankle-specific fatigue were separate domains. Patient quotes from the consensus workshop: "I always feel like I want to try and find a chair to sit down in, and somewhere to put my foot on." "At a certain point I just have to collapse, have to put my feet up and have to rest." **No domain combination made. Quotes added to the domain definition for foot/ankle-specific fatigue.**
	Combine pain domains	-	2	Researchers maintained that each target domain for pain would be measured separately in a trial. **No changes made.**

	Combine necessary and discretionary activities and participation	2	3	Patients felt that necessary and discretionary activities and participation could differ depending on the individual patient, therefore the two concepts should be combined in to a broader domain. **Domains combined into one domain: physical function (activities and participation)**
	Combine overall stiffness and morning stiffness	-	1	Patients suggested stiffness was a single domain, regardless of when it occurred. **Stiffness domains combined for Round 2.**
	Combine emotional wellbeing and body image	1	-	Patients recognised that there could be some overlap, but felt that emotional wellbeing and body image were two separate domains. **No changes made.**
Domain additions	Balance	3	4	Patients, HCPs, and researchers perceived balance as an important domain, but suggested it should be included as part of the gait domain. **Domain definition updated for normality of gait: Any differences from expected 'normal' walking, which may impact on other areas of the body and balance.**

	Pain duration	3	-	HCPs and researchers agreed that duration was a measure of pain, not a domain itself. **No changes made.**
	Falls	1	2	Patients agreed that falls are important and should be included in the Delphi study. **Falls added as a Round 2 domain.**
	Psychological condition of patient	1	-	Patients, HCPs, and researchers agreed this was covered by emotional wellbeing. **No changes made.**
	Quality of life	-	1	HCPs and researchers maintained that “quality of life” is too broad, overlapping with activities and participation. **No changes made.**
	Neuropathic pain, joint pain, soft tissue pain	-	1	Patients expressed that it was not always possible to distinguish between types of foot/ankle pain. **No changes made.**
	Muscle strength	-	1	Patients, HCPs, and researchers suggested that muscle strength was not as important as the other domains included in the Delphi study.

				No changes made.
	Independence/ reliance on others	-	1	Patients felt that independence/reliance on others was part of the activities and participation domain. **Domain definition for physical function (activities and participation) amended to include independence.**
	Impact of foot and ankle problems on other areas of the body	1	-	Patients agreed that the impact of foot and ankle problems on other areas of the body, such as the knees and hips, was important. HCPs and researchers felt this should be included in the domain definition for normality of gait. **Domain definition for normality of gait amended to: Any differences from expected 'normal' walking, which may impact on other areas of the body and balance.**
Domains to drop	Sleep (covered by pain)	-	1	The PAG noted that sleep could be influenced by pain but also numbness and temperature, based on the qualitative work underpinning this study. **No changes made.**
	Footwear comfort (covered by symptoms elsewhere)	2	-	Patients recognised that footwear comfort overlapped with pain and deformity, but felt that problems with footwear was a distinct domain that should be measured.

				No changes made.
	Treatment satisfaction	-	1	The PAG noted that treatment satisfaction is important to measure, particularly considering the extent of dissatisfaction with foot and ankle treatments among patients identified in qualitative research. **No changes made.**
Wording/ readability	Change physician-global to clinician-global as not all HCPs are physicians.	-	1	The PAG recognised that physician-global was the term used in previous research, but that clinician was a more appropriate term. **Domain name changed to clinician-global.**
Other	Difficulty rating domains that are important for some conditions but not for others / that aren't my condition	2	-	**Survey instructions amended to "Based on your personal experience, we would like to know how important you think each of the domains below are, on a scale of 1-9."**
	Comorbidities change domain importance	-	3	The PAG discussed how most influential comorbidities (e.g. diabetes and peripheral arterial disease) have been excluded from COS scope. **No changes made.**

MU, meaning unit; PAG project advisory group.

6.3.11 Domain ratings (Round 2)

A total of 149 participants (57 patients) completed Round 2. Results are presented in Table 6.13. Seven domains were rated as critically important (highlighted green) by ≥ 70% participants in both groups: pain intensity, pain during weightbearing, falls, physical function, emotional wellbeing, sleep, and personal expenses. Only one additional domain, footwear comfort, was rated as critically important by patients. Comparatively, eight domains (foot/ankle inflammation, joint damage/deformity, patient global, foot/ankle ulceration, sufficiency of blood supply, treatment satisfaction, direct costs, and cost-effectiveness) were rated as critically important by HCPs/researchers.

Table 6.13: Round 2 domain ratings

Domain	**Not important**		**Important but not critical**		**Critically important[a]**	
	Patients	HCPs/ researchers	Patients	HCPs/ researchers	Patients	HCPs/ researchers
Pain intensity	3 (5%)	1 (1%)	23 (23%)	6 (7%)	41 (72%)	85 (92%)
Pain during weightbearing	2 (4%)	1 (1%)	8 (14%)	3 (3%)	47 (82%)	88 (96%)
Pain during non-weightbearing	8 (14%)	4 (4%)	24 (42%)	40 (43%)	25 (44%)	48 (52%)
Cramping	10 (18%)	25 (27%)	33 (58%)	56 (61%)	14 (25%)	11 (12%)
Overall fatigue	6 (11%)	8 (9%)	21 (37%)	54 (59%)	30 (53%)	30 (33%)
Foot/ankle-specific fatigue	3 (5%)	6 (7%)	26 (46%)	50 (54%)	28 (49%)	36 (39%)
Numbness	14 (25%)	7 (8%)	22 (39%)	35 (38%)	21 (37%)	50 (54%)
Foot/ankle inflammation	10 (18%)	2 (2%)	11 (19%)	10 (11%)	36 (63%)	80 (87%)
Joint damage/ deformity	9 (16%)	1 (1%)	13 (23%)	15 (16%)	35 (61%)	76 (83%)
Colour changes	18 (32%)	16 (17%)	27 (47%)	48 (52%)	12 (21%)	28 (30%)
Patient global	2 (4%)	2 (2%)	20 (35%)	22 (24%)	35 (61%)	68 (74%)
Clinician global	8 (14%)	9 (10%)	32 (56%)	54 (59%)	17 (30%)	29 (32%)

Joint ROM	5 (9%)	2 (2%)	18 (32%)	25 (27%)	34 (60%)	65 (71%)
Foot/ankle stiffness[b]	3 (5%)	1 (1%)	19 (33%)	38 (41%)	35 (61%)	53 (58%)
Joint instability	7 (12%)	6 (7%)	16 (28%)	28 (30%)	34 (60%)	58 (63%)
Nail abnormalities	21 (37%)	29 (32%)	24 (42%)	56 (61%)	12 (21%)	7 (8%)
Presence of callosities	17 (30%)	21 (23%)	30 (53%)	61 (66%)	10 (18%)	10 (11%)
Foot/ankle ulceration	21 (37%)	3 (3%)	9 (16%)	19 (21%)	27 (47%)	70 (76%)
Sufficiency of blood supply	15 (26%)	4 (4%)	9 (16%)	19 (21%)	33 (58%)	69 (75%)
Normality of gait	5 (9%)	3 (3%)	18 (32%)	31 (34%)	34 (60%)	58 (63%)
Falls[c]	10 (18%)	6 (7%)	7 (12%)	21 (23%)	40 (70%)	65 (71%)
Physical function (activities and participation)[d]	3 (5%)	1 (1%)	6 (11%)	4 (4%)	48 (84%)	87 (95%)
Emotional wellbeing	6 (11%)	1 (1%)	11 (19%)	24 (26%)	40 (70%)	67 (73%)
Body image	11 (19%)	11 (12%)	28 (49%)	64 (70%)	18 (32%)	17 (18%)
Sleep	6 (11%)	2 (2%)	10 (18%)	22 (24%)	41 (72%)	68 (74%)
Footwear comfort	3 (5%)	2 (2%)	13 (23%)	31 (34%)	41 (72%)	59 (64%)
Footwear acceptability	5 (9%)	2 (2%)	27 (47%)	49 (53%)	25 (44%)	41 (45%)

Treatment satisfaction	3 (5%)	4 (4%)	16 (28%)	12 (13%)	38 (67%)	76 (83%)
Direct costs	6 (11%)	1 (1%)	22 (39%)	24 (26%)	29 (51%)	67 (73%)
Cost-effectiveness	5 (9%)	2 (2%)	20 (35%)	20 (22%)	32 (56%)	70 (76%)
Personal expenses	2 (4%)	1 (1%)	12 (21%)	27 (29%)	43 (75%)	64 (70%)

[a] Cells highlighted green correspond to domains that have met or exceeded the 70% threshold for consensus.

[b]Domain amended after amalgamation of domains following content analysis/Round 1 consensus workshop.

[c]Domain added following content analysis/Round 1 consensus workshop.

6.3.12 Domain ratings (Round 3)

A total of 130 participants (50 patients) completed Round 3. Findings from Round 3 are shown in Table 6.14. Seven domains achieved consensus (highlighted green) between both groups at the end of Round 2. Six of these were held for final rating and not presented in Round 3. Falls was also rated as critically important by ≥ 70% participants in both groups in Round 2, but as it was a newly added domain, it was presented in Round 3 for further rating. Four domains were rated as critically important by ≥ 70% participants in both groups: joint ROM, normality of gait, footwear comfort, treatment satisfaction. No additional domains were rated as critically important by patients, but not by HCPs/researchers. In contrast, ten domains were rated as critically important by HCPs/researchers, but not by patients: foot/ankle inflammation, joint damage/deformity, patient global, foot/ankle stiffness, joint instability, foot/ankle ulceration, sufficiency of blood supply, falls, direct costs, and cost-effectiveness.

Table 6.14 Round 3 domain ratings

Domain	**Not important**		**Important but not critical**		**Critically important[a]**	
	Patients	HCPs/ researchers	Patients	HCPs/ researchers	Patients	HCPs/ researchers
Pain intensity	Domain held for final consensus					
Pain during weightbearing	Domain held for final consensus					
Pain during non-weightbearing	6 (12%)	1 (1%)	16 (32%)	27 (34%)	28 (56%)	52 (65%)
Cramping	10 (20%)	13 (16%)	29 (58%)	58 (73%)	11 (22%)	9 (11%)
Overall fatigue	5 (10%)	4 (5%)	17 (34%)	45 (56%)	28 (56%)	31 (39%)
Foot/ankle-specific fatigue	7 (14%)	2 (3%)	13 (26%)	42 (53%)	30 (60%)	36 (45%)
Numbness	13 (26%)	5 (6%)	19 (38%)	30 (38%)	18 (36%)	45 (56%)
Foot/ankle inflammation	7 (14%)	1 (1%)	12 (24%)	4 (5%)	31 (62%)	75 (94%)
Joint damage/deformity	8 (16%)	0 (0%)	11 (22%)	5 (6%)	31 (62%)	75 (94%)
Colour changes	16 (32%)	11 (14%)	26 (52%)	46 (58%)	8 (16%)	23 (29%)
Patient global	4 (8%)	1 (1%)	14 (28%)	9 (11%)	32 (64%)	70 (88%)
Clinician global	6 (12%)	3 (4%)	31 (62%)	59 (74%)	13 (26%)	18 (23%)

Joint ROM	5 (10%)	1 (1%)	10 (20%)	17 (21%)	35 (70%)	62 (78%)
Foot/ankle stiffness	5 (10%)	1 (1%)	12 (24%)	22 (28%)	33 (66%)	57 (71%)
Joint instability	7 (14%)	1 (1%)	9 (18%)	19 (24%)	34 (68%)	60 (75%)
Nail abnormalities	23 (46%)	18 (23%)	21 (42%)	56 (70%)	6 (12%)	6 (8%)
Presence of callosities	19 (38%)	14 (18%)	27 (54%)	56 (70%)	4 (8%)	10 (13%)
Foot/ankle ulceration	18 (36%)	1 (1%)	5 (10%)	11 (14%)	27 (54%)	68 (85%)
Sufficiency of blood supply	11 (22%)	1 (1%)	8 (16%)	10 (13%)	31 (62%)	69 (86%)
Normality of gait	5 (10%)	1 (1%)	7 (14%)	19 (24%)	38 (76%)	60 (75%)
Falls	6 (12%)	0 (0%)	11 (22%)	11 (14%)	33 (66%)	69 (86%)
Physical function	Domain held for final consensus					
Emotional wellbeing	Domain held for final consensus					
Body image	11 (22%)	9 (11%)	24 (48%)	58 (73%)	15 (30%)	13 (16%)
Sleep	Domain held for final consensus					
Footwear comfort	2 (4%)	0 (0%)	8 (16%)	14 (18%)	40 (80%)	66 (83%)
Footwear acceptability	7 (14%)	1 (1%)	21 (42%)	50 (63%)	22 (44%)	29 (36%)
Treatment satisfaction	2 (4%)	1 (1%)	6 (12%)	8 (10%)	42 (84%)	71 (89%)

Direct costs	4 (8%)	3 (4%)	15 (30%)	16 (20%)	31 (62%)	61 (76%)
Cost-effectiveness	4 (8%)	3 (4%)	12 (24%)	11 (14%)	34 (68%)	66 (83%)
Personal expenses	Domain held for final consensus					

[a] Cells highlighted green correspond to domains that have met or exceeded the 70% threshold for consensus.

6.3.13 Domain ratings (Round 4)

A total of 126 participants (49 patients) completed the in/out and ranking exercise in Round 4. Ten domains achieved consensus from ≥ 70% of both groups in Round 2 or 3 and were taken into to the fourth and final round of the Delphi (Table 6.15).

6.3.13.1 In/out exercise

Four domains were voted into the core domain set by ≥ 70% of both groups: pain intensity, physical function, pain when weightbearing, and treatment satisfaction. Joint ROM was voted into the core domain set by ≥ 70% of the patient group only, but had an overall score of ≥ 70%. Two other domains, normality of gait and sleep, were voted into the core domain set by ≥ 70% of the patient group, but not by the HCP/researcher group, and did not achieve an overall score of ≥ 70%.

Table 6.15: Round 4 in/out domain ratings

Domain[a]	**Overall %** (n = 126)	**Weighted average (1:1), %**	**Patients** (n = 49)	**HCPs / researchers** (n = 77)
Pain intensity	97%	97%	46 (94%)	76 (99%)
Physical function	96%	96%	45 (92%)	76 (99%)
Pain when weightbearing	87%	87%	42 (86%)	67 (87%)
Treatment satisfaction	75%	74%	35 (71%)	59 (77%)
Joint ROM	73%	76%	42 (86%)	50 (65%)
Footwear comfort	63%	63%	30 (61%)	49 (64%)
Emotional wellbeing	62%	62%	29 (59%)	49 (64%)
Normality of gait	59%	61%	36 (73%)	38 (49%)
Sleep	59%	62%	37 (76%)	37 (48%)
Personal expenses	34%	35%	17 (35%)	26 (34%)

[a] Cells highlighted green correspond to domains that met or exceeded the 70% threshold for consensus.

6.3.13.2 Ranking exercise

Results from the Round 4 ranking exercise are displayed in Table 6.16. Both groups (patients and HCP/researchers) ranked pain intensity as the most important domain, whilst pain when weightbearing, physical function and joint ROM were ranked within the top 4. Both groups ranked personal expenses as the least important domain.

Table 6.16: Round 4 domain rankings

Patient rankings			HCP/researcher rankings			Overall rankings		
Rank	**Domain**	**Mean score**	**Rank**	**Domain**	**Mean score**	**Rank**	**Domain**	**Mean score**
1	Pain intensity	8.72	1	Pain intensity	9.4	1	Pain intensity	9.06
2	Pain when weightbearing	7.74	2	Physical function	8.21	2	Pain when weightbearing	7.74
3	Physical function	6.54	3	Pain when weightbearing	7.74	3	Physical function	7.375
4	Joint ROM	6.22	4	Joint ROM	5.27	4	Joint ROM	5.745
5	Sleep	5.48	5	Emotional wellbeing	5.2	5	Emotional wellbeing	4.815
6	Normality of gait	5.07	6	Treatment satisfaction	4.64	6	Sleep	4.625
7	Emotional wellbeing	4.43	7	Footwear comfort	4.5	7	Normality of gait	4.54
8	Footwear comfort	4.28	8	Normality of gait	4.01	8	Treatment satisfaction	4.46
9	Treatment satisfaction	4.28	9	Sleep	3.77	9	Footwear comfort	4.39
10	Personal expenses	2.24	10	Personal expenses	2.24	10	Personal expenses	2.24

6.3.14 Content analysis (Round 4)

6.3.14.1 Decisions regarding inclusion/exclusion

In the pathophysiological section of Round 4, 23 HCPs/researchers and 14 patients provided free-text responses to elaborate on their decisions for including or excluding domains from the final core domain set. In the life impact section, 22 HCPs/researchers and 12 patients provided free-text responses. Finally, in the resource use section, 13 HCPs/researchers and 15 patients provided free-text responses. The number of meaning units corresponding to each sub-category for each Round 4 domain, with supporting verbatim quotes from the free-text responses, is presented in Table 6.17.

Table 6.17: Round 4 content analysis (domain decisions)

Domain	**Sub-category**	**MUs** (patients)	**MUs** (HCP / researchers)	**Verbatim free-text quote**
Pathophysiological manifestations				
Pain when weightbearing	Redundancy	1	1	“Pain intensity and pain during weightbearing reflect mostly similar assessments.” (HCP)
	Confounding factors	1	-	“[Pain when] weightbearing is a little bit relative in my experience, with the time I’ve realised that it depends on how you carry the weight, or your overall state of fatigue, or even if you are used to it or not, it’s something that you can cope with it either by avoiding it or with learn how to do it with a better.” (Patient)
Joint ROM	Not the focus of all interventions	-	2	“Joint range of motion is at the low end of my scale because often ROM is a fixed deformity, not caused by pain but by changes, so not so much affected by treatment. If pain is the limiting factor to ROM that will be picked up in the pain domains.” (Patient)

Domain	Sub-category	MUs (patients)	MUs (HCP / researchers)	Verbatim free-text quote
	Not important in all RMDs	1	3	“I think the importance varies depending on the nature of the condition, and so would not be considered mandatory in my mind.” (HCP)
	Not as important as other domains	-	3	“Even if you don't have full range [of motion] you may still be able to function well.” (HCP)
	Difficult to measure	-	3	“it is not easy enough to measure and compare.” (HCP)
	Redundancy	1	-	“If pain is the limiting factor to ROM that will be picked up in the pain domains.” (Patient)
Gait	Not important in all RMDs	-	4	“Whilst normality of gait is definitely important, I think the importance varies depending on the nature of the condition, and so would not be considered mandatory in my mind.” (HCP)
	Doesn’t always lead to life impact	1	1	“You can have a perfectly acceptable lifestyle and still ‘walk funny.’” (Patient)
	Not feasible/difficult to measure	2	5	“Gait abnormality/normality is very subjective and difficult to assess/quantify.” (HCP)

Domain	Sub-category	MUs (patients)	MUs (HCP / researchers)	Verbatim free-text quote
	Confounding factors	-	1	“Normality of gait depends on hip and knee joint as well as foot so may not be included.” (HCP)
	Redundancy with other domains	1	1	“I decided the normality of gait to be out because it depends on the previous three domains. If they [are] okay, it will be good.” (HCP)
	Subjective	-	2	“It is difficult to know what normal is, particularly in the context of inflammatory arthritis where previous joint damage may render gait ‘abnormal’ which places a ceiling of research outcomes.” (HCP)
	Not as important as other domains	1	-	“Of these three crucial domains, the normality of gait seems to me to be the least important. Life is hugely improved with reduction in pain and improved range of motion; walking ‘normally’ is a luxury here.” (Patient)
Life impact domains				
Treatment satisfaction	Subjective	2	1	“Satisfaction after treatment is subjective. Effects may appear in the longer term.” (HCP)

Domain	Sub-category	MUs (patients)	MUs (HCP / researchers)	Verbatim free-text quote
	Lacks granularity	-	3	“Satisfaction is influenced by many factors beyond the treatment itself, including expectations, communication with healthcare providers, and personal circumstances. This makes it harder to interpret consistently across studies. Additionally, satisfaction doesn’t pinpoint what worked or didn’t work in a treatment. It’s a broad summary that lacks the granularity needed for clinical decision-making or comparing interventions.” (HCP)
	Redundancy/overlap with other domains	-	1	“Treatment satisfaction would overlap with scores from the other domains, i.e. if footwear were comfortable and pain reduced, then the participant would be more likely to be satisfied.” (HCP)
	Difficult to measure	-	1	“Treatment satisfaction is notoriously difficult to assess as participants rarely want to indicate dissatisfaction for fear services will be removed.” (HCP)
Physical function	Confounding factors	1	-	“Other factors may affect the physical activity and the patient wrongfully relates it to his original condition.” (Patient)

Domain	Sub-category	MUs (patients)	MUs (HCP / researchers)	Verbatim free-text quote
Emotional wellbeing	Confounding factors	4	3	“I think emotional wellbeing is important but I don’t think it needs to be a core domain. If we can exercise more and sleep better then emotional wellbeing is likely to improve anyway.” (Patient)
	Not as important as other domains		2	“While I think emotional wellbeing is important, now I am not sure that it is mandatory.” (HCP)
Sleep	Confounding factors/subjective	-	1	“Sleep quality is influenced by many factors beyond foot and ankle health, such as stress, environment, and unrelated medical conditions, which complicate the isolation of treatment effects in this specific context.” (HCP)
	Difficult to measure	-	1	‘Sleep cannot be monitored over a long period as it will improve within a short time and stay the same.” (HCP)
	Not important in all RMDs	-	1	“I think the sleep domain may be important in cases like neuropathic pain conditions.” (HCP)
	Redundancy with other domains	1	3	“Improvements in pain intensity and emotional wellbeing often lead to better sleep, making the inclusion of sleep as a core domain potentially redundant.” (HCP)

Domain	Sub-category	MUs (patients)	MUs (HCP / researchers)	Verbatim free-text quote
Footwear comfort	Too difficult to define/measure	2	6	“The domains I have removed can be very objective and difficult to measure.” (Patient)
	Should only be mandatory in certain circumstances	-	3	“Footwear comfort should be mandatory but only in certain circumstances – footwear being investigated (or as part of a complex intervention).” (HCP)
	Redundancy with other domains	1	1	“Discomfort in footwear often stems from underlying issues like pain, swelling, or deformity, which are already captured in domains such as pain intensity, pain during weight-bearing, and joint range of motion.” (HCP)
Resource use domains				
Personal expenses	Confounding factors	6	10	“I think personal expenses are very important BUT they will also vary tremendously across patients and types of insurance and factors that are not necessarily related to treatment success (or failure).” (HCP)
	Not as important as other domains	2	-	“Whilst this might be important, it slips down my priorities behind the physical improvements that come with treatment.” (Patient)

Domain	Sub-category	MUs (patients)	MUs (HCP / researchers)	Verbatim free-text quote
	Not important in all RMDs	1	2	“I selected ‘OUT’ as I don’t think it’s quite as important for all conditions as the others, particularly for more mild conditions.” (Patient)

MU, meaning unit; RMD, rheumatic and musculoskeletal disease

6.3.14.2 Delphi participation feedback

A total of 22 (29%) HCP/researcher participants and 15 (31%) patients included free-text responses at the end of the Round 4 survey.

Theme 1: Survey design

Thirteen HCP/researcher participants and four patient participants commented on the design, structure and ease of survey completion. Most comments were positive; participants felt the surveys were well organised, clear, easy to follow and an appropriate length. One patient participant highlighted issues using the dragging function on a mobile phone to rank domains in Round 4.

Theme 2: Engagement

Four HCP/researcher participants commented that the process was engaging. One participant indicated that the reminder emails were important, whilst another commented that they liked the personalised emails. One patient participant also commented on engagement, noting that the time period in between surveys was appropriate: “not so close together that project became onerous, not too far apart that project purpose/scope was lost in memory.”

Theme 3: Information

Two HCP/researcher participants and two patient participants commented on the information provided about the Delphi study. Participants appreciated the guidance relating to time commitment, the purpose and process. One patient participant commented on the complex medical terms and explanations in the survey, and suggested that future studies could make sure that all text is in plain English.

Theme 4: Overall perceptions

Overall, participants were positive about their experiences of taking part in the Delphi study. Six participants (two patients) highlighted that the study was valuable and important. Two HCP/researchers expressed that they had enjoyed the study, whilst two patients commented that the study was interesting. Two other patient participants suggested that interviews or focus groups would be a better way of understanding what domains are important to patients, whilst another felt that it would be useful to have

more information in the hypothetical clinical trial examples presented in Round 4 (e.g. age, gender, and RMD).

6.3.15 Final PAG meeting

All Delphi results were discussed with the PAG in an online consensus meeting on 8th September 2025. The PAG agreed that all domains that achieved ≥ 70% consensus from both groups in Round 4 (pain intensity, pain when weightbearing, physical function (activities and participation), and treatment satisfaction) should be included as core domains in the final core domain set. Adverse events (including death) was also included as a core domain. As this is a mandatory inclusion in any OMERACT COS, it did not go through the consensus process.

The inclusion of joint ROM as a core domain was debated. The PAG considered the HCP/researcher and overall rating for this domain, as well as findings from the Delphi content analysis and previous qualitative studies, and discussed whether it should be included as a core domain or a circumstance-specific core domain (for studies where the aim of treatment was to influence joint ROM). The outcome of the discussion was to include it as a core domain, but to amend the name of the domain to “joint movement”, to match the domain definition that had been provided to participants.

The PAG then discussed circumstance-specific core domains. Whilst personal expenses had the lowest ratings in Round 4 (when participants were asked to focus on what should be included as core domains), the PAG considered expenses important to measure in specific circumstances, such as phase 3 clinical trials. Resource/societal use is also a strongly recommended core area in the OMERACT framework. The PAG agreed that “healthcare expenses” should be a circumstance-dependent core domain, taking into consideration the importance of personal expenses throughout the qualitative work and previous Delphi rounds as well as the importance of direct costs/cost-effectiveness to HCPs/researchers (based on Round 1-3 Delphi results). In addition, the PAG agreed that structural pathology (encompassing the target domains of inflammation and joint damage/deformity) should be a circumstance-dependent core domain for clinical trials investigating the effectiveness of treatments targeting these. This decision was based on the consistently high percentage of HCP/researcher participants rating these domains as critically important in Rounds 1-3. The PAG agreed that the remaining domains from Round 4 (emotional wellbeing, sleep, normality of gait, and footwear comfort) were important domains for future consideration.

The final core domain set is presented in Figure 6.3.

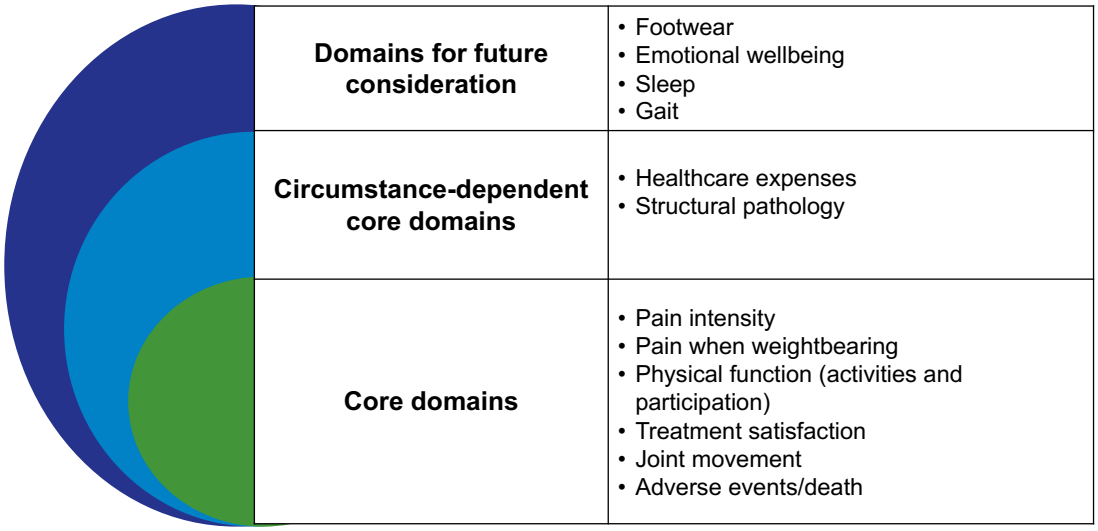


Figure 6.3: COMFORT core domain set

6.3.16 Protocol deviations

The original study protocol (87) stated that the first round of the Delphi study would be open for six weeks. This was reviewed prior to launching the survey and reduced to two weeks in an attempt to facilitate participant engagement. The protocol also stated that a final OMERACT-hosted consensus meeting would take place. However, OMERACT subsequently amended their standard process, replacing consensus meetings with a definitive fourth Delphi round, a PAG consensus meeting, and final voting through an online module. The COMFORT study followed the updated OMERACT process.

6.4 Discussion

This study reports findings from an online, four-round, modified Delphi consensus study. It achieved its aim to develop a core domain set for future clinical trials and observational studies involving patients with foot and ankle disorders in RMDs. The Delphi was developed using established methods (63, 109), with international collaboration across a broad spectrum of patients, HCPs, and researchers with experience in this area. This study incorporated findings from the scoping review of domains in existing studies (73) (Chapter 1, section 1.5.1), the qualitative synthesis of existing studies (Chapter 3), the secondary analysis of focus group transcripts (Chapter 4), and the qualitative interview study (Chapter 5).

6.4.1 Main findings

Consensus was reached on the inclusion of five core domains: pain intensity, pain when weightbearing, physical function (activities and participation), joint movement, and treatment satisfaction. As a mandatory domain within any OMERACT COS, adverse events (including death) was included as a sixth core domain. Over 70% of patients and HCPs/researchers agreed on the inclusion of these domains in the core domain set in Round 4 of the Delphi. These domains should be measured and reported, as a minimum, in all future clinical studies in this area of health. In the following section, findings from this Delphi study will be discussed in relation to previous studies within the COMFORT project, and other OMERACT Delphi studies focusing on developing a COS for specific joints: hip and knee OA, (320), hand OA (340), and shoulder disorders (149). Comparisons between the COMFORT core domain set and other COS, including joint-specific OMERACT COS, disease-specific OMERACT COS, and COS in the broader literature, will be discussed in further depth in Chapter 7.

6.4.1.1 Core domains

Consensus for the inclusion of pain intensity and pain when weightbearing was achieved among both patients and HCPs/researchers in Round 2 of the Delphi, whilst consensus for physical function (activities and participation) was achieved earlier, in Round 1. These three core domains achieved the highest ratings in Round 4, in both the in/out and ranking exercises. In a previous scoping review of the domains measured in existing clinical trials and observational studies of foot and ankle disorders in RMDs (73), pain (118/150 studies) and function (102 studies) were the most commonly identified domains. Correspondingly, pain and function were the most prominent themes in both the synthesis of existing qualitative studies and the primary qualitative interview study with patients.

Previous OMERACT Delphi studies conducted in the COS development process for specific joints demonstrated similar findings. In the final round of the OMERACT hip and knee OA Delphi study, 98% of patients and 97% of other contributors rated pain as a critically important domain, whilst 100% of patients and 95% of other contributors rated function as a critically important domain (320). The OMERACT Delphi studies for shoulder disorders and hand OA reported similarly high consensus levels (≥ 90%) for these domains, among patients, HCPs, and researchers (354).

Joint movement (30/150 studies, measured as joint ROM) has been less frequently measured in previous foot and ankle RMD studies, but was identified as an important

domain to patients in the qualitative research underpinning the Delphi study. Joint movement was included as a core domain based on its overall score (patients and HCPs/researchers combined), although there was a difference in Round 4 between patients (86%) and HCPs/researchers (65%) rating this domain into the core domain set.

Comparatively, whilst joint ROM was included as a candidate domain in both the hip and knee OA and shoulder Delphi studies, it did not achieve consensus for inclusion in either final core domain set (149, 320). In the Delphi for hip and knee OA, 81% of patients rated joint ROM as critically important, compared to 29% of other contributors (320). Similarly, a higher percentage of patients, compared to other contributors, considered joint ROM to be a critical domain in the Delphi for shoulder disorders. The marked differences between patient and HCP/researcher priorities could be explained by findings from the COMFORT Delphi content analysis; HCPs/researchers indicated that joint movement is difficult to measure, and that its relevance is intervention- and RMD-dependent. In congruence, the hand OA core domain set, joint mobility was included as a circumstance-specific core domain for trials of structure modification and observational studies (340).

It is possible that improving joint movement will not be the explicit intention of every future foot and ankle RMD study. However, the Delphi and the qualitative research underpinning it included patients with multiple RMDs. Joint movement has been deemed important to patients and should therefore be measured, even if HCP/researchers did not reach consensus on its relevance. Challenges relating to implementing the COMFORT core domain set will be explored in more detail in Chapter 7.

Treatment satisfaction (36/150 studies) has also been measured less frequently than pain and function in existing foot and ankle RMD studies (73). The content analysis identified potential difficulties in measuring this domain. Satisfaction is a complex domain with no standardised definition, and can reflect a patient's values, expectations, and emotions in addition to the effectiveness of a treatment (355, 356). Satisfaction with treatment services was included in the first round of the Delphi study for shoulder disorders, where 60% of participants rated it as important enough to be included in a core domain set (354). The breakdown of contributor scores was not reported, and the domain was not included in subsequent rounds. Satisfaction was not included in the hip and knee OA or hand OA Delphi studies (320, 357). These differences may reflect disparities in patients' experiences of treatments and care. As discussed in Chapter 1, evidence for foot and ankle treatments lags behind other sites, including hip and knee OA where many robust clinical trials have been conducted and definitive treatments are

available (338, 339). Therefore, dissatisfaction among patients is likely to be more pertinent in the area of foot and ankle disorders in RMDs.

6.4.1.2 Circumstance-specific core domains

Two circumstance-specific core domains, to be measured and reported in all relevant trials, were also included in the COMFORT core domain set: structural pathology and healthcare expenses. Structural pathology is recommended for studies where the target of the intervention is to improve joint structure or deformity, or local disease activity/inflammation (e.g. in IA and OA). This domain was captured in the previous scoping review as disease progression/deformity on imaging (51/150 studies), clinician-assessed foot/ankle disease activity (6 studies), disease activity on imaging (9 studies), and disease activity on laboratory markers (6 studies).

The OMERACT Delphi for hip and knee OA included a similar circumstance-specific core domain: joint structure (320). This was defined as imaging (such as radiograph, MRI, ultrasound) reflecting changes in joint structure, and was recommended for clinical trials investigating structure-modifying interventions. Joint structure was rated as a critically important domain by 71% of patients and 40% of other contributors in the third round of the Delphi. In a subsequent consensus meeting, participants expressed concerns that this domain would not be relevant for non-structure-modifiable interventions, hence its inclusion as a circumstance-specific core domain. The OMERACT core domain set for hand OA also included structural progression (to be measured radiographically) as a circumstance-specific core domain in clinical trials of structure modification, as well as including disease activity as a core domain. Surgical process outcomes was a candidate domain in the OMERACT Delphi for shoulder disorders, but this did not progress further than Round 1 (354).

In the current Delphi, the domains joint damage/deformity and inflammation were both rated as critically important to include in a core domain set by ≥ 70% HCPs/researchers in all three rounds of the Delphi, but did not achieve ≥ 70% consensus among patients, thus did not progress to Round 4. This contrasted with findings from the Delphi for hip and knee OA, where patient ratings were higher than other contributor groups, and potentially reflects differences in patient participants. For example, the COMFORT Delphi included patient participants with foot and ankle disorders across many different RMDs, some of which do not cause structural damage, as opposed to all patient participants having OA, where structural damage is common.

The decision to include the broad domain structural pathology (encompassing the target domains of joint damage/deformity and inflammation) as a circumstance-specific

core domain was made by the PAG, influenced by Delphi scores that suggested it is an important domain to HCPs and researchers, and by the previous qualitative research (254, 317). Both the qualitative synthesis (Chapter 3) and interview study (Chapter 5) identified joint deformity as an important domain to many patients with IA and OA, but it was less important to patients with other RMDs. The inclusion of structural pathology as a core domain for all clinical studies is therefore not appropriate, given the intended scope of the COS, which applies to multiple RMDs. The decision to measure this circumstance-specific core domain will ultimately depend on what a study is trying to achieve; researchers must consider if the domain is important to the type of patients the study is including (e.g. do they have joint damage/deformity or inflammation?), and if the aim of the intervention within their study is to modify any of these manifestations (e.g. a surgical intervention addressing joint damage/deformity, or a pharmacological intervention addressing foot and ankle inflammation).

Healthcare expenses, encompassing the target domains cost-effectiveness and personal expenses, was also included as a circumstance-specific core domain in the COMFORT core domain set. Healthcare expenses is an under-researched area in foot and ankle studies; only 11/150 studies in the scoping review measured domains relating to cost (358-368). Similarly, a systematic review of the cost-effectiveness of interventions for foot and ankle MSK conditions highlighted a limited evidence base in this area (73, 369). In the development of the OMERACT core domain set for hip and knee OA, a healthcare expenses domain was included in the three-round Delphi study (direct costs). This broad domain included healthcare use (costs, painkiller use, hospital admission and consultations) and time to surgery (e.g. joint replacement). In Round 3 of the Delphi, 79% of patients with hip and knee OA rated healthcare use as critical to measure, compared to 66% of other contributors, and 83% of patients with OA rated time to surgery as critical to measure, compared to 42% of other contributors. In the final core domain set, ‘costs’ was included as an important but optional domain. No domains relating to healthcare expenses were included in the hand OA or shoulder Delphi studies (320, 340, 357).

In the current Delphi study, cost-effectiveness was rated highly by HCPs/researchers in the first three rounds of the Delphi, but did not achieve ≥ 70% consensus among patients, thus did not progress to Round 4. In contrast, the personal expenses domain reached consensus from both groups at the end of Round 3, but only 34% of overall participants in Round 4 rated it into the core domain set. Nevertheless, resource/societal use is a recommended core area within the OMERACT framework (63), and inclusion of cost-effectiveness and personal expenses represents HCP/researcher-focused and patient-focused domains. The inclusion of healthcare

expenses as a circumstance-specific domain reflects that it may not be plausible to measure in all future studies.

6.4.1.3 Domains for future consideration

Emotional wellbeing, sleep, gait, and footwear were identified as domains for future consideration. These can be considered 'important but optional' domains that may require further research (63). These domains reached consensus between patients and HCPs at the end of Round 3 of the Delphi, but were not rated into the core domain set in Round 4. Sleep and emotional wellbeing/psychosocial impact were both deemed 'important but optional' domains in the core domain sets for OMERACT hip and knee OA and shoulder disorders. In both cases, ≥ 70% patients rated sleep and emotional wellbeing as critically important, but this consensus threshold was not met by other contributors. Neither domain was included in the OMERACT hand OA Delphi, despite a preceding qualitative study identifying the psychological consequences of hand OA (370). Neither footwear nor gait domains have been included in previous OMERACT core domain sets, reflecting their specificity to foot and ankle disorders in RMDs.

6.4.1.4 Key differences between contributors

In Round 3 of the Delphi study, there were no domains that reached consensus amongst patient but not HCP/researcher participants, whilst ten domains reached consensus amongst HCPs/researchers but not patients. This reflects findings in previous rounds, where patients' ratings tended to be lower than HCP/researcher ratings. Whilst this was not explored further in the current study, patients may provide more conservative ratings in Delphi studies because they feel less confident in their decision-making (371). Most of the domains rated as critically important by HCPs/researchers and not by patients were more objective, pathophysiological domains, such as joint instability and sufficiency of blood supply, in addition to the structural pathology domains discussed above.

In addition to differences in ratings for joint movement, as discussed above, two other domains in Round 4 (normality of gait, and sleep) reached consensus among patients but not HCP/researcher participants. These findings may reflect patient participants' personal experiences of their own condition, compared to HCPs/researchers' consideration of all RMDs. Additionally, in the content analysis, HCPs/researchers highlighted the difficulties in measuring gait.

6.4.2 Strengths and limitations

The main strengths of this study include the application of well established, standardised COS methodology, the active involvement of a multidisciplinary and multi-contributor PAG and PPI contributors throughout the study design and consensus building process, international recruitment of a diverse range of patients and other contributors, and strong participant retention rates between rounds.

OMERACT is well-established in the RMD community and the organisation's standardised methodology and expertise is based on almost three decades of COS development experience. Developing a core domain set with OMERACT endorsement has the potential to increase its uptake (242). However, aspects of OMERACT methodology employed in the Delphi study may have impacted on the consensus process. For example, it is possible that holding domains that achieved consensus in Round 2 for final rating in Round 4 (i.e. not presenting them in Round 3) increased the consensus level of the remaining domains presented in Round 3. Further research could assess the impact of holding domains that have achieved consensus versus presenting all domain scores from the previous round. Additionally, the domains presented in the Delphi study were ordered within the OMERACT core areas of pathophysiological manifestations, life impact, and societal/resource use, as per OMERACT guidance, with the exception of the ranking exercise in Round 4. This conflicts with evidence suggesting that the order of the domains presented affects consensus (209), and the subsequent recommendation that the order of objective (HCP/researcher-focused) and subjective (patient-focused) domains should be randomised. Carrying out four Delphi rounds is a new stipulation within OMERACT methodology, and this study was the first to go through this process. As such, analysis of the ranking task was not prespecified. Guidance to do an in/out exercise *and* a domain ranking exercise exposed unprecedented issues, with discrepancy between the two exercises for some domains. For example, whilst pain intensity, pain when weightbearing, physical function (activities and participation) and joint ROM were rated consistently across both exercises, treatment satisfaction ranked much lower than expected based on the in/out exercise. This could have been overcome by asking participants to rank only the domains they voted into the core domain set, as opposed to all domains, although it is debatable whether both exercises are required in Round 4. The survey software utilised in this study may also have influenced the ranking exercise, as one participant highlighted an issue when completing this exercise on a mobile phone. A limitation of this study is that the fourth round of the Delphi was not piloted on different devices.

The Delphi study benefited from the PAG's extensive experience of selecting, measuring and reporting domains in previous clinical trials and observational studies. However, in some cases there was a conflict between findings from the Delphi (and the preceding qualitative research), in terms of the importance of a domain to patients, and how feasible some members of the PAG felt the domain would be to measure. For example, joint ROM was rated highly by patients, and achieved ≥ 70% consensus overall in Round 4, but HCP/researcher members of the PAG expressed hesitations about including it as a core domain due to the difficulties in measuring it objectively. This may underline why it did not achieve ≥ 70% consensus among HCPs/researchers. These apprehensions echoed those from HCP/researcher participants in the free-text comment content analysis. COS development guidance emphasises that determining *what* to measure should be based on what is important to measure, without being influenced by *how* to measure. This conflict partially influenced the decision to reword joint ROM to 'joint movement', as joint ROM is associated with objective measurement (such as goniometers), whereas movement could be a patient-reported domain.

Input from the PPI members within the PAG, and the additional PPI perspectives that were sought during the study, was fundamental to the development of a core domain set that reflects what is meaningful to people who live with foot and ankle disorders in RMDs. PPI contributors maintained the patient voice throughout the Delphi, influencing the readability and accessibility of the surveys, and inputting into domain decisions at every stage. This ensured that domains of importance to patients were protected.

Another strength of this Delphi is that the first round was piloted with patients and HCPs outside of OMERACT, to ensure the survey was understandable. However, the majority of those who piloted the study were based in the UK; the Delphi could have been piloted among non-research active patients and HCPs, including those whose first language was not English, more widely outside of the UK to improve inclusivity. Whilst the presentation of individual and group scores from each round followed accessibility guidelines and was piloted with the PAG, these aspects of the study may also have benefited from wider piloting. This is particularly pertinent given that many individuals lack health numeracy skills (221). Additionally, whilst the use of a readability calculator for the survey background and instructions was a strength of this study, the readability of domain definitions was not formally assessed and may not have been fully understand by all participants. However, only one free-text comment revealed a potential issue with the domain definitions, with other participants highlighting that these were clear. Understandability may have benefited further from conducting think-aloud interviews during piloting (372).

International representation was also a major strength of this Delphi study; participants from 22 countries were involved, including 70 participants whose first language was not English. A total of 27 participants from four lower-middle income countries (Egypt, Morocco, India, and Bangladesh) completed all four rounds of the Delphi. A systematic review assessing the involvement of LMICs in international Delphi surveys for COS development identified that only 25% of studies included Delphi participants from LMICs (104). Due to resource limitations, the current Delphi study was only presented in the English language and participants had to be able to access the internet and email in order to participate. Consequently, findings may not represent the priorities of non-English speaking patients and HCPs/researchers, those with limited digital literacy, or those in digital poverty. This reflects most COS development Delphi studies, where translation from English into other languages is rarely conducted (104). OMERACT Delphi studies have typically been restricted to the English language, although an ongoing Delphi study prioritising domains for SLE clinical trials is translating into ten different languages using AI (373). The advantages and disadvantages of translating Delphi studies into different languages will be discussed further in Chapter 7. International representation was strongest from HCP/researcher participants, with six continents represented, following distribution of the study through professional organisations internationally. Whilst patient representation from Egypt was strong, likely due to the involvement of a clinical collaborator during recruitment, most other patient participants lived in high-income, English-speaking countries. This is a recurrent limitation in COS development studies (103, 104), and methods to increase and maintain engagement from under-represented countries in Delphi studies remains an important area for ongoing COS methodology research. Notwithstanding, the Delphi study was designed to be as inclusive as possible for participants whose first language was not English, and the views of non-English speaking participants were represented in the qualitative study (Chapter 5), findings from which informed the design of the Delphi survey.

This study was strengthened by good representation from three key contributor groups (patients, HCPs, and researchers). Some of these participants have known dual roles as guideline developers, healthcare managers, journal editors, and positions on funding panels. However, there was no representation from other potentially relevant contributor groups, such as industry representatives and commissioners. Other COS developers have also had difficulties engaging with these contributor groups (374). Another potential limitation is the lack of representation from certain HCP/researcher professions. Relevant professions were explored in Chapter (section 2 2.9.4); these groups have experience in managing or conducting research with patients with foot and ankle disorders in RMDs. The majority of HCP/researcher participants were

rheumatologists, podiatrists, or physiotherapists by background, with a limited number of orthotists and GPs. No nurses, orthopaedic surgeons, or podiatric surgeons took part in the study, despite prior engagement and wide promotion of the study among relevant professional groups. Thus, the core domain set may not represent the perspectives of all relevant professional backgrounds. Lack of surgical representation is particularly relevant, given the frequency of studies in this area that focus on comparing surgical interventions. In the previous scoping review of domains, 75/150 studies (50%) involved at least one surgical intervention. Participation of surgeons in the OMERACT core domain set for hip and knee OA was also much lower than participation of other HCP/researcher groups (320). A key objective for the next phase of COS development, where the focus is on OMI selection, will be to increase engagement among surgeons, especially as the top priority for foot and ankle surgeons in a recent James Lind Alliance priority setting partnership was to determine the best outcome measures for use after foot and ankle surgery (375). Despite lack of input from surgeons in the current study, the core domain set likely reflects the perspectives of other HCPs who manage patients prior to and following foot and ankle surgery, and does represent the perspectives of patients who are awaiting or have had foot and ankle surgery. Whilst this information was not formally recorded in the Delphi study, some of patients' free-text responses related to their surgical experiences. Surgical history was also explored in the preceding qualitative interview study and synthesis, the findings from which informed the domains included in the Delphi. Additionally, surgical trials were discussed in the final PAG consensus meeting and the broad domain of structural pathology (including the target domain of joint damage/deformity) was added as a circumstance-dependent domain in recognition of its importance in this context.

Another strength of this study was the high participant retention rate between rounds. A total of 206 participants (83 patients) took part in Round 1, with a dropout rate of 27.7% (31.3% patient dropout rate) between Rounds 1 and 2, 12.8% (12.3% patient dropout rate) between Rounds 2 and 3, and 3.1% (2% patient dropout rate) between Rounds 3 and 4. In comparison, the three-round OMERACT hip and knee OA Delphi (320) had a dropout rate of 48% (n = 166 participants) between Rounds 1 and 2, and 33% (n = 58 participants) between Rounds 2 and 3, whilst the two-round OMERACT Delphi for shared decision-making in rheumatology (376) had a dropout rate of 31.7% (n = 54 participants) between Rounds 1 and 2. The OMERACT CPPD Delphi (350), in which the authors reported specific software issues and delays, had a dropout rate of 41.4% (n = 50 participants) between Rounds 1 and 2, and 39.5% (n = 28 participants) between Rounds 2 and 3. The higher retention rate in the current study may have been due to a personalised email approach rather than the use of automatically generated emails (350), and the influence of recruitment and reminder emails sent from PAG

members. However, impact on recruitment and retention (e.g. the increase in responses after reminder emails were sent) was not formally measured, and is an area for future research. Additionally, the use of patient-centred approach in reminder emails (Figure 6.1) may have improved retention rate. This could also be explored further in future Delphi studies. Information about a gift card incentive may also have influenced retention. Whilst an incentive could potentially affect the reliability of responses (377), the information was provided at the end of Round 2 when participants had already invested time into the study without knowing about an incentive. Further research is needed to look at the impact of incentives at different points in the consensus development process. All participants were offered an acknowledgement in the publication arising from this study as an incentive; 75 (97%) of Round 4 HCP/researcher participants confirmed they would like their name to be included, indicating the potential of this method for recruitment and retention. Additional recruitment and retention methods include selecting motivated participants through an initial invitation (184) and setting up a dedicated website to share information and updates about the Delphi with participants (193, 378). The impact of these methods on recruitment and retention were not feasible in the current study due to time and resource limitations. A previous study reported showing short demonstration videos to supplement written instructions for a Delphi, to promote usability of the survey and reduce technical problems by giving participants a visual step-by-step guide on how to complete each round (193). However, this was deemed unnecessary in the current study due to the simplicity of the survey platform, and supported by the lack of email requests to the research team for assistance with the survey.

6.4.3 Implications

6.4.3.1 Implications for research

This study has achieved consensus on *what* domains should be measured. Future work must focus on *how* these domains should be measured, as the use of heterogenous OMIs to measure the same domain can also contribute to difficulties comparing and combining findings from different studies (65). This phase of the COS development process was outlined in Chapter 2 (section 2 2.7). Preliminary OMI work in relation to the development of this core domain set will be discussed in Chapter 7.

6.4.3.2 Implications for the COS methodology base

Multiple methods to increase recruitment and retention were employed in this Delphi study (Table 6.7). These methods were not formally assessed in this study, but positive feedback regarding participants' experiences of completing the Delphi study provided valuable insight to inform wider COS methodology. However, not all participants left free-text responses in this section, and responses may have reflected a bias among those who had a positive experience. Future research could explore experiences of participation further in more depth, such as in interviews or focus groups, and could include an exploration of the experiences of participants who dropped out of the study.

6.4.3.3 Implications for clinical practice

In the long term, findings from this study have the potential to improve the quality of evidence for treatments for patients with foot and ankle disorders in RMDs. At present, the substantial heterogeneity of domains measured and reported in research hinders the ability to compare and combine findings from different studies, reducing the quality of evidence for treatments. Standardisation of domains in foot and ankle research, through the use of the core domain set developed in this study, will facilitate meta-analyses. This in turn will improve clinical decision-making for HCPs. Input from patients throughout the development of this core domain set should also increase the relevance of future research findings, making them more transferrable to practice. This study also has more immediate implications for HCPs and patients. Following OMERACT methodology, the intended scope of the core domain set applies to research, and participants were not asked to consider what to measure in clinical practice. However, the domains that matter most to patients are likely important in both settings. The core domain set can guide clinical conversations about foot and ankle disorders, focusing appointments on what patients want treatments to achieve. The potential use of the COMFORT COS in routine clinical practice will be explored in Chapter 7.

6.4.3.4 Implementation

Further work is also crucial to facilitate implementation of these findings into future studies. COS uptake, and some of the issues surrounding it, was explored in Chapter 2 (section 2.8). Implementation (including a COMFORT dissemination strategy) will be discussed in Chapter 7.

6.5 Conclusion

The Delphi study reported in this chapter achieved its aim of developing an internationally agreed core domain set to be measured and reported, as a minimum, in clinical trials involving treatments for patients with foot and ankle disorders in RMDs, in line with the COMFORT project's second objective. The core domain set represents the priorities of a diverse range of patients, HCPs, and researchers, from 22 countries, with a strong retention rate across four rounds of surveys. Five domains reached consensus for inclusion as core domains (mandatory for all trials): pain intensity, pain when weightbearing, function/participation, joint movement, treatment satisfaction. Adverse events (including death) is also a mandatory core domain. Two circumstance-dependent core domains were also included: structural pathology (encompassing joint damage/deformity and inflammation) and healthcare expenses (encompassing cost-effectiveness and personal expenses). Four domains for future consideration were also identified: emotional wellbeing, sleep, gait, footwear. The study also provided insight for COS methodologists, by collecting feedback from participants about their experience of taking part in the Delphi. Uptake of this core domain set in future trials has the potential to improve the quality of the evidence for foot and ankle treatments, by facilitating effective data synthesis and improving translation of findings into clinical practice, supporting evidence based decision-making. The next chapter will summarise and discuss the overall COMFORT project.

Chapter 7 Discussion and conclusion

7.1 Introduction

This chapter discusses the overall COMFORT project, including its original contributions to the literature. Firstly, a summary of each phase of the project is provided, and an interpretation of key findings is presented. Reflections on the methodological process are then discussed, followed by the strengths and limitations, and implications for clinical practice and future research. Finally, the overall conclusion of the COMFORT project is presented.

7.2 Project summary

This work sought to overcome two major issues in the area of foot and ankle RMD research. The first is that domains are measured inconsistently, compromising data synthesis between otherwise comparable studies. The second is that the domains measured are not necessarily those that are relevant to patients, limiting the transferability of study findings to clinical practice. The consequence of these issues is a lack of high-quality evidence for foot and ankle interventions to inform clinical guidelines and decision-making.

An important part of the solution to these problems is the development of a COS: an agreed, standardised set of domains that are clinically relevant, reflecting what is important to patients and other key contributors, measured and reported (as a minimum) in all studies in this area of health (65). The overarching purpose of the COMFORT programme of work, therefore, was to improve the quality of evidence relating to the effectiveness of interventions for foot and ankle disorders in RMDs, by improving the consistency and relevance of domains measured in foot and ankle RMD research through development of a core domain set. The core domain set will inform the design of future research studies by specifying *what* domains should be measured and reported as a minimum. Research studies measuring the same domains that are meaningful to all key contributors will facilitate data synthesis and improve the relevance of future research findings.

7.3 Review of exploratory hypothesis

The first chapter of this thesis outlined an exploratory hypothesis underpinning this doctoral programme of work. The exploratory hypothesis stated that by using a mix of

methods and with input from multiple key contributor groups (including patients) internationally, it would be possible to develop a core set of domains to be measured and reported in future research involving foot and ankle disorders in a range of RMDs. This hypothesis was explored in each of the empirical chapters of this thesis and the completed programme indicates that this approach results in a comprehensive process and robust outcomes.

7.4 Overview of thesis

COMFORT adopted a rigorous evidence-based approach, using a mix of methods. Each phase of the work built on the findings of the preceding phase. The empirical chapters within this thesis will now be summarised.

Systematic review and thematic synthesis of existing qualitative studies (Chapter 3)

This study aimed to identify and synthesise existing qualitative literature describing foot and ankle symptoms, and the impact of these symptoms, among patients living with RMDs. This was achieved through a systematic review and thematic synthesis and focused on exploring patients' perspectives and experiences of foot and ankle disorders, to subsequently determine which domains are important to patients and should be considered for inclusion in a COS. The 34 studies within the review included 503 patients with nine different RMDs. Descriptive findings were further investigated to produce analytical themes corresponding to broad and target domains of importance. Important domains to patients included pain, function (physical, social, occupational), appearance of the feet, financial burden, footwear impact, and treatment satisfaction. Findings from this systematic review highlighted gaps in the qualitative evidence base with regards to under-represented RMDs, predominantly SSc, as well as ethnicities and geographic locations.

Secondary analysis of focus groups (Chapter 4)

This study sought to explore the perspectives of patients with SSc-related RP in the feet, addressing the paucity of qualitative research for this RMD, as identified in Chapter 3. A secondary analysis of focus group transcripts was conducted using a directed approach to content analysis. Focus groups with 40 patients had been carried out previously by a different group of COS developers intending to standardise domains for studies involving patients with SSc-related RP, but omitting foot-related data. The secondary analysis identified six key symptoms, which were presented as

themes corresponding to domains: temperature changes, pain, cramping and stiffness, numbness, and colour changes. These symptoms impacted on activities and footwear choices. This study resulted in domains that had not been identified in previous quantitative or qualitative literature, and informed further qualitative exploration in a wider range of RMDs.

Qualitative interview study (Chapter 5)

The aim of this study was to identify any additional domains of importance to patients living with foot and ankle disorders in RMDs, through a direct exploration of their experiences. This was achieved through a qualitative descriptive study involving online, telephone, and face-to-face interviews with 56 patients from eight countries, including one LMIC. Patients with a range of RMDs, including some that had not been explored in previous qualitative literature (adult JIA, AS, enteropathic arthritis, CPPD, mixed CTDs), ethnicities and first languages participated. The interviews sought to identify domains of importance to patients and look at how domains compared across different characteristics. The interviews were guided by a schedule, and transcripts were analysed using the framework method. The interview schedule and coding framework were co-produced with PPI contributors, based on findings from the thematic synthesis and secondary analysis of focus groups, in addition to a previous scoping review of domains. Three domains, foot- and ankle-specific fatigue, cognitive fatigue and impact on other joints, had not been identified previously.

Online Delphi consensus study (Chapter 6)

This study aimed to establish international, expert consensus on a core set of domains to measure in future foot and ankle RMD research. This was achieved through an online Delphi study with patients, HCPs, and researchers, involving four rounds of surveys, following established COS methodology. Findings from the previous scoping review and qualitative studies were used to develop an initial list of domains and domain definitions for the first round, with additional PPI input. Novel methods to aid recruitment and retention were adopted. A total of 206 participants (83 patients) from 22 countries took part in Round 1 of the Delphi consensus study. A total of 126 participants from 17 countries were represented in all four rounds, including 49 patients with a range of RMDs in addition to HCPs/researchers from different professional backgrounds. The final core domain set consisted of six core domains: pain intensity, pain when weightbearing, physical function (activities and participation), joint movement, treatment satisfaction and adverse events (including death). Two

circumstance-dependent core domains were also identified: structural pathology (inflammation and joint damage/deformity) and healthcare expenses (cost-effectiveness and personal expenses). Finally, four domains were deemed important for future consideration: emotional wellbeing, sleep, gait, footwear.

7.5 Interpretation of findings

7.5.1 Findings in relation to the existing literature

7.5.1.1 Joint-specific COS

Findings from the COMFORT Delphi study were discussed in relation to previous joint-specific OMERACT Delphi studies in Chapter 6 (section 6.4.1). Since the induction of this programme of work, a core domain set for ankle OA has been published independently of OMERACT (379). Ankle OA is one of the RMDs within COMFORT's scope. Reassuringly, the core domain set for ankle OA is similar to COMFORT; the core domains are pain severity, health-related quality of life, disability, and ankle ROM. A comparison of the core domain definitions between both core domain sets is presented in Table 7.1. Although disability and health-related quality of life do not feature in the COMFORT core domain set, their definitions overlap with physical function (activities and participation). As yet, no OMI-related work for this core domain set has been published.

An early point of discussion prior to seeking funding for the COMFORT project was whether each RMD where the foot and/or ankle is affected should have a separate COS. As most COS take many years to produce and require a significant amount of funding to execute rigorously, this was neither practical nor feasible. Furthermore, the qualitative work in this thesis has demonstrated that the domains that are important to patients are similar regardless of the specific RMD. This suggests that focusing on individual foot and ankle disorders would lead to a range of very similar COS, especially in a niche research area where many of the same HCPs/researchers would be taking part regardless of the specific condition targeted. Development of one COS for multiple conditions facilitates quicker application in future research, and ensures harmonisation of outcomes across studies investigating the same treatments in different RMDs. The existence of two core domain sets covering the same RMD, ankle OA, will likely be confusing for researchers, and future collaborative work is crucial to streamline the two sets at the OMI selection stage.

Table 7.1: COMFORT and ankle OA core domain set comparison

COMFORT core domains	**Domain definition**	**Ankle OA core domains**	**Doman definition**
Pain intensity	How severe pain in the foot/ankle is.	Pain severity	Participant-rated intensity of their pain.
Pain when weightbearing	Foot/ankle pain experienced whilst weightbearing (e.g. walking, standing).		
Physical function (activities and participation)	Impact of foot and ankle problems on walking (including ability to walk, walking distance), driving, personal care, household tasks, caring responsibilities, work (paid or voluntary), and being able to participate and enjoy life events, e.g. sports, exercise, social and recreational activities, religious activities, hobbies.	Disability	A multi-dimensional entity encompassing elements from the ICF domains of impairments and activity limitations, participation restrictions.
		Function	The perceived level of difficulty/ability a person has in performing certain activities/tasks that a person attributes to their ankle OA.
Joint movement	The current amount of movement in a foot/ankle joint.	Ankle ROM	Ankle joint ROM.
Treatment satisfaction	How good the patient feels the treatment received has been.	Health-related quality of life	The physical, social, and psychological dimensions of health that are being considered as separate health areas influenced by a person's experiences, beliefs, expectations, and perceptions.

As discussed in Chapter 6 (section 6.4.1), the core domain set developed in this study is also similar to the OMERACT core domain sets for hip and knee OA (320), hand OA (340) and shoulder disorders (149), although domain definitions have varied. COMFORT is the first OMERACT core domain set to include two core pain domains (pain intensity and pain when weightbearing). In the OMERACT core domain set for shoulder disorders, the broad domain of pain was included as a core domain, defined as *how much a person's shoulder hurts, reflecting the overall magnitude of the pain experience (rest, during and after activity, at night)* (149). Similarly, the OMERACT core domain set for hip and knee OA included the core domain of pain, which encompassed the following target domains: pain overall, pain with activity, pain at rest, pain during the night, and pain during the day. In the COMFORT core domain set, inclusion of two target domains relating to pain, rather than merging these into one broad domain, reflects qualitative findings indicating that these are distinct domains of importance to patients. It also attempts to overcome the challenges of measuring a broad domain that may need multiple OMIs.

The final definition of physical function in the COMFORT core domain set was *impact of foot and ankle problems on walking (including ability to walk, walking distance), driving, personal care, household tasks, caring responsibilities, work (paid or voluntary), and being able to participate and enjoy life events, e.g. sports, exercise, social and recreational, activities, religious activities, and hobbies.* This differed from definitions in other COS. For example, OMERACT has recently developed a standardised definition for physical function: *one's ability to carry out various activities that require physical capability, ranging from self-care (activities of daily living) to more vigorous activities that require increasing degrees of mobility, strength or endurance* (380). Similarly, the OMERACT core domain set for shoulder disorders separated physical function (activities) and participation (recreation and work). Physical function was defined as *a person's ability to carry out daily physical activities, ranging from self-care (e.g. bathing, combing hair) to more complex activities that require a combination of skills (e.g. driving a car)*, whilst participation was defined as *a person's ability to engage in a life situation, in any form of play, recreational, or leisure activity acts (e.g., sports of any kind or levels), and the ability to meet physical and/or psychological demands of work)* (149). Physical function was included in the final core domain set for shoulder disorders, whereas participation was an important but optional domain. The core domain set for hip and knee OA considered physical function from two different perspectives, one definition included the target domains leg function, personal activities of daily living, and sports, exercise and physical activity. The second definition encompassed the more objective measures of joint control (e.g. giving way), balance, muscle strength, joint ROM, exercise tolerance and endurance. Participation was

considered as a separate domain, defined as role function (ability to do work or vocational activities). Physical function was included as a core domain, whereas participation was deemed an important but optional domain (320). In the COMFORT qualitative study, reduced ROM, stiffness, muscle weakness, instability, and balance, and altered gait/walking were all categorised within physical function, whereas activities/participation was a separate broad domain. The final definition of physical function in the COMFORT project reflected the importance of participation to patients, as well as the considerable overlap between activities and participation, as identified in the qualitative research and through PPI input.

7.5.1.2 Condition-specific COS

Pain and function are commonly included as core domains in disease-specific OMERACT COS (72, 250, 350, 381). Comparatively, no OMERACT COS to date have specified satisfaction as a core domain. Additionally, despite being strongly recommended by OMERACT, few OMERACT core domain sets have included societal/resource use domains. The OMERACT core domain set for SSc-RP includes hospitalisation or need for urgent intervention for RP as a mandatory domain within the core area of societal/resource use, although the results from the Delphi consensus process have not yet been published (382). The lack of consideration for societal/resource use domains in COS also appears to reflect the wider COS literature. In the latest systematic review of COS published in the COMET database, 210 out of 589 (35.7%) published COS included domains from this category, compared to 91.3% of COS that included pathophysiological domains and 73.3% that included life impact domains (94).

Beyond OMERACT, pain and physical function are commonly included core domains across other MSK conditions. In a systematic review of 40 COS (383), pain was included as a core domain in 90% and physical function was included in 88%. In contrast, participation (recreation, sport, and work) was included as a core domain in 15% of COS, role functioning (social, work, household, and caregiving) in 8%, and social functioning in 5%. Delivery of care, which included satisfaction with outcome, acceptability and availability, withdrawal from treatment, appropriateness of treatment, process implementation, and service outcome and utility (usefulness, value, or benefit perceived from treatment), was included as a core domain in 18% of COS. The inclusion of participation (within the definition of physical function) and treatment satisfaction as core domains in COMFORT may reflect experiences that are unique to patients with foot and ankle disorders in RMDs, compared to other health conditions.

Qualitative findings throughout this thesis have demonstrated that the consequences of foot and ankle disorders are far-reaching, with the impact on weightbearing hugely disruptive to almost all aspects of patients' lives. Additionally, as discussed throughout this thesis, satisfaction is particularly pertinent in this area given the lack of high quality evidence for foot and ankle treatments. In some cases, foot and ankle treatments might be effective (e.g. in reducing pain and improving function), but deemed unsatisfactory by patients. This was particularly evident in qualitative findings relating to footwear interventions in Chapters 3 and 5; for example, some patients expressed that their shoes were comfortable, but did not look good. As treatment satisfaction is associated with patient adherence (384), its inclusion as a core domain is important in this context.

7.5.1.3 COMFORT in the context of whole-disease COS

Given the broad scope of the COMFORT core domain set in terms of the RMDs covered, it must also be considered in the context of relevant whole-disease COS. The concept of overlapping core domain sets was introduced in Chapter 4, in relation to the SSc-RP core domain set. In summary, the primary intention of the COMFORT core domain set is for outcome measurement in foot and ankle-specific studies, e.g. where the researchers are specifically assessing the effectiveness of an intervention that targets the foot and ankle. This could be a pharmacological (e.g. local steroid injections into the foot and ankle), conservative (e.g. foot orthoses, footwear, or foot-specific exercises), or surgical (e.g. foot or ankle joint fusion or replacement) intervention. For studies investigating the effectiveness of a treatment for an RMD overall, use of the disease-specific COS is crucial for data synthesis across the disease. The core domain set for foot and ankle disorders can then be measured in addition if researchers are also interested in these outcomes. As the COMFORT core domain set includes the circumstance-specific broad domain structural pathology (encompassing inflammation), there is overlap with OMERACT core domain sets that include disease activity (71, 233, 250). These whole-disease COS focus on systemic disease activity/inflammation, whereas COMFORT is concerned with disease activity/inflammation in the foot and ankle. This is particularly important in IA, where foot and ankle symptoms can persist even when clinical remission of disease activity is achieved (18).

An example of how COMFORT might intersect with a whole-disease COS (PsA) (233) is presented in Figure 7.1.

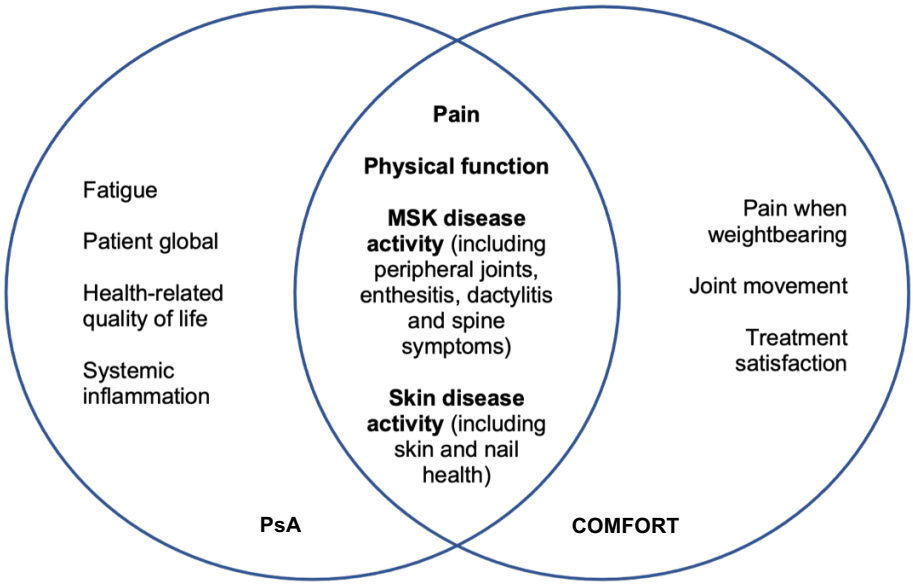


Figure 7.1: Overlap between PsA and COMFORT core domain sets

7.5.1.4 Broader COS development

The COMFORT core domain set can also be considered in light of broader COS development initiatives endeavouring to standardise the measurement of specific domains, For example, INTEGRATE-Pain has developed consensus on separate, overarching core domain sets for acute, the transition from acute to chronic, recurrent/episodic, and chronic pain, through a three-round Delphi process, to encourage standardisation within and across different pain categories internationally (385). Recurrent/episodic pain and chronic pain are particularly relevant to patients with foot and ankle disorders. The final core domain set for chronic pain included pain (including measures of pain intensity and pain interference as subdomains), quality of life, and activities of daily living (385). Pain intensity and activities of daily living are covered in the COMFORT core domain set. Pain interference and quality of life could be measured in foot and ankle studies in addition to the foot and ankle core domain set, recognising that they are meaningful in broader pain research. This could potentially improve the integration of findings from foot and ankle research into broader pain policy and future meta-analyses of pain interventions.

7.5.2 Original contributions to the literature

Findings from the research within this doctoral programme of work offer several original contributions to the literature. Historically, foot and ankle disorders in RMDs have been overlooked. For example, despite similar prevalence to that of knee OA and more than

double that of hip OA (386, 387), the evidence base for foot OA lags significantly behind other commonly affected sites and there is a lack of guidance for clinical management (388). Evidence for treatments for other RMDs where the foot and ankle are affected is also lacking; for example, a recent BSR guideline for the management of foot health in IA found limited evidence for interventions for foot problems in RA and none for interventions for foot problems in SpA (60). The need for high-quality RCTs for foot and ankle problems is consistently highlighted as a priority in research agendas (60, 375, 388). As foot and ankle research continues to build momentum, the need to measure standardised and meaningful domains is crucial, as "clinical trials are only as credible as their outcomes" (389). This thesis presents the first set of domains, achieved through consensus, to be measured and reported in future foot and ankle research across multiple RMDs. The core domain set itself, therefore, is a novel contribution to the literature.

To the candidate's knowledge, the secondary analysis of focus groups in Chapter 4 is the first qualitative study to have explored the experiences of foot and ankle problems among patients with SSc, whilst the qualitative interview study in Chapter 5 is the first to have explored the experiences of foot and ankle problems among patients with adult JIA, different types of SpA (AS, enteropathic arthritis), pseudogout and mixed CTDs. The preceding synthesis of existing qualitative studies highlighted that ethnicity was poorly reported, there was no inclusion of patients from LMICs, and limited inclusion of patients with a first language other than English. By including patients from LMICs, with a diverse range of ethnicities and first languages, the qualitative interview study addressed these areas, filling gaps in the literature, and produced novel findings that can be applied to clinical practice and wider research. This is discussed in more depth in section 7.5.3.2.

Finally, the methodological areas explored in this programme of work generated novel learning points that could be applied to COS development studies in the future, including incorporating the perspectives of non-English speakers into qualitative work, Delphi recruitment and retention methods, and the conduct of a fourth, definitive Delphi round. These methodological areas are discussed further in the subsequent sections of this chapter.

7.5.3 Reflections on the methodological process

Two areas of discussion will now be examined in more detail, bringing together the candidate's reflections of the overall methodological process throughout the project as a whole.

7.5.3.1 Working within a standardised framework

The COMFORT programme of work is supported by OMERACT, an international outcomes organisation that is recognised for driving COS development in RMDs by developing standardised methodological processes and frameworks. Areas where these processes conflict with the best available COS methodology evidence (e.g. order of the presentation of domains) were outlined in Chapter 6 (section 6.4.2). OMERACT has produced multiple resources to aid COS developers. For Delphi consensus studies, COS developers are expected to use OMERACT templates for each page of the survey and the PIS, in addition to ordering the survey by OMERACT core area with an accompanying definition. This presented a conflict between the need to adhere to OMERACT standards and wanting to be inclusive and increase COMFORT-Delphi's recruitment and retention rates by keeping the surveys short, clear, and readable, especially for participants who did not speak English as a first language. For example, OMERACT states that the title *core area: pathophysiological manifestations* must be presented at the start of a Delphi study, with the following accompanying core area definition:

> *"These are typically defined as observable or measurable phenomena associated with a health condition. These may include clinical signs, symptoms, and laboratory or imaging findings. They are a disease or condition's concrete, identifiable features that reflect its presence and characteristics. Examples of pathophysiological manifestations include symptom severity, joint count, range of motion, or laboratory findings such as sedimentation rate. Imaging outcomes looking at joint space or existing subchondral cysts are also pathophysiological manifestations."* (390).

This text was deemed 'difficult to read' (corresponding to undergraduate level reading age) when entered into a validated readability calculator (391). The term *pathophysiological manifestations* is technical, and is the first term that participants see when starting an OMERACT Delphi survey that uses the template. The information about imaging outcomes in the core area definition is also highly technical. Whilst there was no flexibility to change the title, it was argued that information about each OMERACT core area was not crucial for understanding how to complete the Delphi, or why the study was important. A compromise was made to include OMERACT core area definitions as a PDF link for any participants who wanted to understand more about these.

Additionally, the OMERACT Delphi survey template asks participants to select the collaborator group that best represents their affiliation or role in relation to their involvement with OMERACT, followed by the OMERACT classification of collaborator groups according to the '11 Ps': patient research partners, patients/consumers, caregivers, and patient groups, payers and purchasers of health services, payers/funders of research, policymakers, principal investigators, producers and commissioners of guidelines, product makers, program managers, providers, public, publishers, and others (63). It was contested that this terminology may well be confusing and off-putting to participants who are unfamiliar with OMERACT; consequently, more accessible terms were used in the COMFORT Delphi study. Further, in an attempt to be inclusive the OMERACT participant demographics template lists ten gender identities, but this excludes multiple other gender identities that have been described (392, 393). A pragmatic decision was therefore made to list the four most common gender identities, with a 'prefer not to say' option as well as an option for participants to describe their own gender.

In recent years, OMERACT has demonstrated commitment to improving EDI (394). For example, the organisation has implemented reduced costs/fee waivers for LMIC participants and placed more emphasis on cultural competency at consensus workshops. Improving the accessibility of OMERACT materials also has the potential to improve EDI. Potential survey readability issues were disputed throughout the design of the Delphi consensus study, and whilst the final surveys represented a more inclusive approach as a result of coming to a compromise with OMERACT, the process subsequently took longer due to OMERACT's reviews of wording that did not adhere to standard templates.

Following a formalised process where there is a lack of flexibility to be innovative can inevitably slow process, and any large, well-established organisation presents a risk of power dynamics, whereby overall decision-making can be influenced by more senior members, discounting marginalised perspectives. Balancing a formalised process with undertaking a PhD, which by definition is an original contribution to knowledge, was a challenge throughout this work. However, challenging OMERACT's processes, with the aim of improving inclusivity, strengthened this research.

7.5.3.2 Learning from international collaboration

Inclusion of a diverse range of perspectives in COS development is crucial to ensure that core domains are meaningful to the diverse range of patients who are affected by foot and ankle disorders in RMDs. In COMFORT, diversity was a central consideration

in the qualitative interviews and the Delphi consensus study. In addition to maximising inclusion of participants of different ethnicities and with first languages, the qualitative research was conducted in Arabic with patients based in Egypt. Recruitment was facilitated by two local collaborators at different hospital sites, who conducted the interviews with each other's patients. These collaborators attended an early OMERACT Foot and Ankle Working Group meeting focusing on the feasibility of the COMFORT project. The candidate's initial plan was to involve patients from three continents (Europe, North America, and Australasia), based on OMERACT's guidelines (63). However, during the meeting, one of the collaborators pointed out that "patients in Egypt have foot problems, too!". After this meeting, the candidate contacted both collaborators about the planned qualitative research, and both were keen to facilitate this. As qualitative rheumatology research in non-Western cultures is uncommon (395) and neither collaborator had qualitative research expertise, this posed challenges. The candidate considered limiting recruitment to English-speaking and digitally literate participants from Egypt, to be able to conduct the interviews online herself. However, the strengths of including more authentic cultural insights outweighed the limitations identified. One of the Egypt-based collaborators involved in the qualitative work also facilitated recruitment of English-speaking patients from clinics in Egypt to the Delphi study. This led to the retention of 12/14 (85.7%) of patients from Egypt between Round 1 and Round 4.

Recommendations for future COS developers based on these experiences include:

- reaching out as early as possible to potential collaborators who are based in LMICs;
- meeting the collaborators online regularly (before, during and after data collection) to facilitate their understanding of the aims and conduct of the research;
- ensuring availability to address any issues as they arose (for example, the candidate made it clear she was available anytime for contact through email or mobile phone, including Sundays, which are working days in Egypt).

Due to resource limitations, the Delphi was restricted to the English language. It is highly likely that translation of the Delphi study into other languages would have further improved the diversity of participants, especially patients (378). The potential of using artificial intelligence (AI) to translate the Delphi surveys into other languages was considered during the design of the study. However, whilst AI can be cost-effective way of translating text quickly, the quality of translations can vary (396). Prioritisation of foot and ankle domains for a COS is a complex area whereby loss of context,

misinterpretation, and lack of cultural sensitivity could occur. Input from native speakers with subject matter expertise and Delphi experience would therefore be needed. Further work is needed to understand the full potential of AI in COS development; OMERACT has outlined in a recent newsletter that it is currently exploring how AI can support the COS development process and improve efficiency.

7.6 Strengths and limitations

The strengths and limitations of each phase of the project are detailed in Chapters 3-6. The strengths and limitations of the overall project will now be discussed.

7.6.1 OMERACT

The challenges of completing a PhD that involved developing a COS within the confines of a well-established organisation were explored in the previous section. However, there were also major strengths to working with OMERACT that benefitted this overall project, which should not be overlooked. COS developed with OMERACT have credibility within the rheumatology community, strengthening the international visibility of COMFORT. Uptake of the COS in the field of foot and ankle RMD research is therefore more likely. OMERACT provided infrastructure and logistical assistance, as well as methodological support. This saved time and resources, e.g. through the organisation of meetings, provision of experienced facilitators, access to comprehensive training, and a platform to engage with other OMERACT Working Groups, to share best practice and collaborate. OMERACT's endorsement of COMFORT also undoubtedly enhanced recruitment of participants during both the qualitive interview study and Delphi consensus study. As a leading example of meaningfully involving patients in research, working with the OMERACT PRP Network was also a key strength. This is considered further in the next section.

7.6.2 PPI

PPI input was critical from the outset of COMFORT and patients' lived experiences were integrated into every phase of the project. Early input from PPI shaped this work. Locally, PPI contributors identified that foot and ankle problems are an important and highly neglected area, and that research addressing the lack of evidence for treatments was worthwhile. Internationally, PPI contributors within the OMERACT PRP Network advocated for a foot and ankle COS based on their lived experiences, and were pivotal in driving overall OMERACT support and endorsement of the Foot and Ankle Working Group.

Planning for PPI in the COMFORT project took place during the funding application. PPI activities were costed for according to NIHR guidelines (86), and PPI contributors were recruited to the project through the NIHR Leeds BRC and OMERACT. Crucially, one OMERACT PPI contributor, with lived experience of foot and ankle problems in RA and extensive knowledge of COS development was one of the candidate's PhD supervisors, ensuring that every decision relating to this project, and the candidate's personal growth as an independent researcher, was grounded in the patient perspective.

Examples of how PPI inputted into each study, and the impact that their input had, were informed by reflective logs and are discussed in each individual chapter. Overall, PPI input ensured that the core domain set developed during this project is relevant to patients, and has genuinely taken into consideration the patient perspective throughout. Given that this was an international project with a diverse range of patient participants, one aspect of PPI that could have further strengthened this work would have been the involvement of PPI contributors from LMICs. Whilst the OMERACT PRP Network is continuing to diversify, lack of engagement from patients in LMICs, in PPI roles, needs addressing in future work. The impact of PPI input into the COMFORT project has been assessed by the OMERACT PRP Network. PPI contributors who have been involved in the COMFORT project (including those within the PAG and members of the wider network) have evaluated their own participation, reflecting on their engagement using the conceptual patient engagement in research (PEIR) framework. This framework covers themes such as procedural requirements, convenience, contributions team interaction, the research environment, support, and feeling valued.

Findings from a survey of PPI participation across all OMERACT Working Groups were presented at the OMERACT 2025 workshop (manuscript in preparation). Feedback highlighted that PPI contributors involved in the COMFORT project felt valued as equal members of the research team, e.g. through mutual respect, being acknowledged for their input, given opportunities to share their own experiences, lead meetings and presentations, and being compensated appropriately. PPI contributors also noted that they appreciated the candidate's personalised and flexible approach to communication, and felt they had improved skills in meaningfully contributing to the co-design of qualitative interview schedules and in analysing qualitative data. From a condition perspective, PPI contributors highlighted that they had a better understanding of the wider consequences of foot and ankle disorders across different RMDs.

7.6.3 Involvement of key contributor groups

The range of contributor groups (patients with different RMDs, and HCPs/researchers from different professional backgrounds) involved in this work is one of its strengths. However, the qualitative research underpinning the development of the core domain set only included patients, which is a limitation. Whilst HCPs and researchers from different professional backgrounds were involved in the Delphi consensus study, and were able to provide free-text comments, their views would have been better represented through qualitative interviews or focus groups. The scoping review undertaken in preparation for this programme of work indicated which domains are likely to be important to HCPs/researchers, but not why they are important. This would have provided a more comprehensive understanding of domain priorities, particularly if different countries had been represented, and this may have led to different or additional domains being included in the Delphi consensus study.

7.7 Implementation

In order to facilitate meta-analyses and enable the translation of findings based on relevant domains to clinical practice and patient care, a COS must be implemented successfully into the design, conduct, and reporting of research studies. However, uptake of COS is variable (82, 242-245). In order to understand how to implement COMFORT successfully, potential barriers to implementation must be identified. COS implementation has recently been explored from a behavioural science theory perspective; Matvienko-Sikar et al. (2024) proposed the use of the Behaviour Change Wheel and COM-B model to identify methods to facilitate COS adoption (397). COM-B conceptualises behaviour as the result of three interacting components: capability, opportunity, and motivation. In the case of COMFORT, researchers working in this clinical area may lack knowledge and awareness about the core domain set (capability), support (opportunity), and/or motivation to use the core domain set (e.g. if they don't understand why it should be used, or if they have uncertainty about OMIs). Potential barriers to COS uptake were outlined in Chapter 2 (section 2.8). These barriers, along with proposed strategies that have been or will be adopted to improve up take of the COMFORT core domain set, will now be further explored.

7.7.1 Visibility

An obvious barrier to COS implementation is lack of awareness that the COS exists. From the beginning of this work in 2018, and with the support of OMERACT, the

COMFORT project has been highly promoted. It is included on the OMERACT website (398), registered in the COMET database (88), and a protocol is published in *Trials* (87). A preliminary COS dissemination strategy was developed at the beginning of this project and has been refined throughout, based on established methods (65, 246, 399), with ongoing PAG input (Table 7.2).

Table 7.2: COMFORT dissemination strategy

Method of dissemination	**Strategies**	**Examples of how strategies have been/will be utilised in the COMFORT project**
Patient/PPI events	Engagement with PPI contributors, working with the committees and communities they are involved with to promote COS uptake; dissemination through newsletters and articles to charities and patient organisations.	Findings from each phase of the COMFORT project have been shared through the international OMERACT PRP Network, through newsletters and workshops, and through social media. A local NIHR Leeds BRC PPI dissemination event is planned for 2026, to provide feedback to the PPI members who helped to conceive the COMFORT project before funding was sought, and to those involved in co-designing the Delphi study. A summary of the qualitative interview findings was sent to patient participants who indicated their interest, and a lay summary of Delphi findings is in preparation for feedback to the patient organisations/support groups involved in recruitment. This is being co-produced with PPI contributors.
Relevant websites	Publication of the COS on relevant websites, where it is possible to search for a specific COS.	COMFORT is registered on the COMET and OMERACT websites, where details of all phases of the project have been published. The final core domain set will be accessible on both websites.
Publications	Publications of findings from each stage of the research in relevant, peer-reviewed academic journals.	The scoping review, qualitative synthesis, secondary analysis of qualitative focus groups and qualitative interview study have been published in *Seminars in Arthritis and Rheumatism*, a relevant, peer-reviewed academic journal. The Delphi study manuscript is in preparation and will be submitted to the same journal. Seminars in

		Arthritis and Rheumatism is the official OMERACT journal, where researchers can find rheumatology COS studies (open access).
Social media	Dissemination of the COS on social media.	Findings from each phase of the COMFORT project have been shared on social media platforms by members of the PAG and wider OMERACT Foot and Ankle Working Group, including PPI contributors.
Conferences and meetings	Dissemination at professional conferences and meetings	Findings from the COMFORT project have been shared extensively at local, national and international conferences across the disciplines of the foot and ankle, rheumatology, and outcome measurement. Conferences and meetings have primarily involved disseminating to HCPs and researchers, although some conferences have also included patient delegates (400, 401). Findings from the scoping review and qualitative studies have been disseminated at the following conferences and meetings, through oral and poster presentations: • Royal College of Podiatry conference • Northern and Yorkshire Rheumatology meeting • NHS England Yorkshire and Humber meeting • Academy of Medical Sciences Clinical Academics in Training conference • NIHR Academy conference • NHS Research and Innovation conferences and meetings • NHS podiatry and rheumatology meetings • EULAR Congress

		• EULAR foot and ankle SIG meetings • American College of Rheumatology Convergence • OARSI International Foot and Ankle OA Consortium meetings • OMERACT workshops and meetings The candidate has been invited to present findings from all phases of the COMFORT project at the BSR conference in 2026, and at the OMERACT workshop in 2027. The intention is to co-present the project with a PPI contributor (PR). Abstracts will also be submitted to the OARSI 2026 World Congress, EULAR 2026 Congress, and PROMs Annual Research conference. All members of the PAG, and the clinical collaborators based in different countries who were involved in recruiting for the qualitative interviews and Delphi study, will be invited to share the overall project findings at local conferences and meetings.
Journals	Engagement with journal editors and Cochrane review groups to raise awareness of the COS.	Journal editors are part of the OMERACT Foot and Ankle Working Group and wider OMERACT community. Others participated in the Delphi study (e.g. editors for the *Journal of Foot and Ankle Research* and *Rheumatology*).
Funding bodies	Engagement with funding bodies, asking them to consider including the need to use the COS in	The COMFORT project was funded by the NIHR and a report will be prepared at the end of the fellowship funding period, asking the funder to consider including the need to use the COMFORT core domain set in relevant future funding applications.

	applications for financial support for future studies.	
Guideline developers	Engagement with guideline developers to facilitate inclusion of the COS in recommendations on the management of patients.	Key foot and ankle guideline developers, in the UK and internationally, have been involved in the development of the COS, including members of the PAG and the wider OMERACT Foot and Ankle Working Group.

However, dissemination alone could be considered a passive approach that does not address behaviour change amongst researchers (398). It has been proposed that for the COM-B component ‘capability’, a strategy could include education/training, e.g. workshops, webinars, or guidance materials to inform researchers about the core domain set and how to use it.

7.7.2 Facilitating and monitoring uptake

Engagement with the core domain set is crucial for uptake. Ongoing engagement with key contributors, including patients, HCPs, and researchers, increases the relevance and legitimacy of a COS. Engagement has been prominent throughout the COMFORT project, recognising that if people help to design a COS, they are potentially more likely to then use it (65). The PAG and wider OMERACT Foot and Ankle Working Group include world-leading foot and ankle RMD researchers who will be key users of this core domain set in the future. All members are highly influential in the field of foot and ankle disorders in RMDs, with the potential to facilitate uptake of the core domain set among other researchers and HCPs.

The main outcome of this project is a core domain set that can now be used in clinical trials and observational studies for foot and ankle disorders in RMDs. However, a barrier to implementation is choice and application of the OMIs to measure the core domains. Researchers may be apprehensive to engage with the core domain set if there are no agreed, high-quality and feasible OMIs for the core domains. This will be addressed in future work (section 7.8.2). The previous scoping review of domains in existing clinical trials and observational studies also extracted a range of candidate OMIs, which are presented in the supplementary material of the article (73: supplementary table 2). As a preliminary step towards the next phase of developing a COS for foot and ankle disorders in RMDs, OMIs have been tabulated for each core domain, and an overview of each OMI has been summarised Appendix T. Additionally, OMIs that have been assessed in relevant systematic reviews of measurement properties, which will inform the next phase of this work, were introduced in Chapter 1 (section 1.5.1) and are tabulated in Appendix A. HCPs and researchers can therefore select from a range of OMIs to measure the core domain set as it stands.

Whilst some foot and ankle RMD researchers have been engaged throughout the COS development process, others may consider that the COMFORT core domains are irrelevant or too generic, have other preferences, or feel it is too burdensome to measure the core domain set. For example, as discussed in Chapter 6, some researchers may deem joint movement irrelevant or too difficult to measure. Increased

understanding among foot and ankle researchers of the importance of using the core domain set is therefore important. Endorsements from funders and journals may help to overcome this, creating an expectation or requirement for the core domain set to be used. Various government research funders in the UK now endorse the use of COS in submissions to their awards (65, 402), and some reporting guidelines have included consideration of a COS (403).

The need for further work with ethics committees, journals, trial registries, and protocol reporting initiatives to embed COS reporting in relevant guidance and templates has been identified (404). Uptake of the core domain set will be monitored using citation analysis and by reviewing trial registry entries. This may include monitoring COS use in trial registrations and publications, surveying researcher awareness, and examining barriers and enablers over time.

7.8 Future research

This doctoral project has involved working with an international network of patients, HCPs, and researchers who have extensive experience in foot and ankle disorders in RMDs, and in wider outcome measurement. With the support of the international outcomes organisation OMERACT, the PAG and the wider Foot and Ankle Working Group have been united in their goal to improve outcome measurement in this field. Ongoing collaboration with these networks will support future work in this area. Four key areas for future research have been identified: further exploration of contextual factors, developing a core outcome measurement set, COS methodology, and a review of the COMFORT scope. Each of these areas will now be explored in more detail.

7.8.1 Exploration of contextual factors

Contextual factors were briefly explored in the qualitative interviews, whereby it was indicated that age, sex, ethnicity, disease duration, geographic location, treatment experiences, and healthcare systems could influence the results of foot and ankle studies and should be considered alongside a core domain set. Other previous OMERACT Delphi studies have included contextual factors for prioritisation in Delphi studies (72, 320). Further work is needed to understand the contextual factors in the area of foot and ankle disorders in RMDs, alongside work that is underway to produce a consensus-based set of endorsed contextual factors in rheumatology (405).

7.8.2 Developing a core outcome measurement set

A COS involves developing both a core domain set and a core outcome measurement set. Methods for assessing OMI measurement properties and selecting OMIs for a COS were discussed in Chapter 2 (section 2.7). As discussed above, candidate OMIs for the COMFORT core outcome measurement set have been identified from an existing scoping review of clinical trials and observational studies (73). However, this scoping review was published in 2020, and additional OMIs may have been developed since. Additionally, the scoping review excluded studies of interventions for non-systemic MSK conditions (e.g. plantar heel pain, tendinopathies, and hallux valgus), therefore candidate OMIs may have been missed. The formal process of selecting OMIs for core domains using the OMERACT framework is extensive, with each core domain assessed individually. This process includes assessing the psychometric properties, feasibility and acceptability for each candidate OMI, followed by an individual consensus process. For instance, the OMERACT Shoulder Working Group was established in 2015, with a core domain set for shoulder disorders endorsed in 2018. Preliminary work into OMIs for two of the core domains, pain and physical function, was published in 2023 (149), with further work on these particular domains conducted at the OMERACT 2025 workshop. OMI selection in COMFORT presents a particular challenge, given the broad range of RMDs and interventions included within the scope. Multiple candidate OMIs are available for most of the core domains, but some are RMD- or intervention-specific. Further work is therefore needed to establish the validity of these OMIs in different populations. Additionally, many of the OMIs identified are composite OMIs, measuring both core and non-core domains, with combined scoring (406). Lessons will inevitably be learnt from similar COS development groups as they move through this complicated process. Establishing *how* each core domain should be measured is a significant area for future work, enabling researchers to measure the core domains with feasible and acceptable OMIs that have good psychometric properties. This will further support the synthesis of evidence for treatments for foot and ankle disorders in RMDs.

7.8.3 COS methodology

The Delphi study sought feedback from participants to inform the conduct of future Delphi studies (Chapter 6, section 6.3.14.2). Learning points for future COS developers based on this feedback are in line with existing recommendations (65, 193). These include designing a survey that is easy to follow and concise, with clear information about the expected time commitment, purpose of the Delphi, and nature of the process.

The language within a Delphi survey should be in plain English, and the survey should be piloted across all devices. Additionally, recruitment and reminder emails should be personalised.

Several methodological questions also arose from the Delphi study (Chapter 6). In the feedback, one participant suggested the need for qualitative work with patients. This had been conducted to inform the Delphi, but not presented or alluded to in the survey itself. This leads to the methodological questions: *Does presenting previous COS development work influence ratings in a Delphi study, and how should previous COS development work be presented?* For example, findings from previous work could be presented at the start of the survey, or by each domain.

The Delphi study conducted in this project contained a fourth round where participants were asked to rate each domain as *in* or *out* of the final core domain set. Whilst this is strongly recommended by OMERACT (63), one previous study indicated that multiple rounds of Delphi may not be advantageous (212). Another methodological question arising from this work, therefore, is: *What impact does a fourth Delphi round have on consensus?* The fourth round of the Delphi study also presented domains in the context of hypothetical clinical trials, which may have aided participants' understanding of core domains in comparison to previous rounds where participants rated domains in terms of their importance with no clinical trial examples. The methodological questions arising here are: *What is the impact of presenting domains in the context of example clinical trials on consensus and on participants' understanding of COS?*

During the Delphi study, patients were sent personalised reminder emails featuring a photograph of a PPI contributor and a quote detailing why they were taking part in the COMFORT project. This concept could be developed further, e.g. by involving more PPI contributors with quotes tailored to specific RMDs. Further, the same concept could applied to HCP/researcher reminder emails, either with patient quotes or with a HCP/researcher equivalent email. The methodological question here is: *What is the impact of personalised recruitment/reminder emails, tailored to the participant's condition, on recruitment and retention in Delphi studies conducted as part of COS development*?

Methodological questions of this nature could be addressed by conducting Study Within A Trial (SWAT), also known as a 'nested trial', 'embedded trial', or 'trial within a trial'. A SWAT is a methodological sub-study whereby an intervention to improve the conduct of a trial is tested in the context of an ongoing trial. The aim of a SWAT is to generate new knowledge to improve the design and delivery of future trials (407). SWATs can be randomised trials themselves (i.e. trial within a trial) or non-randomised

evaluations, such as comparing electronic data collection methods alongside existing paper based data collection. In the context of COS development, SWATs can be embedded within the process of developing a COS, e.g. in Delphi surveys. For example, researchers developing a COS for pulmonary sarcoidosis research registered a SWAT that assesses the effects of providing video guidance on how to complete a Delphi survey and an email helpdesk in an online Delphi survey on response and retention. This is published within the SWAT repository, an online database that publishes SWAT protocols (408). As this was the candidate's first experience of conducting a Delphi consensus study, embedding a SWAT within the Delphi study was considered to be beyond the scope of the PhD. Future Delphi studies conducted as part of COS development could formally test the impact of various methods employed during the Delphi study on consensus, and recruitment and retention, through SWAT methodology.

7.8.4 Review of the COMFORT scope

The broad scope of COMFORT was continually revaluated throughout this programme of work. However, there are some conditions that were not included in the scope that are potentially relevant. Like most COS within rheumatology, this work focused on adults with foot and ankle disorders in RMDs. Previous research has demonstrated that foot and ankle problems are also prevalent in children and young people, such as those with JIA (409-411). Current foot health in IA guidelines apply to adults, children and young people, and domains measured in existing JIA foot studies are similar to those in the core domain set (e.g. pain, function) (60). Recent work also indicates that foot involvement occurs in a preclinical phase of RA (412), signifying that the COS may be relevant for future studies focusing on preventive interventions for those at risk of developing the condition, prior to diagnosis (413). The methodological challenges of developing a COS for prevention have been recognised (e.g. presenting domains that describe the effects of a treatment to patients who do not have a disease yet) (218). Future work is needed to establish domain priorities among patients in these additional relevant groups, as well as HCPs and researchers who work in these areas, and whether these populations can be integrated into the existing COS.

7.9 Implications for clinical practice

This work has implications for clinical practice. In the long term, uptake of the core domain set will lead to improvements in the quality of evidence for treatments, and

ultimately more informed decision-making for patients and HCPs in clinical practice settings. Some COS developers have specified that the intention of their COS is for use in clinical studies *and* routine clinical practice (414, 415), and the number of COS with this dual focus is increasing (416). Following OMERACT methodology, the COMFORT core domain set was developed with the intention of being measured and reported in research studies. As a result, Delphi participants were not asked to rate the importance of domains to be measured in the context of routine clinical practice, although it is plausible that their priorities would be the same or similar regardless of the context.

The COMFORT core domain set reflects *what* is important to measure and is therefore relevant to clinical consultations, shared decision-making between patients and HCPs regarding treatment options through enhanced understanding of what patients want treatments to achieve, and in assessing whether or not a treatment has been effective. Understanding which domains are important to patients and HCPs can therefore inform the design of information resources, guidelines, audits and registries (417).

Much like in research, there is variation in foot and ankle outcome measurement in routine clinical practice (418). The next stage of this work (*how* to measure domains) may differ between clinical practice and research, particularly in terms of feasibility of OMIs. It has not yet been established whether an existing COS, developed for research studies, can be used in routine clinical practice, or whether adaptation and validation is required (415). A recent systematic review exploring the scope, contributor involvement and development methods in COS intended for routine clinical practice identified that these were more likely to include OMIs than COS for research, with OMI feasibility and availability informing the selection of core domains (415). Interestingly, the main method for identifying important domains was to review research, rather than routine health records, and less than half of the COS identified involved patient input. The need for improving the methodology for developing COS for use in clinical practice has been highlighted (415). Future engagement with clinical commissioners and a wider range of HCPs is important when considering implementation of the COMFORT core domain set into clinical practice settings.

Whilst this was an international project, consideration has been given to the specific implications for NHS practice. Outcomes are increasingly used in NHS clinical practice to evaluate health service provision, aid the management of patients, and quantitatively measure the impact of clinical interventions. However, outcomes are not always measured for foot and ankle disorders in RMDs in the NHS and measures that are used are not typically standardised. Standardising outcome measurement would

facilitate comparison across NHS providers and regions, whilst also underpinning quality improvement and resource allocation. This aligns with a shift towards value-based healthcare, defined as “the equitable, sustainable and transparent use of the available resources to achieve better outcomes and experiences for every person” (419). Value-based healthcare focuses on delivering outcomes that matter most to patients, so that the resources available have maximum impact (415). The COMFORT project has established the domains that matter most to patients with foot and ankle disorders in RMDs. As discussed above, core domain set could potentially inform *what* is measured in the NHS, with further work needed to establish *how* to do this.

The diversity of patients included throughout this programme of work also has implications for the NHS. Around 4.2 million (8%) people in England and Wales do not speak English as their first language (420); 726,000 (1.3%) state that they cannot speak English well, whilst 138,000 (0.3%) cannot speak English at all. Patients who do not speak English are known to have worse health outcomes, delays in diagnosis and barriers in accessing care (421). Their experiences and priorities may differ, therefore, of those who do speak English well. Health inequalities between ethnic minority and white groups are also well established; people from some ethnic minority groups are more likely to have worse health outcomes and to report poorer experiences of using health services than their white counterparts. Some groups, e.g. people from Bangladeshi and Pakistani communities have the poorest health outcomes across a range of indicators (422). In the 2021 census, 19% of people in England identified themselves as belonging to a non-White ethnic minority group.

Whilst data relating specifically to foot and ankle disorders in RMDs is lacking, there are known ethnic disparities in wider RMDs where foot and ankle problems are common. The National Early Inflammatory Arthritis Audit (NEIAA) indicated that minority ethnic groups are less likely to achieve remission at three months for early IA, despite there being no substantial differences in several process metrics (e.g. referral speed and time from referral to assessment) (423). Additionally, Asian patients are about half as likely as White patients to be prescribed biologic or targeted synthetic anti-rheumatic therapies early after RA diagnosis (424). Further, the profile of MSK pain among people from Indian, Pakistani, Bangladeshi and African Caribbean communities differs from that among people of a white background, with a higher prevalence of MSK pain in multiple joints reported among ethnic minority groups (425). However, there is less published research about how people from minority backgrounds experience care (426, 427). This issue was reflected in the qualitative synthesis in Chapter 3; Chapters 5 and 6 endeavoured to address EDI. Findings from the qualitative research and Delphi consensus study reflect the perspectives of a

diverse range of patients, including those from non-White ethnic minority groups, and those who do not speak English well, or at all. Considering the perspectives and experiences of these patients in NHS clinical practice is a positive step towards addressing health inequalities.

7.10 Conclusion

The COMFORT project achieved its aim of developing a core set of domains for foot and ankle disorders in RMDs, to be measured and reported in future research. A rigorous evidence-based approach, using a mix of methods, was employed to address each project objective successfully. The project's first objective was to identify and explore domains of importance to patients with foot and ankle disorders in RMDs. This was achieved through a synthesis of existing qualitative studies, a secondary analysis of qualitative data and a primary qualitative interview study. The project's second objective was to achieve multidisciplinary, multi-contributor and expert international consensus on a core domain set for foot and ankle disorders in RMDs. This was achieved through a Delphi consensus study with patients, HCPs, and researchers. The central role of patient involvement in this work helped to ensure that it prioritised the views of people living with foot and ankle disorders in RMDs throughout. Findings provide various novel contributions to the literature, including:

- The first set of domains for measurement in future foot and ankle RMD research;
- Novel understanding of the experiences of patients with foot and ankle problems in specific RMDs, of differing races/ethnicities and from LMICs that had not been explored in previous qualitative literature;
- Original contributions to the COS methodology evidence base.

Findings also have important implications for clinical practice and future research. The key output from this project is a set of core domains that can now be measured in research, and can be used to inform clinical consultations as well as the design of information resources, guidelines, audits and registries. The next step will involve establishing which OMIs should be used in future research, ensuring further standardisation of outcome measurement for foot and ankle disorders in RMDs. This will continue to advance the quality of evidence for foot and ankle treatments, ultimately leading to improved patient care.

References

1. Redmond, A. (2025). Ankle and Foot Systemic Anatomy. In Standring S, ed. Gray's Anatomy: The Anatomical Basis of Clinical Practice (43rd ed.): Elsevier, 2025. p1524-1536.

2. Global, regional, and national burden of other musculoskeletal disorders, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol. 2023;5(11):e670-e82.

3. Cross M, Smith E, Hoy D, Carmona L, Wolfe F, Vos T, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis. 2014;73(7):1316-22.

4. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis. 2014;73(7):1323-30.

5. Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014;73(6):968-74.

6. Smith E, Hoy D, Cross M, Merriman TR, Vos T, Buchbinder R, et al. The global burden of gout: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014;73(8):1470-6.

7. Smith E, Hoy DG, Cross M, Vos T, Naghavi M, Buchbinder R, et al. The global burden of other musculoskeletal disorders: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014;73(8):1462-9.

8. World Health Organization (WHO). Musculoskeletal health. 2022 https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions Accessed 1st February 2023.

9. Grondal L, Tengstrand B, Nordmark B, Wretenberg P, Stark A. The foot: still the most important reason for walking incapacity in rheumatoid arthritis: distribution of symptomatic joints in 1,000 RA patients. Acta Orthop. 2008;79(2):257-61.

10. van der Leeden M, Steultjens MP, Ursum J, Dahmen R, Roorda LD, Schaardenburg DV, et al. Prevalence and course of forefoot impairments and walking disability in the first eight years of rheumatoid arthritis. Arthritis Rheum. 2008;59(11):1596-602.

11. Wickman AM, Pinzur MS, Kadanoff R, Juknelis D. Health-related quality of life for patients with rheumatoid arthritis foot involvement. Foot Ankle Int. 2004;25(1):19-26.

12. Thomas MJ, Peat G, Rathod T, Marshall M, Moore A, Menz HB, et al. The epidemiology of symptomatic midfoot osteoarthritis in community-dwelling older adults: cross-sectional findings from the Clinical Assessment Study of the Foot. Arthritis Res Ther. 2015;17(1):178.

13. Paget LDA, Tol JL, Kerkhoffs G, Reurink G. Health-related quality of life in ankle osteoarthritis: a case-control study. Cartilage. 2021;13(1_suppl):1438s-44s.

14. Roddy E, Muller S, Rome K, Chandratre P, Hider SL, Richardson J, et al. Foot problems in people with gout in primary care: baseline findings from a prospective cohort study. J Foot Ankle Res. 2015;8:31.

15. Reinoso-Cobo A, Gijon-Nogueron G, Caliz-Caliz R, Ferrer-Gonzalez MA, Vallejo-Velazquez MT, Miguel Morales-Asencio J, et al. Foot health and quality of life in patients with rheumatoid arthritis: a cross-sectional study. BMJ Open. 2020;10(5):e036903.

16. Hansen L, Krogh TP, Ellingsen T, Bolvig L, Fredberg U. Long-term prognosis of plantar fasciitis: a 5- to 15-year follow-up study of 174 patients with ultrasound examination. Orthop J Sports Med. 2018;6(3):2325967118757983.

17. Scott DL, Wolfe F, Huizinga TWJ. Rheumatoid arthritis. Lancet. 2010;376(9746):1094-108.

18. van der Leeden M, Steultjens MPM, van Schaardenburg D, Dekker J. Forefoot disease activity in rheumatoid arthritis patients in remission: results of a cohort study. Arthritis Res Ther. 2010;12(1):R3.

19. Hooper L, Bowen CJ, Gates L, Culliford DJ, Ball C, Edwards CJ, et al. Prognostic indicators of foot-related disability in patients with rheumatoid arthritis: results of a prospective three-year study. Arthritis Care Res (Hoboken). 2012;64(8):1116-24.

20. Turner DE, Helliwell PS, Siegel KL, Woodburn J. Biomechanics of the foot in rheumatoid arthritis: Identifying abnormal function and the factors associated with localised disease ‘impact’. Clin Biomech. 2008;23(1):93-100.

21. Hendry G. Juvenile idiopathic arthritis. In: Helliwell PS, Backhouse MR, Siddle HJ, editors. The Foot and Ankle in Rheumatology: Oxford University Press; 2019.

22. Turner D, Helliwell P. Spondyloarthropathies. In: Helliwell PS, Backhouse MR, Siddle HJ, editors. The Foot and Ankle in Rheumatology: Oxford University Press; 2019.

23. Turner DE, Helliwell PS, Emery P, Woodburn J. The impact of rheumatoid arthritis on foot function in the early stages of disease: a clinical case series. BMC Musculoskelet Disord. 2006;7:102.

24. Gassara Z, Feki A, Hakim Z, Ben Djmeaa S, Abid C, Kallel MH, et al. Foot involvement in psoriatic arthritis: Prevalence, clinical and radiological features. Foot Ankle Surg. 2024;30(6):465-70.

25. Agarwal V, Singh R, Wiclaf, Chauhan S, Tahlan A, Ahuja CK, et al. A clinical, electrophysiological, and pathological study of neuropathy in rheumatoid arthritis. Clin Rheumatol. 2008;27(7):841-4.

26. Kaeley N, Ahmad S, Pathania M, Kakkar R. Prevalence and patterns of peripheral neuropathy in patients of rheumatoid arthritis. J Family Med Prim Care. 2019;8(1):22-6.

27. Alkaabi JK, Ho M, Levison R, Pullar T, Belch JJ. Rheumatoid arthritis and macrovascular disease. Rheumatology (Oxford). 2003;42(2):292-7.

28. Gotzsche PC, Johansen HK. Short-term low-dose corticosteroids vs placebo and nonsteroidal antiinflammatory drugs in rheumatoid arthritis. Cochrane Database Syst Rev. 2004;2005(3):Cd000189.

29. Firth J, Waxman R, Law G, Nelson EA, Helliwell P, Siddle H, et al. The predictors of foot ulceration in patients with rheumatoid arthritis. Clin Rheumatol. 2014;33(5):615-21.

30. Siddle HJ, Firth J, Waxman R, Nelson EA, Helliwell PS. A case series to describe the clinical characteristics of foot ulceration in patients with rheumatoid arthritis. Clin Rheumatol. 2012;31(3):541-5.

31. Firth J, Hale C, Helliwell P, Hill J, Nelson EA. The prevalence of foot ulceration in patients with rheumatoid arthritis. Arthritis Rheum. 2008;59(2):200-5.

32. Hunter DJ, Felson DT. Osteoarthritis. BMJ. 2006;332(7542):639-42.

33. Roddy E, Thomas MJ, Marshall M, Rathod T, Myers H, Menz HB, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: cross-sectional findings from the clinical assessment study of the foot. Ann Rheum Dis. 2015;74(1):156-63.

34. Murray C, Marshall M, Rathod T, Bowen CJ, Menz HB, Roddy E. Population prevalence and distribution of ankle pain and symptomatic radiographic ankle osteoarthritis in community dwelling older adults: A systematic review and cross-sectional study. PLoS One. 2018;13(4):e0193662.

35. Roddy E, Doherty M. Epidemiology of gout. Arthritis Res Ther. 2010;12(6):223.

36. Singh JA, Gaffo A. Gout epidemiology and comorbidities. Semin Arthritis Rheum. 2020;50(3s):S11-s6.

37. Rome K, Roddy E. Crystal arthritis. In: Helliwell PS, Backhouse MR, Siddle HJ, editors. The Foot and Ankle in Rheumatology: Oxford University Press; 2019. p. 0.

38. Voulgari PV, Alamanos Y, Papazisi D, Christou K, Papanikolaou C, Drosos AA. Prevalence of Raynaud's phenomenon in a healthy Greek population. Ann Rheum Dis. 2000;59(3):206-10.

39. Haque A, Hughes M. Raynaud's phenomenon. Clin Med (Lond). 2020;20(6):580-7.

40. Tian J, Kang S, Zhang D, Huang Y, Zhao M, Gui X, et al. Global, regional, and national incidence and prevalence of systemic sclerosis. Clin Immunol. 2023;248:109267.

41. Bairkdar M, Rossides M, Westerlind H, Hesselstrand R, Arkema EV, Holmqvist M. Incidence and prevalence of systemic sclerosis globally: a comprehensive systematic review and meta-analysis. Rheumatology. 2021;60(7):3121-33.

42. Allanore Y, Simms R, Distler O, Trojanowska M, Pope J, Denton CP, et al. Systemic sclerosis. Nat Rev Dis Primers. 2015;1(1):15002.

43. Walker UA, Tyndall A, Czirják L, Denton C, Farge-Bancel D, Kowal-Bielecka O, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. Ann Rheum Dis. 2007;66(6):754-63.

44. Herrick AL. Raynaud's phenomenon. J Scleroderma Relat Disord. 2019;4(2):89-101.

45. Sari-Kouzel H, Hutchinson CE, Middleton A, Webb F, Moore T, Griffin K, et al. Foot problems in patients with systemic sclerosis. Rheumatology. 2001;40(4):410-3.

46. La Montagna G, Baruffo A, Tirri R, Buono G, Valentini G. Foot involvement in systemic sclerosis: a longitudinal study of 100 patients. Semin Arthritis Rheum. 2002;31(4):248-55.

47. La Montagna G, Sodano A, Capurro V, Malesci D, Valentini G. The arthropathy of systemic sclerosis: a 12 month prospective clinical and imaging study. Skeletal Radiol. 2005;34(1):35-41.

48. Tian J, Zhang D, Yao X, Huang Y, Lu Q. Global epidemiology of systemic lupus erythematosus: a comprehensive systematic analysis and modelling study. Ann Rheum Dis. 2023;82(3):351-6.

49. Otter SJ, Kumar S, Gow P, Dalbeth N, Corkill M, Rohan M, et al. Patterns of foot complaints in systemic lupus erythematosus: a cross sectional survey. J Foot Ankle Res. 2016;9:10.

50. Koc TA, Bise CG, Neville C, Carreira D, Martin RL, McDonough CM. Heel Pain – Plantar Fasciitis: Revision 2023. J Orthop Sports Phys Ther. 2023;53(12):CPG1-CPG39.

51. Thomas MJ, Whittle R, Menz HB, Rathod-Mistry T, Marshall M, Roddy E. Plantar heel pain in middle-aged and older adults: population prevalence, associations with health status and lifestyle factors, and frequency of healthcare use. BMC Musculoskelet Disord. 2019;20(1):337.

52. Sullivan J, Pappas E, Burns J. Role of mechanical factors in the clinical presentation of plantar heel pain: Implications for management. Foot (Edinb). 2020;42:101636.

53. de Jonge S, van den Berg C, de Vos RJ, van der Heide HJL, Weir A, Verhaar JAN, et al. Incidence of midportion Achilles tendinopathy in the general population. Br J Sports Med. 2011;45(13):1026-8.

54. Landorf KB, Keenan A-M, Kearney R. Soft-tissue disorders. In: Helliwell PS, Backhouse MR, Siddle HJ, editors. The Foot and Ankle in Rheumatology: Oxford University Press; 2019. p. 0.

55. Nix S, Smith M, Vicenzino B. Prevalence of hallux valgus in the general population: a systematic review and meta-analysis. J Foot Ankle Res. 2010;3:21.

56. Lisacek-Kiosoglous A, Ristanis S, Georgiou A. The trend of hallux valgus diagnosis and treatment: NHS England, population level data (1999-2019). J Orthop Bone Res. 2021;2(1):1-14.

57. Tenten-Diepenmaat M, Dekker J, Heymans MW, Roorda LD, Vliet Vlieland TPM, van der Leeden M. Systematic review on the comparative effectiveness of foot orthoses in patients with rheumatoid arthritis. J Foot Ankle Res. 2019;12:32.

58. Conceição CSd, Gomes Neto M, Mendes SMD, Sá KN, Baptista AF. Systematic review and meta-analysis of effects of foot orthoses on pain and disability in rheumatoid arthritis patients. Disability and Rehabilitation. 2015;37(14):1209-13.

59. National Institute for Health and Care Excellence (NICE). Rheumatoid arthritis in adults: management. NICE guideline NG100. Published: 11 July 2018. Last updated: 12 October 2020. https://www.nice.org.uk/guidance/ng100 [

60. Chapman LS, Backhouse M, Corp N, van der Windt D, Bearne L, Cherry L, et al. The British Society for Rheumatology guideline for the management of foot health in inflammatory arthritis. Rheumatology. 2025;64(5):2355-68.

61. Joyce C, Rajak R. A clinical audit into the adherence of foot health management standards of rheumatoid arthritis compared with the foot health management standards of diabetes mellitus in North-East London. Rheumatol Adv Pract. 2021;5(1):rkab006.

62. McCulloch L, Borthwick A, Redmond A, Edwards K, Pinedo-Villanueva R, Prieto-Alhambra D, et al. UK podiatrists' experiences of podiatry services for people living with arthritis: a qualitative investigation. J Foot Ankle Res. 2018;11(1):27.

63. Beaton D ML, Grosskleg S, Shea B, Tugwell B (editors). The OMERACT Handbook Version 2.1. OMERACT; 2021. https://omeract.org/wp-content/uploads/2025/11/Handbook-June-2021-1.pdf Accessed on 1st December 2025.

64. Riel H, Plinsinga ML, Delahunt E, Jensen MB, Landorf KB, van Middelkoop M, et al. Large variation in participant eligibility criteria used in plantar heel pain research studies - a systematic review. J Foot Ankle Res. 2022;15(1):69.

65. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials. 2017;18(3):280.

66. Webbe J, Sinha I, Gale C. Core outcome sets. Arch Dis Child Educ Pract Ed. 2018;103(3):163-6.

67. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials. 2012;13:132.

68. Kirkham JJ, Williamson P. Core outcome sets in medical research. BMJ Med. 2022;1(1):e000284.

69. Barton JL, Imboden J, Graf J, Glidden D, Yelin EH, Schillinger D. Patient-physician discordance in assessments of global disease severity in rheumatoid arthritis. Arthritis Care Res (Hoboken). 2010;62(6):857-64.

70. Smolen JS, Strand V, Koenig AS, Szumski A, Kotak S, Jones TV. Discordance between patient and physician assessments of global disease activity in rheumatoid arthritis and association with work productivity. Arthritis Res Ther. 2016;18(1):114.

71. Bartlett SJ, Hewlett S, Bingham CO, 3rd, Woodworth TG, Alten R, Pohl C, et al. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. Ann Rheum Dis. 2012;71(11):1855-60.

72. Mackie SL, Twohig H, Neill LM, Harrison E, Shea B, Black RJ, et al. The OMERACT core domain set for outcome measures for clinical trials in polymyalgia rheumatica. J Rheumatol. 2017;44(10):1515-21.

73. Chapman LS, Jones J, Redmond AC, Flurey CA, Richards P, Hofstetter C, et al. Developing a core outcome set for foot and ankle disorders in rheumatic and musculoskeletal diseases: A scoping review and report from the OMERACT 2022 foot and ankle special interest group session. Semin Arthritis Rheum. 2023;61:152210.

74. Ortega-Avila AB, Ramos-Petersen L, Cervera-Garvi P, Nester CJ, Morales-Asencio JM, Gijon-Nogueron G. Systematic review of the psychometric properties of patient-reported outcome measures for rheumatoid arthritis in the foot and ankle. Clin Rehabil. 2019;33(11):1788-99.

75. Jia Y, Huang H, Gagnier JJ. A systematic review of measurement properties of patient-reported outcome measures for use in patients with foot or ankle diseases. Qual Life Res. 2017;26(8):1969-2010.

76. Martin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. J Orthop Sports Phys Ther. 2007;37(2):72-84.

77. van der Leeden M, Steultjens MP, Terwee CB, Rosenbaum D, Turner D, Woodburn J, et al. A systematic review of instruments measuring foot function, foot pain, and foot-related disability in patients with rheumatoid arthritis. Arthritis Rheum. 2008;59(9):1257-69.

78. Naal FD, Impellizzeri FM, Rippstein PF. Which are the most frequently used outcome instruments in studies on total ankle arthroplasty? Clin Orthop Relat Res. 2010;468(3):815-26.

79. Schrier JC, Palmen LN, Verheyen CC, Jansen J, Koëter S. Patient-reported outcome measures in hallux valgus surgery. A review of literature. Foot Ankle Surg. 2015;21(1):11-5.

80. Murphy M, Travers M, Gibson W, Chivers P, Debenham J, Docking S, et al. Rate of improvement of pain and function in mid-portion Achilles tendinopathy with loading protocols: a systematic review and longitudinal meta-analysis. Sports Med. 2018;48(8):1875-91.

81. Prudêncio DA, Serafim TT, Marinho Mateus Lopes APSR, Maffulli N, Okubo R. Questionnaires and scales for assessment of ankle function: a systematic review of instruments translated and validated for Brazilian Portuguese. Disabil Rehabil. 2021;43(3):309-16.

82. Saldanha IJ, Hughes KL, Dodd S, Lasserson T, Kirkham JJ, Wu Y, et al. Study found increasing use of core outcome sets in Cochrane systematic reviews and identified facilitators and barriers. J Clin Epidemiol. 2024;169:111277.

83. Plys E, Moss KO, Jacklin K, Yazzie M, Tambor E, Luers E, et al. Discontinuing the term "stakeholder" from the NIA IMPACT Collaboratory Engaging Partners Team: an example of the process of language change in an organization. Am J Geriatr Psychiatry. 2025;33(2):192-9.

84. Akl EA, Khabsa J, Petkovic J, Magwood O, Lytvyn L, Motilall A, et al. "Interest-holders": A new term to replace "stakeholders" in the context of health research and policy. Cochrane Evid Synth Methods. 2024;2(11):e70007.

85. Håkansson Eklund J, Holmström IK, Kumlin T, Kaminsky E, Skoglund K, Höglander J, et al. "Same same or different?" A review of reviews of person-centered and patient-centered care. Patient Educ Couns. 2019;102(1):3-11.

86. National Institute for Health and Care Research (NIHR). Public invovlement in research. 2023. https://www.nihr.ac.uk/get-involved/public-involvement Accessed 5 April 2023.

87. Chapman LS, Redmond AC, Flurey CA, Richards P, Smith TO, Arnold JB, et al. Developing an Outcome Measures in Rheumatology (OMERACT) Core set of Outcome Measures for FOot and ankle disorders in RheumaTic and musculoskeletal diseases (COMFORT): core domain set study protocol. Trials. 2023;24(1):65.

88. Core Outcome Measures in Effectiveness Trials (COMET Initiative). www.comet-initiative.org Accessed on 14th August 2024.

89. Core Outcome Measures in Effectiveness Trials (COMET Initiative). https://www.comet-initiative.org/Studies/Details/2081 Accessed on 1st November 2025.

90. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on

measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539-49.

91. Leshem YA, Simpson EL, Apfelbacher C, Spuls PI, Thomas KS, Schmitt J, et al. The Harmonising Outcome Measures for Eczema (HOME) implementation roadmap. Br J Dermatol. 2023;189(6):710-8.

92. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain. 2003;106(3):337-45.

93. Molloy EJ, Gale C, Marsh M, Bearer CF, Devane D, Modi N. Developing core outcome set for women's, newborn, and child health: the CROWN Initiative. Pediatr Res. 2018;84(3):316-7.

94. Dodd S, Gorst SL, Young A, Lucas SW, Williamson PR. Patient participation impacts outcome domain selection in core outcome sets for research: an updated systematic review. J Clin Epidemiol. 2023;158:127-33.

95. Young B, Bagley H. Including patients in core outcome set development: issues to consider based on three workshops with around 100 international delegates. Res Involv Engagem. 2016;2(1):25.

96. Cheung PP, de Wit M, Bingham CO, 3rd, Kirwan JR, Leong A, March LM, et al. Recommendations for the Involvement of Patient Research Partners (PRP) in OMERACT Working Groups. A Report from the OMERACT 2014 Working Group on PRP. J Rheumatol. 2016;43(1):187-93.

97. de Wit M, Kirwan JR, Tugwell P, Beaton D, Boers M, Brooks P, et al. Successful stepwise development of patient research partnership: 14 years' experience of actions and consequences in Outcome Measures in Rheumatology (OMERACT). Patient. 2017;10(2):141-52.

98. de Wit MP, Abma TA, Koelewijn-van Loon MS, Collins S, Kirwan J. What has been the effect on trial outcome assessments of a decade of patient participation in OMERACT? J Rheumatol. 2014;41(1):177-84.

99. de Wit M, Abma T, Koelewijn-van Loon M, Collins S, Kirwan J. Involving patient research partners has a significant impact on outcomes research: a responsive evaluation of the international OMERACT conferences. BMJ Open. 2013;3(5).

100. Hofstetter C, Grosskleg S, Hamilton CB, Hoens AM, Shea B, Tugwell P, et al. Patient research partner engagement in OMERACT: Enhancing engagement through

the implementation of patient engagement in research tools. Semin Arthritis Rheum. 2024;67:152447.

101. Core Outcome Measures in Effectiveness Trials (COMET Initiative). COMET PPI Toolkit. https://comet-ppi-toolkit.liverpool.ac.uk/toolkit/ Accessed on 14th July 2025.

102. https://omeract.org/wp-content/uploads/2025/11/OMERACT-PRP-Toolkit-v2-Sept-2025.pdf.

103. Karumbi J, Gorst SL, Gathara D, Gargon E, Young B, Williamson PR. Inclusion of participants from low-income and middle-income countries in core outcome sets development: a systematic review. BMJ Open. 2021;11(10):e049981.

104. Lee A, Davies A, Young AE. Systematic review of international Delphi surveys for core outcome set development: representation of international patients. BMJ Open. 2020;10(11):e040223.

105. Karumbi J, Gorst S, Gathara D, Young B, Williamson P. Awareness and experiences on core outcome set development and use amongst stakeholders from low- and middle- income countries: An online survey. PLOS Glob Public Health. 2023;3(12):e0002574.

106. Alkhaffaf B, Metryka A, Blazeby JM, Glenny AM, Adeyeye A, Costa PM, et al. Core outcome set for surgical trials in gastric cancer (GASTROS study): international patient and healthcare professional consensus. Br J Surg. 2021;108(10):1216-24.

107. Davies PA, Davies AK, Kirkham JJ, Young AE. Secondary analysis of data from a core outcome set for burns demonstrated the need for involvement of lower income countries. J Clin Epidemiol. 2022;144:56-71.

108. Vidyasagaran AL, Ayesha R, Boehnke JR, Kirkham J, Rose L, Hurst JR, et al. Core outcome sets for trials of interventions to prevent and to treat multimorbidity in adults in low and middle-income countries: the COSMOS study. BMJ Global Health. 2024;9(8):e015120.

109. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. PLoS Med. 2017;14(11):e1002447.

110. Kirkham JJ, Gorst S, Altman DG, Blazeby JM, Clarke M, Devane D, et al. Core Outcome Set-STAndards for Reporting: The COS-STAR Statement. PLoS Med. 2016;13(10):e1002148.

111. Lasserson TJ TJ, Higgins JP. . Chapter 1: Starting a review [last updated August 2021]. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al,

editor(s). Cochrane Handbook for Systematic Reviews of Interventions version 6.5. Cochrane, 2024.

112. International prospective register of systematic reviews (PROSPERO). 2025. https://www.crd.york.ac.uk/prospero/ Accessed on 14th October 2025.

113. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32.

114. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and explanation. Ann Intern Med. 2018;169(7):467-73.

115. Page MJ, Huang H, Verhagen AP, Gagnier JJ, Buchbinder R. Outcome reporting in randomized trials for shoulder disorders: literature review to inform the development of a core outcome set. Arthritis Care Res (Hoboken). 2018;70(2):252-9.

116. Hodgson R, Bushe C, Hunter R. Measurement of long-term outcomes in observational and randomised controlled trials. Br J Psychiatry. 2007;191(S50):s78-s84.

117. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol. 2007;34(5):1174-7.

118. Keeley T, Williamson P, Callery P, Jones LL, Mathers J, Jones J, et al. The use of qualitative methods to inform Delphi surveys in core outcome set development. Trials. 2016;17(1):230.

119. Creswell JW, Tashakkori A. Editorial: Differing perspectives on mixed eethods research. J Mix Methods Res. 2007;1(4):303-8.

120. Guba EG, Lincoln YS. Competing paradigms in qualitative research. In: Denzin NK, Lincoln YS (eds.) Handbook of qualitative research. Thousand Oaks, California: Sage Publications, Inc; 1994. p. 105-117.

121. Clarke V, Braun, V. Successful qualitative research. Thousand Oaks, California: Sage Publications, Inc; 2013.

122. Kaushik V, Walsh CA. Pragmatism as a research paradigm and its implications for social work research. Social Sciences. 2019;8(9):255.

123. Downe S, Finlayson KW, Lawrie TA, Lewin SA, Glenton C, Rosenbaum S, et al. Qualitative Evidence Synthesis (QES) for Guidelines: Paper 1 – Using qualitative evidence synthesis to inform guideline scope and develop qualitative findings statements. Health Res Policy Syst. 2019;17(1):76.

124. Flemming K, Booth A, Garside R, Tunçalp Ö, Noyes J. Qualitative evidence synthesis for complex interventions and guideline development: clarification of the purpose, designs and relevant methods. BMJ Global Health. 2019;4(Suppl 1):e000882.

125. Dixon-Woods M, Shaw RL, Agarwal S, Smith JA. The problem of appraising qualitative research. Qual Saf Health Care. 2004;13(3):223-5.

126. Korstjens I, Moser A. Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. Eur J Gen Pract. 2018;24(1):120-4.

127. Duden GS. Challenges to qualitative evidence synthesis – Aiming for diversity and abstracting without losing meaning. Methods in Psychology. 2021;5:100070.

128. Barnett-Page E, Thomas J. Methods for the synthesis of qualitative research: a critical review. BMC Med Res Methodol. 2009;9(1):59.

129. Eaves YD. A synthesis technique for grounded theory data analysis. J Adv Nurs. 2001;35(5):654-63.

130. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc. 2015;13(3):179-87.

131. Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: meta-narrative reviews. BMC Med. 2013;11(1):20.

132. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O, Peacock R. Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review. Soc Sci Med. 2005;61(2):417-30.

133. Lucas PJ, Baird J, Arai L, Law C, Roberts HM. Worked examples of alternative methods for the synthesis of qualitative and quantitative research in systematic reviews. BMC Med Res Methodol. 2007;7(1):4.

134. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol. 2008;8(1):45.

135. Oliver SR, Rees RW, Clarke-Jones L, Milne R, Oakley AR, Gabbay J, et al. A multidimensional conceptual framework for analysing public involvement in health services research. Health Expect. 2008;11(1):72-84.

136. Noyes J BA, Cargo M, Flemming K, Harden A, Harris J, Garside R, Hannes K, Pantoja T, Thomas J. Chapter 21: Qualitative evidence. In: Higgins JP TJ, Chandler J, Cumpston M, Li T, Page MJ, et al., editor. Cochrane Handbook for Systematic Reviews of Interventions version 65. Cochrane, 2019.

137. Katrak P, Bialocerkowski AE, Massy-Westropp N, Kumar VSS, Grimmer KA. A systematic review of the content of critical appraisal tools. BMC Med Res Methodol. 2004;4(1):22.

138. Munthe-Kaas HM, Glenton C, Booth A, Noyes J, Lewin S. Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: first stage in the development of the CAMELOT tool. BMC Med Res Methodol. 2019;19(1):113.

139. Dalton J, Booth A, Noyes J, Sowden AJ. Potential value of systematic reviews of qualitative evidence in informing user-centered health and social care: findings from a descriptive overview. J Clin Epidemiol. 2017;88:37-46.

140. Granholm A, Alhazzani W, Møller MH. Use of the GRADE approach in systematic reviews and guidelines. Br J Anaesth. 2019;123(5):554-9.

141. Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implement Sci. 2018;13(1):2.

142. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. Reviewing evidence. [Available from: https://www.nice.org.uk/consultations/1950/3/reviewing-evidence.

143. Munthe-Kaas H, Bohren MA, Glenton C, Lewin S, Noyes J, Tunçalp Ö, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 3: how to assess methodological limitations. Implement Sci. 2018;13(1):9.

144. Munn Z, Porritt K, Lockwood C, Aromataris E, Pearson A. Establishing confidence in the output of qualitative research synthesis: the ConQual approach. BMC Med Res Methodol. 2014;14:108.

145. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol. 2012;12(1):181.

146. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

147. de Jong Y, van der Willik EM, Milders J, Voorend CGN, Morton RL, Dekker FW, et al. A meta-review demonstrates improved reporting quality of qualitative reviews following the publication of COREQ- and ENTREQ-checklists, regardless of modest uptake. BMC Med Res Methodol. 2021;21(1):184.

148. Gorst SL, Young B, Williamson PR, Wilding JPH, Harman NL. Incorporating patients' perspectives into the initial stages of core outcome set development: a rapid review of qualitative studies of type 2 diabetes. BMJ Open Diabetes Res Care. 2019;7(1):e000615.

149. Ramiro S, Page MJ, Whittle SL, Huang H, Verhagen AP, Beaton DE, et al. The OMERACT core domain set for clinical trials of shoulder disorders. J Rheumatol. 2019;46(8):969-75.

150. Smithson J. Using and analysing focus groups: Limitations and possibilities. Int J Soc Res Methodol. 2000;3(2):103-19.

151. Guest G, Namey E, Taylor J, Eley N, McKenna K. Comparing focus groups and individual interviews: findings from a randomized study. Int J Soc Res Methodol. 2017;20(6):693-708.

152. Sim J, Waterfield J. Focus group methodology: some ethical challenges. Qual Quant. 2019;53(6):3003-22.

153. Abrams KM, Wang Z, Song YJ, Galindo-Gonzalez S. Data richness trade-offs between face-to-face, online audiovisual, and online text-only focus groups. Soc Sci Comput Review. 2015;33(1):80-96.

154. Sturges JE, Hanrahan KJ. Comparing telephone and face-to-face qualitative interviewing: a research note. Qualitative Research. 2004;4(1):107-18.

155. Lambert SD, Loiselle CG. Combining individual interviews and focus groups to enhance data richness. J Adv Nurs. 2008;62(2):228-37.

156. Gargon E, Williamson PR, Young B. Improving core outcome set development: qualitative interviews with developers provided pointers to inform guidance. J Clin Epidemiol. 2017;86:140-52.

157. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res. 2016;26(13):1753-60.

158. Creswell JW, Creswell, JD. Research design: qualitative, quantitative, and mixed methods approaches. 5th ed. Los Angeles: SAGE; 2018.

159. Pauling JD, Domsic RT, Saketkoo LA, Almeida C, Withey J, Jay H, et al. Multinational qualitative research study exploring the patient experience of Raynaud's phenomenon in systemic sclerosis. Arthritis Care Res (Hoboken). 2018;70(9):1373-84.

160. Nielsen SM, Boers M, de Wit M, Shea B, van der Windt DA, Reeves BC, et al. OMERACT consensus-based operational definition of contextual factors in

rheumatology clinical trials: A mixed methods study. Semin Arthritis Rheum. 2021;51(3):601-6.

161. Mackie SL, Hughes R, Walsh M, Day J, Newton M, Pease C, et al. "An impediment to living life": why and how should we measure stiffness in polymyalgia rheumatica? PLoS One. 2015;10(5):e0126758.

162. Flurey CA, Jones B, Gazel U, Uzoka C, Rosser K, Khoo T, et al. "It means almost forgetting that you've got a disease": An OMERACT study to define independence in the context of rheumatoid arthritis remission from the patient perspective. Semin Arthritis Rheum. 2024;68:152526.

163. Salmasi S, Kelly A, Bartlett SJ, de Wit M, March L, Tong A, et al. Researchers' perspectives on methodological challenges and outcomes selection in interventional studies targeting medication adherence in rheumatic diseases: an OMERACT-adherence study. BMC Rheumatol. 2021;5(1):26.

164. Creswell JW PC. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, California: SAGE Publications, Inc.; 2018.

165. Starks H, Trinidad SB. Choose your method: a comparison of phenomenology, discourse analysis, and grounded theory. Qual Health Res. 2007;17(10):1372-80.

166. Santiago-Delefosse M, Gavin A, Bruchez C, Roux P, Stephen SL. Quality of qualitative research in the health sciences: Analysis of the common criteria present in 58 assessment guidelines by expert users. Soc Sci Med. 2016;148:142-51.

167. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med. 2014;89(9):1245-51.

168. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. BMC Med Res Methodol. 2007;19(6):349-57.

169. Braun V, Clarke V. How do you solve a problem like COREQ? A critique of Tong et al.’s (2007) Consolidated Criteria for Reporting Qualitative Research. Methods Psychol. 2024;11:100155.

170. Braun V, Clarke V. Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: A review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). Palliat Med. 2024;38(6):608-16.

171. Lincoln YS, Guba EG. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. New Directions for Program Evaluation. 1986;1986(30):73-84.

172. Johnson JL, Adkins D, Chauvin S. A review of the quality indicators of rigor in qualitative research. Am J Pharm Educ. 2020;84(1):7120.

173. Harvey N, Holmes CA. Nominal group technique: an effective method for obtaining group consensus. Int J Nurs Pract. 2012;18(2):188-94.

174. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. Int J Clin Pharm. 2016;38(3):655-62.

175. Tillett W, Dures E, Hewlett S, Helliwell PS, FitzGerald O, Brooke M, et al. A multicenter nominal group study to rank outcomes important to patients, and their representation in existing composite outcome measures for psoriatic arthritis. J Rheumatol. 2017;44(10):1445-52.

176. Kelly A, Tymms K, de Wit M, Bartlett SJ, Cross M, Dawson T, et al. Patient and Caregiver Priorities for Medication Adherence in Gout, Osteoporosis, and Rheumatoid Arthritis: Nominal Group Technique. Arthritis Care Res (Hoboken). 2020;72(10):1410-9.

177. Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. Semin Arthritis Rheum. 2011;41(2):95-105.

178. Waggoner J, Carline JD, Durning SJ. Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. Acad Med. 2016;91(5):663-8.

179. Grosskleg S, Beaton D, Conaghan P, Hofstetter C, Maxwell L, Shea B, et al. OMERACT 2020: A virtual (R)evolution. Semin Arthritis Rheum. 2021;51(3):588-92.

180. OMERACT. Downloadable forms: Core domain set development. 5.1 Core domain set online module content. 2025. https://omeract.org/resources/downloadable-forms/ Accessed on 1st October 2025.

181. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualitative Research in Sport, Exercise and Health. 2021;13(2):201-16.

182. Firth J, Nelson EA, Briggs M, Gorecki C. A qualitative study to explore the impact of foot ulceration on health-related quality of life in patients with rheumatoid arthritis. Int J Nurs Stud. 2011;48(11):1401-8.

183. Bloor M, Wood F. Keywords in qualitative methods: A vocabulary of research concepts. 2006.

184. Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus. Practical assessment, research, and evaluation. 2007;12(1).

185. Biggane AM, Brading L, Ravaud P, Young B, Williamson PR. Survey indicated that core outcome set development is increasingly including patients, being conducted internationally and using Delphi surveys. Trials. 2018;19(1):113.

186. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: How to decide its appropriateness. World J Methodol. 2021;11(4):116-29.

187. Fish R, MacLennan S, Alkhaffaf B, Williamson PR. “Vicarious thinking” was a key driver of score change in Delphi surveys for COS development and is facilitated by feedback of results. J Clin Epidemiol. 2020;128:118-29.

188. Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. Int J Nurs Stud. 2001;38(2):195-200.

189. Hutchings A, Raine R, Sanderson C, Black N. A comparison of formal consensus methods used for developing clinical guidelines. J Health Serv Res Policy. 2006;11(4):218-24.

190. Brookes ST, Macefield RC, Williamson PR, McNair AG, Potter S, Blencowe NS, et al. Three nested randomized controlled trials of peer-only or multiple stakeholder group feedback within Delphi surveys during core outcome and information set development. Trials. 2016;17(1):409.

191. MacLennan S, Kirkham J, Lam TBL, Williamson PR. A randomized trial comparing three Delphi feedback strategies found no evidence of a difference in a setting with high initial agreement. J Clin Epidemiol. 2018;93:1-8.

192. Fish R, Brookes S, Sanders C, Williamson P. Comprehension of feedback by patients within a Delphi survey. Preprint. 2018.

193. Hall DA, Smith H, Heffernan E, Fackrell K. Recruiting and retaining participants in e-Delphi surveys for core outcome set development: Evaluating the COMiT'ID study. PLoS One. 2018;13(7):e0201378.

194. Grant S, Booth M, Khodyakov D. Lack of preregistered analysis plans allows unacceptable data mining for and selective reporting of consensus in Delphi studies. J Clin Epidemiol. 2018;99:96-105.

195. Howell SJ. The need for rigour in consensus statements and guidelines. Br J Anaesth. 2024 Dec;133(6):1123-1125.

196. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: Recommendations based on a methodological systematic review. Palliat Med. 2017;31(8):684-706.

197. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol. 2014;67(4):401-9.

198. Seylanova N, Chernyavskaya A, Degtyareva N, Mursalova A, Ajam A, Xiao L, et al. Core outcome measurement set for research and clinical practice in post-COVID-19 condition (long COVID) in children and young people: an international Delphi consensus study "PC-COS Children". Eur Respir J. 2024;63(3).

199. Schmitt J, Langan S, Stamm T, Williams HC. Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. J Invest Dermatol. 2011;131(3):623-30.

200. Bennett C, Vakil N, Bergman J, Harrison R, Odze R, Vieth M, et al. Consensus statements for management of Barrett's dysplasia and early-stage esophageal adenocarcinoma, based on a Delphi process. Gastroenterology. 2012;143(2):336-46.

201. Harman NL, Bruce IA, Kirkham JJ, Tierney S, Callery P, O'Brien K, et al. The importance of integration of stakeholder views in core outcome set development: otitis media with effusion in children with cleft palate. PLoS One. 2015;10(6):e0129514.

202. Blazeby JM, Macefield R, Blencowe NS, Jacobs M, McNair AG, Sprangers M, et al. Core information set for oesophageal cancer surgery. Br J Surg. 2015;102(8):936-43.

203. Barrington H, Young B, Williamson PR. Patient participation in Delphi surveys to develop core outcome sets: systematic review. BMJ Open. 2021;11(9):e051066.

204. Crew R, Williamson PR. COMET Initiative. DelphiManager brochure. https://www.comet-initiative.org/delphimanager/docs/DelphiManagerBrochureV4.0.pdf Accessed 2nd October 2024.

205. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.

206. De Meyer D, Kottner J, Beele H, Schmitt J, Lange T, Van Hecke A, et al. Delphi procedure in core outcome set development: rating scale and consensus criteria determined outcome selection. J Clin Epidemiol. 2019;111:23-31.

207. Remus A, Smith V, Wuytack F. Methodology in core outcome set (COS) development: the impact of patient interviews and using a 5-point versus a 9-point Delphi rating scale on core outcome selection in a COS development study. BMC Med Res Methodol. 2021;21(1):10.

208. Alwin DF, Krosnick JA. The reliability of survey attitude measurement: The influence of question and respondent attributes. Sociological Methods and Research. 1991;20(1):139-81.

209. Brookes ST, Chalmers KA, Avery KNL, Coulman K, Blazeby JM. Impact of question order on prioritisation of outcomes in the development of a core outcome set: a randomised controlled trial. Trials. 2018;19(1):66.

210. Jorm A. Steps in Carrying out a Delphi Study. In: Jorm A, editor. Using the Delphi Method to Establish Expert Consensus: A Practical Guide. Singapore: Springer Nature Singapore; 2025. p. 41-93.

211. Veugelers R, Gaakeer MI, Patka P, Huijsman R. Improving design choices in Delphi studies in medicine: the case of an exemplary physician multi-round panel study with 100% response. BMC Med Res Methodol. 2020;20(1):156.

212. Quirke FA, Battin MR, Bernard C, Biesty L, Bloomfield FH, Daly M, et al. Multi-round versus real-time Delphi survey approach for achieving consensus in the COHESION core outcome set: a randomised trial. Trials. 2023;24(1):461.

213. Gargon E, Crew R, Burnside G, Williamson PR. Higher number of items associated with significantly lower response rates in COS Delphi surveys. J Clin Epidemiol. 2019;108:110-20.

214. Kottner J, Beaton D, Clarke M, Dodd S, Kirkham J, Lange T, et al. Core outcome set developers should consider and specify the level of granularity of outcome domains. J Clin Epidemiol. 2024;169:111307.

215. D'Agostino MA, Beaton DE, Maxwell LJ, Cembalo SM, Hoens AM, Hofstetter C, Zabalan C, Bird P, Christensen R, de Wit M, Doria AS, Maksymowych WP, Oo WM, Østergaard M, Serban T, Sloan VS, Terslev L, van Rossum MA, Conaghan PG, Boers M. Improving domain definition and outcome instrument selection: Lessons learned for OMERACT from imaging. Semin Arthritis Rheum. 2021 Oct;51(5):1125-1133.

216. Alkhaffaf B, Blazeby JM, Metryka A, Glenny A-M, Adeyeye A, Costa PM, et al. Methods for conducting international Delphi surveys to optimise global participation in core outcome set development: a case study in gastric cancer informed by a comprehensive literature review. Trials. 2021;22(1):410.

217. Schirmer J. Ethical issues in the use of multiple survey reminders. J Acad Ethics. 2009;7(1):125-39.

218. Lechner A, Coleman S, Balzer K, Kirkham JJ, Muir D, Nixon J, et al. Core outcomes for pressure ulcer prevention trials: results of an international consensus study: Classification: Outcomes and qualitative research. Br J Dermatol. 2022;187(5):743-52.

219. Smith H, Horobin A, Fackrell K, Colley V, Thacker B, Hall DA, et al. Defining and evaluating novel procedures for involving patients in Core Outcome Set research: creating a meaningful long list of candidate outcome domains. Res Involv Engagem. 2018;4(1):8.

220. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev. 2009;2009(3):Mr000008.

221. Rowlands G, Protheroe J, Winkley J, Richardson M, Seed PT, Rudd R. A mismatch between population health literacy and the complexity of health information: an observational study. Br J Gen Pract. 2015;65(635):e379-86.

222. Simonds VW, Garroutte EM, Buchwald D. Health literacy and informed consent materials: designed for documentation, not comprehension of health research. J Health Commun. 2017;22(8):682-91.

223. Biggane AM, Williamson PR, Ravaud P, Young B. Participating in core outcome set development via Delphi surveys: qualitative interviews provide pointers to inform guidance. BMJ Open. 2019;9(11):e032338.

224. Kostanjsek N. Use of the International Classification of Functioning, Disability and Health (ICF) as a conceptual framework and common language for disability statistics and health information systems. BMC Public Health. 2011;11(Suppl 4):S3.

225. de Kleuver M, Faraj SSA, Holewijn RM, Germscheid NM, Adobor RD, Andersen M, et al. Defining a core outcome set for adolescent and young adult patients with a spinal deformity. Acta Orthop. 2017;88(6):612-8.

226. Clarke L, Ridgewell E, Dillon MP. Identifying and linking prosthetic outcomes to the ICF framework: a step to inform the benefits measured in prosthetic health economic evaluations. Disabil Rehabil. 2023;45(6):1103-13.

227. Doubblestein D, Koehler L, Anderson E, Scheiman N, Stewart P, Schaverien M, et al. Development of a core outcome set for breast cancer-related lymphedema: a Delphi study. Breast Cancer Res Treat. 2024.

228. Boers M, Beaton DE, Shea BJ, Maxwell LJ, Bartlett SJ, Bingham CO, 3rd, et al. OMERACT Filter 2.1: Elaboration of the conceptual framework for outcome measurement in health intervention studies. J Rheumatol. 2019;46(8):1021-7.

229. Creative Commons. Attribution 4.0 International. 2025. https://creativecommons.org/licenses/by/2.0/ Accessed on 1st October 2025.

230. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. J Clin Epidemiol. 2018;96:84-92.

231. Vanhala A, Lehto A-R, Maksimow A, Torkki P, Kivivuori S-M. Classifying outcomes in secondary and tertiary care clinical quality registries—an organizational case study with the COMET taxonomy. BMC Health Serv Res. 2022;22(1):806.

232. Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. J Invest Dermatol. 2015;135(1):24-30.

233. Orbai AM, de Wit M, Mease PJ, Callis Duffin K, Elmamoun M, Tillett W, et al. Updating the psoriatic arthritis (PsA) core domain set: a report from the PsA workshop at omeract 2016. J Rheumatol. 2017;44(10):1522-8.

234. OMERACT. Downloadable forms: Core domain set development. 4.3 OMERACT Onion template. 2025. https://omeract.org/resources/downloadable-forms/ Accessed on 1st October 2025.

235. COSMIN. COSMIN database of systematic reviews of outcome measurement instruments. https://database.cosmin.nl/ Accessed on 6th July 2025.

236. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA. 1995;273(1):59-65.

237. Emery M-P, Perrier L-L, Acquadro C. Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID): Frequently asked questions. Health Qual Life Outcomes. 2005;3(1):12.

238. Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID). https://eprovide.mapi-trust.org/about/about-proqolid Accessed on 6th July 2025.

239. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. Trials. 2016;17(1):449.

240. Beaton DE, Maxwell LJ, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Instrument selection using the OMERACT Filter 2.1: the OMERACT Methodology. J Rheumatol. 2019;46(8):1028-35.

241. Leung YY, Orbai AM, Tillett W, Ogdie A, Eder L, Goel N, et al. Instruments measuring physical function for psoriatic arthritis endorsed at GRAPPA 2020 annual meeting: updates of the GRAPPA-OMERACT Working Group. J Rheumatol Suppl. 2021;97:60-3.

242. Hughes KL, Clarke M, Williamson PR. A systematic review finds Core Outcome Set uptake varies widely across different areas of health. J Clin Epidemiol. 2021;129:114-23.

243. Kirkham JJ, Bracken M, Hind L, Pennington K, Clarke M, Williamson PR. Industry funding was associated with increased use of core outcome sets. J Clin Epidemiol. 2019;115:90-7.

244. Kirkham JJ, Clarke M, Williamson PR. A methodological approach for assessing the uptake of core outcome sets using ClinicalTrials.gov: findings from a review of randomised controlled trials of rheumatoid arthritis. BMJ. 2017;357:j2262.

245. Smith TO, Mansfield M, Hawker GA, Hunter DJ, March LM, Boers M, et al. Uptake of the OMERACT-OARSI Hip and Knee Osteoarthritis Core Outcome Set: Review of Randomized Controlled Trials from 1997 to 2017. J Rheumatol. 2019;46(8):976-80.

246. Tunis SR, Maxwell LJ, Graham ID, Shea BJ, Beaton DE, Bingham CO, 3rd, et al. Engaging stakeholders and promoting uptake of OMERACT core outcome instrument sets. J Rheumatol. 2017;44(10):1551-9.

247. Williamson PR, Barrington H, Blazeby JM, Clarke M, Gargon E, Gorst S, et al. Review finds core outcome set uptake in new studies and systematic reviews needs improvement. J Clin Epidemiol. 2022;150:154-64.

248. Hughes KL, Williamson PR, Young B. In-depth qualitative interviews identified barriers and facilitators that influenced chief investigators' use of core outcome sets in randomised controlled trials. J Clin Epidemiol. 2022;144:111-20.

249. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. J Rheumatol Suppl. 1994;41:86-9.

250. Navarro-Compán V, Boel A, Boonen A, Mease P, Landewé R, Kiltz U, et al. The ASAS-OMERACT core domain set for axial spondyloarthritis. Semin Arthritis Rheum. 2021;51(6):1342-9.

251. UK Standards for Public Involvement. https://sites.google.com/nihr.ac.uk/pi-standards/home Accessed on 19th February 2025.

252. National Institute for Health and Care Research (NIHR). Payment guidance for researchers and professionals involving people in research. https://www.nihr.ac.uk/payment-guidance-researchers-and-professionals Accessed on 19th February 2025.

253. Anderson AM, Brading L, Swaithes L, Evans N, Fedorowicz SE, Murinas D, et al. Building trust and inclusion with under-served groups: a public involvement project employing a knowledge mobilisation approach. Res Involv Engagem. 2024;10(1):122.

254. Chapman LS, Flurey CA, Redmond AC, Richards P, Hofstetter C, Tapster B, et al. Living with foot and ankle disorders in rheumatic and musculoskeletal diseases: A systematic review of qualitative studies to inform the work of the OMERACT Foot and Ankle Working Group. Semin Arthritis Rheum. 2023;61:152212.

255. CASP qualitative checklist. https://casp-uk.net/casp-tools-checklists/qualitative-studies-checklist Accessed on 1st June 2022.

256. Lockwood C MZ, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015;13(3):179–187.

257. Page MJ, O'Connor DA, Malek M, Haas R, Beaton D, Huang H, et al. Patients' experience of shoulder disorders: a systematic review of qualitative studies for the OMERACT shoulder core domain set. Rheumatology (Oxford). 2019.

258. Wainwright M, Zahroh RI, Tunçalp Ö, Booth A, Bohren MA, Noyes J, et al. The use of GRADE-CERQual in qualitative evidence synthesis: an evaluation of fidelity and reporting. Health Res Pol Syst. 2023;21(1):77.

259. Colvin CJ, Garside R, Wainwright M, Munthe-Kaas H, Glenton C, Bohren MA, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 4: how to assess coherence. Implement. Sci. 2018;13(1):13.

260. Glenton C, Carlsen B, Lewin S, Munthe-Kaas H, Colvin CJ, Tunçalp Ö, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 5: how to assess adequacy of data. Implement. Sci. 2018;13(1):14.

261. Lewin S, Bohren M, Rashidian A, Munthe-Kaas H, Glenton C, Colvin CJ, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 2: how to make an overall CERQual assessment of confidence and create a Summary of Qualitative Findings table. Implement Sci. 2018;13(1):10.

262. Noyes J, Booth A, Lewin S, Carlsen B, Glenton C, Colvin CJ, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings–paper 6: how to assess relevance of the data. Implementation Science. 2018;13(1):4.

263. Backhouse MR, Vinall-Collier KA, Redmond AC, Helliwell PS, Keenan AM. Interpreting outcome following foot surgery in people with rheumatoid arthritis. J Foot Ankle Res. 2016;9:20.

264. Björk M, Thyberg I, Valtersson E, Östlund G, Stenström B, Sverker A. Foot barriers in patients with early rheumatoid arthritis: an interview study among Swedish women and men. Arthritis Care Res (Hoboken). 2018;70(9):1348-54.

265. Blake A, Mandy PJ, Stew G. Factors influencing the patient with rheumatoid arthritis in their decision to seek podiatry. Musculoskelet. Care. 2013;11(4):218-28.

266. de Souza S, Williams R, Lempp H. Patient and clinician views on the quality of foot health care for rheumatoid arthritis outpatients: a mixed methods service evaluation. J Foot Ankle Res. 2016;9:1.

267. Firth J, Nelson A, Briggs M, Gorecki C. Experiences of healthcare provision for foot ulceration occurring in people with rheumatoid arthritis. Musculoskelet. Care. 2013;11(3):159-67.

268. Goodacre LJ, Candy FJ. ‘If I didn’t have RA I wouldn’t give them house room’: the relationship between RA, footwear and clothing choices. Rheumatology. 2010;50(3):513-7.

269. Hoque A, Steultjens M, Dickson DM, Hendry GJ. Patients' and clinicians' perspectives on the clinical utility of the Rheumatoid Arthritis Foot Disease Activity Index. Rheumatol Int. 2022;42(10):1807-17.

270. Hendry GJ, Gibson KA, Pile K, Taylor L, Du Toit V, Burns J, et al. "They just scraped off the calluses": a mixed methods exploration of foot care access and provision for people with rheumatoid arthritis in south-western Sydney, Australia. J Foot Ankle Res. 2013;6(1):34.

271. Laitinen AM, Boström C, Hyytiä S, Stolt M. Experiences of foot health in patients with rheumatoid arthritis: a qualitative study. Disabil Rehabil. 2022;44(1):88-95.

272. Naidoo S, Anderson S, Mills J, Parsons S, Breeden S, Bevan E, et al. "I could cry, the amount of shoes I can't get into": A qualitative exploration of the factors that influence retail footwear selection in women with rheumatoid arthritis. J Foot Ankle Res. 2011;4:21.

273. Ramos-Petersen L, Nester CJ, Ortega-Avila AB, Skidmore S, Gijon-Nogueron G. A qualitative study exploring the experiences and perceptions of patients with rheumatoid arthritis before and after wearing foot orthoses for 6 months. Health Soc Care Community. 2021;29(3):829-36.

274. Sanders L, Donovan-Hall M, Borthwick A, Bowen CJ. Experiences of mobility for people living with rheumatoid arthritis who are receiving biologic drug therapy: implications for podiatry services. J Foot Ankle Res. 2017;10(1):14.

275. Tehan PE, Morpeth T, Williams AE, Dalbeth N, Rome K. "Come and live with my feet and you'll understand" - a qualitative study exploring the experiences of retail footwear in women with rheumatoid arthritis. J Foot Ankle Res. 2019;12:15.

276. Williams AE, Nester CJ, Ravey MI. Rheumatoid arthritis patients' experiences of wearing therapeutic footwear - a qualitative investigation. BMC Musculoskelet Disord. 2007;8:104.

277. Williams AE, Nester CJ, Ravey MI, Kottink A, Klapsing M-G. Women's experiences of wearing therapeutic footwear in three European countries. J Foot Ankle Res. 2010;3(1):23.

278. Williams AE, Graham AS. 'My feet: visible, but ignored . . .' A qualitative study of foot care for people with rheumatoid arthritis. Clin Rehabil. 2012;26(10):952-9.

279. Wilson O, Kirwan J, Dures E, Quest E, Hewlett S. The experience of foot problems and decisions to access foot care in patients with rheumatoid arthritis: a qualitative study. J Foot Ankle Res. 2017;10:4.

280. Conlin C, Khan RM, Wilson I, Daniels TR, Halai M, Pinsker EB. Living with both a total ankle replacement and an ankle fusion: a qualitative study from the patients' perspective. Foot Ankle Int. 2021;42(9):1153-61.

281. Pinsker EB, Sale JEM, Gignac MAM, Daniels TR, Beaton DE. "I don't have to think about watching the ground": a qualitative study exploring the concept of vigilance as an important outcome for ankle reconstruction. Arthritis Care Res (Hoboken). 2020;72(10):1367-73.

282. Thomas MJ, Moore A, Roddy E, Peat G. "Somebody to say 'come on we can sort this'": a qualitative study of primary care consultation among older adults with symptomatic foot osteoarthritis. Arthritis Care Res (Hoboken). 2013;65(12):2051-5.

283. Yeowell G, Samarji RA, Callaghan MJ. An exploration of the experiences of people living with painful ankle osteoarthritis and the non-surgical management of this condition. Physiotherapy. 2021;110:70-6.

284. Zaidi R, Pfeil M, Macgregor AJ, Goldberg A. How do patients with end-stage ankle arthritis decide between two surgical treatments? A qualitative study. BMJ Open. 2013;3(7).

285. Frecklington M, Williams A, Dalbeth N, McNair P, Gow P, Rome K. The footwear experiences of people with gout: a qualitative study. J Foot Ankle Res. 2019;12(1):38.

286. Liddle J, Roddy E, Mallen CD, Hider SL, Prinjha S, Ziebland S, et al. Mapping patients' experiences from initial symptoms to gout diagnosis: A qualitative exploration. BMJ Open. 2015;5(9):e008323.

287. Richardson JC, Liddle J, Mallen CD, Roddy E, Prinjha S, Ziebland S, et al. "Why me? I don't fit the mould … I am a freak of nature": a qualitative study of women's experience of gout. BMC Womens Health. 2015;15:122.

288. Carter K, Walmsley S, Chessman D, Rome K, Turner DE. Perspectives of patients and health professionals on the experience of living with psoriatic arthritis-related foot problems: a qualitative investigation. Clin Rheumatol. 2019;38(6):1605-13.

289. Williams AE, Blake A, Cherry L, Alcacer-Pitarch B, Edwards CJ, Hopkinson N, et al. Patients' experiences of lupus-related foot problems: A qualitative investigation. Lupus. 2017;26(11):1174-81.

290. Campbell RF, Morriss-Roberts C, Durrant B, Cahill S. "I need somebody who knows about feet" a qualitative study investigating the lived experiences of conservative treatment for patients with posterior tibial tendon dysfunction. J Foot Ankle Res. 2019;12:51.

291. Cotchett M, Rathleff MS, Dilnot M, Landorf KB, Morrissey D, Barton C. Lived experience and attitudes of people with plantar heel pain: a qualitative exploration. J Foot Ankle Res. 2020;13(1):12.

292. Ceravolo ML, Gaida JE, Keegan RJ. Quality-of-life in Achilles tendinopathy: an exploratory study Clin J Sport Med. 2020;30(5):495-502.

293. Dismore LL. A qualitative study to understand patients' experiences of foot pain and surgical consultation for hallux valgus and/or hallux rigidus. Pain and Rehabilitation: The Journal of the Physiotherapy Pain Association. 2021.

294. Dismore LL, van Wersch A, Critchley R, Murty A, Swainston K. A qualitative study to understand patients' experiences of their post-operative outcomes following forefoot surgery. Br J Pain. 2022;16(3):317-25.

295. Dando C, Bacon D, Borthwick A, Bowen C. Stakeholder views of podiatry services in the UK for people living with arthritis: a qualitative study. J Foot Ankle Res. 2020;13(1):58.

296. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: Striving to meet the trustworthiness criteria. Int J Qual Methods. 2017;16(1):1609406917733847.

297. Lam J, Aldridge R, Blackburn R, Harron K. How is ethnicity reported, described, and analysed in health research in the UK? A bibliographical review and focus group discussions with young refugees. BMC Public Health. 2023;23(1):2025.

298. Turner BE, Steinberg JR, Weeks BT, Rodriguez F, Cullen MR. Race/ethnicity reporting and representation in US clinical trials: a cohort study. Lancet Reg Health Am. 2022;11.

299. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. 'It's this whole picture, this well-being': patients' understanding of 'feeling well' with rheumatoid arthritis. Chronic Illn. 2010;6(3):228-40.

300. Tian J, Kang S, Zhang D, Huang Y, Zhao M, Gui X, et al. Global, regional, and national incidence and prevalence of systemic sclerosis. Clin Immunol. 2023;248:109267.

301. Rubio-Rivas M, Royo C, Simeón CP, Corbella X, Fonollosa V. Mortality and survival in systemic sclerosis: Systematic review and meta-analysis. Semin Arthritis and Rheum. 2014;44(2):208-19.

302. BMJ Best Practice. Limited cutaneous systemic sclerosis. 2024. https://bestpractice.bmj.com/topics/en-gb/593 Accessed on 1st February 2025.

303. Botzoris V, Drosos AA. Management of Raynaud's phenomenon and digital ulcers in systemic sclerosis. Joint Bone Spine. 2011;78(4):341-6.

304. Schmidt WA, Krause A, Schicke B, Wernicke D. Color Doppler ultrasonography of hand and finger arteries to differentiate primary from secondary forms of Raynaud's phenomenon. J Rheumatol. 2008;35(8):1591-8.

305. Sari-Kouzel H, Hutchinson CE, Middleton A, Webb F, Moore T, Griffin K, et al. Foot problems in patients with systemic sclerosis. Rheumatology (Oxford). 2001;40(4):410-3.

306. Alcacer-Pitarch B, Siddle HJ, Buch MH, Emery P, Hashmi F, Redmond AC. Foot health needs in people with systemic sclerosis: an audit of foot health care provision. Clin Rheumatol. 2011;30(12):1611-5.

307. Rirash F, Tingey PC, Harding SE, Maxwell LJ, Tanjong Ghogomu E, Wells GA, et al. Calcium channel blockers for primary and secondary Raynaud's phenomenon. Cochrane Database Syst Rev. 2017;12(12):Cd000467.

308. Bello RJ, Cooney CM, Melamed E, Follmar K, Yenokyan G, Leatherman G, et al. The therapeutic efficacy of botulinum toxin in treating scleroderma-associated Raynaud's phenomenon: a randomized, double-blind, placebo-controlled clinical trial. Arthritis Rheumatol. 2017;69(8):1661-9.

309. Denton CP, De Lorenzis E, Roblin E, Goldman N, Alcacer-Pitarch B, Blamont E, et al. The 2024 British Society for Rheumatology guideline for management of systemic sclerosis. Rheumatology. 2024;63(11):2956-75.

310. Maltez N, Hughes M, Brown E, Hickey V, Park H, Shea B, et al. Developing a core set of outcome measure domains to study Raynaud's phenomenon and digital ulcers in systemic sclerosis: Report from OMERACT 2020. Semin Arthritis Rheum. 2021;51(3):640-3.

311. OMERACT Scleroderma Vascular Disease Working Group. 2025. https://omeract.org/working-groups/systemic-sclerosis Accessed on 19th November 2025.

312. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res. 2005;15(9):1277-88.

313. Chapman LS, Alcacer-Pitarch B, Pauling JD, Flurey CA, Redmond AC, Richards P, et al. Patients' perspectives on systemic sclerosis-related Raynaud's phenomenon in the feet: A qualitative study from the OMERACT Foot and Ankle Working Group. Semin Arthritis Rheum. 2024;65:152372.

314. Williams AE, Nester CJ, Ravey MI, Kottink A, Klapsing MG. Women's experiences of wearing therapeutic footwear in three European countries. J Foot Ankle Res. 2010;3:23.

315. Alcacer-Pitarch B, Di Battista M, Redmond AC, Keenan A-M, Di Donato S, Buch MH, et al. Foot peripheral sensory neuropathy: a frequent disabling manifestation in systemic sclerosis. Rheumatology. 2025;64(6):3635-42.

316. Del Galdo F, Lescoat A, Conaghan PG, Bertoldo E, Čolić J, Santiago T, et al. EULAR recommendations for the treatment of systemic sclerosis: 2023 update. Ann Rheum Dis. 2024:ard-2024-226430.

317. Chapman LS, Flurey CA, Richards P, Redmond AC, Soliman E, Moshrif A, et al. What outcomes are important to people with foot and ankle disorders in rheumatic and musculoskeletal diseases? An OMERACT qualitative interview study across four continents. Semin Arthritis Rheum. 2025:152671.

318. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol. 2018;18(1):148.

319. Tight M. Saturation: An overworked and misunderstood Concept? Qualitative Inquiry. 2024;30(7):577-83.

320. Smith TO, Hawker GA, Hunter DJ, March LM, Boers M, Shea BJ, et al. The OMERACT-OARSI core domain set for measurement in clinical trials of hip and/or knee osteoarthritis. J Rheumatol. 2019;46(8):981-9.

321. Negrón JB, Lopez-Olivo MA, Carmona L, Christensen R, Ingegnoli F, Zamora NV, et al. Patient perspectives on long-term outcomes in rheumatoid arthritis. A qualitative study from the OMERACT patient outcomes in longitudinal studies working group. Semin Arthritis Rheum. 2023;58:152028.

322. Fuller A, Cai K, Filippou G, Pascart T, Diaz-Torne C, Hensey O, et al. Experience and impact of crystal pyrophosphate deposition (CPPD) from a patient and caregiver perspective: A qualitative exploration from the OMERACT CPPD working group. Semin Arthritis Rheum. 2021;51(3):655-60.

323. Sedgwick M, Spiers J. The Use of Videoconferencing as a Medium for the Qualitative Interview. Int J Qual Methods. 2009;8(1):1-11.

324. Hanna P. Using internet technologies (such as Skype) as a research medium: a research note. Qualitative Research. 2012;12(2):239-42.

325. Deakin H, Wakefield K. Skype interviewing: reflections of two PhD researchers. Qualitative Research. 2014;14(5):603-16.

326. Carter SM, Shih P, Williams J, Degeling C, Mooney-Somers J. Conducting Qualitative Research Online: Challenges and Solutions. Patient. 2021;14(6):711-8.

327. Richards HM, Schwartz LJ. Ethics of qualitative research: are there special issues for health services research? Family Practice. 2002;19(2):135-9.

328. Braun, V., Clarke, V. Successful Qualitative Research: A Practical Guide for Beginners. London: Sage Publications, Inc. 2013. p. 28-32.

329. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology. 2013;13(1):117.

330. Chun Tie Y, Birks M, Francis K. Grounded theory research: A design framework for novice researchers. SAGE Open Med. 2019;7:2050312118822927.

331. Sandelowski M. Whatever happened to qualitative description? Research in Nursing & Health. 2000;23(4):334-40.

332. Verbrugge LM, Jette AM. The disablement process. Soc Sci Med. 1994;38(1):1-14.

333. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol. 2014;67(7):745-53.

334. Jootun D, McGhee G, Marland GR. Reflexivity: promoting rigour in qualitative research. Nurs Stand. 2009;23(23):42-6.

335. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. Med Teach. 2022:1-11.

336. Hay-Smith EJC, Brown M, Anderson L, Treharne GJ. Once a clinician, always a clinician: a systematic review to develop a typology of clinician-researcher dual-role experiences in health research with patient-participants. BMC Med Res Methodol. 2016;16(1):95.

337. Feldmann M, Maini RN, Soriano ER, Strand V, Takeuchi T. 25 years of biologic DMARDs in rheumatology. Nat Rev Rheumatol. 2023;19(12):761-6.

338. Liddle AD, Pandit H, Judge A, Murray DW. Patient-reported outcomes after total and unicompartmental knee arthroplasty: a study of 14,076 matched patients from the National Joint Registry for England and Wales. Bone Joint J. 2015;97-b(6):793-801.

339. Evans JT, Evans JP, Walker RW, Blom AW, Whitehouse MR, Sayers A. How long does a hip replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. Lancet. 2019;393(10172):647-54.

340. Kloppenburg M, Bøyesen P, Visser AW, Haugen IK, Boers M, Boonen A, et al. Report from the OMERACT Hand Osteoarthritis Working Group: set of core domains and preliminary set of instruments for use in clinical trials and observational studies. J Rheumatol. 2015;42(11):2190-7.

341. Farrar N, Elliott D, Jepson M, Young B, Donovan JL, Conefrey C, et al. The role of healthcare professionals' communication in trial participation decisions: a qualitative investigation of recruitment consultations and patient interviews across three RCTs. Trials. 2024;25(1):829.

342. Abfalter D, Mueller-Seeger J, Raich M. Translation decisions in qualitative research: a systematic framework. Int J Soc Res Methodol. 2021;24(4):469-86.

343. Nielsen W, Strand V, Simon LS, Parodis I, Kim AHJ, Desai M, et al. OMERACT 2023 Systemic Lupus Erythematosus Special Interest Group: Winnowing and binning preliminary candidate domains for the core outcome set. Semin Arthritis Rheum. 2024;65:152380.

344. University of Leeds Privacy Notice. 2025. https://ahc.leeds.ac.uk/site/doc/privacy-notice Accessed 19th October 2024.

345. Manyara AM, Purvis A, Ciani O, Collins GS, Taylor RS. Sample size in multistakeholder Delphi surveys: at what minimum sample size do replicability of results stabilize? J Clin Epidemiol. 2024;174:111485.

346. Callaghan MJ, Gala JP, Roddy E. Ankle osteoarthritis: an online survey of current treatment practices of UK-based podiatrists and physiotherapists. J Foot Ankle Res. 2023;16(1):89.

347. Government Digital Service. Digital accessibility: Understanding accessibility requirements for public sector. 2018. https://www.gov.uk/guidance/accessibility-requirements-for-public-sector-websites-and-apps Accessed on 4th July 2024.

348. Government Analysis Function. Accessible charts: a checklist of the basics. 2023. https://analysisfunction.civilservice.gov.uk/policy-store/charts-a-checklist/ Accessed on 4th July 2024.

349. Mc Laughlin GH. SMOG Grading-a New Readability Formula. Journal of Reading. 1969;12(8):639-46.

350. Zhang Y, Tedeschi SK, Abhishek A, Hensey O, Grossberg D, Cai K, et al. Core domain set for chronic and/or recurrent manifestations of calcium pyrophosphate deposition disease: OMERACT delphi survey to establish consensus. Semin Arthritis Rheum. 2025;72:152669.

351. Beiderbeck D, Frevel N, von der Gracht HA, Schmidt SL, Schweitzer VM. Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. MethodsX. 2021;8:101401.

352. Teague S, Youssef GJ, Macdonald JA, Sciberras E, Shatte A, Fuller-Tyszkiewicz M, et al. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. BMC Med Res Methodol. 2018;18(1):151.

353. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-12.

354. Page MJ, Huang H, Verhagen AP, Buchbinder R, Gagnier JJ. Identifying a core set of outcome domains to measure in clinical trials for shoulder disorders: a modified Delphi study. RMD Open. 2016;2(2):e000380.

355. Navas C, Minton AP, Rodriguez-Leboeuf AM. The role of patient-reported outcomes to measure treatment satisfaction in drug development. Patient. 2024;17(6):603-17.

356. Baker SC, Watson BM, Jamieson B, Jamieson R. How do patients define satisfaction? The role of patient perceptions of their participation and health provider emotional expression. Health Commun. 2021;36(14):1970-9.

357. Kloppenburg M, Bøyesen P, Smeets W, Haugen IK, Liu R, Visser W, et al. Report from the OMERACT Hand Osteoarthritis Special Interest Group: advances and future research priorities. J Rheumatol. 2014;41(4):810-8.

358. Gallagher KS, Godwin J, Hendry GJ, Steultjens M, Woodburn J. A protocol for a randomised controlled trial of prefabricated versus customised foot orthoses for people with rheumatoid arthritis: the FOCOS RA trial [Foot Orthoses - Customised v Off-the-Shelf in Rheumatoid Arthritis]. J Foot Ankle Res. 2018;11:24.

359. Rome K, Clark H, Gray J, McMeekin P, Plant M, Dixon J. Clinical effectiveness and cost-effectiveness of foot orthoses for people with established rheumatoid arthritis: an exploratory clinical trial. Scand J Rheumatol. 2017;46(3):187-93.

360. Steultjens M. The Gait Rehabilitation in Early Arthritis Trial: trial registration. 2017. https://www.isrctn.com/ISRCTN14277030 Accessed on 19th October 2024.

361. Gagne OJ, Veljkovic A, Glazebrook M, Daniels TR, Penner MJ, Wing KJ, et al. Prospective cohort study on the employment status of working age patients after recovery from ankle arthritis surgery. Foot Ankle Int. 2018;39(6):657-63.

362. Loveday DT, Barr LV, Loizou CL, Barton G, Smith G. A comparative prospective cohort health economic analysis comparing ankle fusion, isolated great toe fusion and hallux valgus surgery. Foot Ankle Surg. 2018;24(1):54-9.

363. Munteanu SE, Landorf KB, Mc Clelland JA, Roddy E, Cicuttini FM, Shiell A, et al. Shoe-stiffening inserts for first metatarsophalangeal joint osteoarthritis (the SIMPLE trial): Study protocol for a randomised controlled trial. Trials. 2017;18(1).

364. Gibson JNA, Thomson CE. Arthrodesis or total replacement arthroplasty for hallux rigidus: A randomized controlled trial. Foot and Ankle International. 2005;26(9):680-90.

365. Stone OD, Ray R, Thomson CE, Gibson JNA. Long-Term Follow-up of Arthrodesis vs Total Joint Arthroplasty for Hallux Rigidus. Foot Ankle Int. 2017;38(4):375-80.

366. Goldberg AJ, Zaidi R, Thomson C, Dore CJ, Skene SS, Cro S, et al. Total ankle replacement versus arthrodesis (TARVA): protocol for a multicentre randomised controlled trial. BMJ Open. 2016;6(9).

367. Paget LDA, Bierma-Zeinstra SMA, Goedegebuure S, Kerkhoffs G, Krips R, Maas M, et al. Platelet-Rich plasma Injection Management for Ankle osteoarthritis study (PRIMA): protocol of a Dutch multicentre, stratified, block-randomised, double-blind, placebo-controlled trial. BMJ Open. 2019;9(10).

368. Divecha HM, Zubairy AI, Barrie JL, Aithal S, Fischer B, Fanshawe T, et al. First metatarsophalangeal joint arthrodesis versus proximal phalanx hemiarthroplasty for hallux rigidus: feasibility study for a randomised controlled trial. Trials. 2014;15:79.

369. Walsh TP, Merlo GB, Rutter C, Abell B, Platt SR, Arnold JB. Cost-Effectiveness of Interventions for Musculoskeletal Foot and Ankle Conditions: A Systematic Review. Arthritis Care Res (Hoboken). 2022;74(4):626-37.

370. Stamm T, van der Giesen F, Thorstensson C, Steen E, Birrell F, Bauernfeind B, et al. Patient perspective of hand osteoarthritis in relation to concepts covered by instruments measuring functioning: a qualitative European multicentre study. Ann Rheum Dis. 2009;68(9):1453-60.

371. Turnbull AE, Sahetya SK, Needham DM. Aligning critical care interventions with patient goals: A modified Delphi study. Heart Lung. 2016;45(6):517-24.

372. Fienberg SE, Loftus EF, Tanur JM. Cognitive aspects of health survey methodology: an overview. Milbank Mem Fund Q Health Soc. 1985;63(3):547-64.

373. OMERACT Systematic Lupus Erythematosus (SLE) Working Group. Delphi registration. 2025 https://clinicaltrials.gov/study/NCT07136090. Accessed on 28th August 2025.

374. Millward CP, Armstrong TS, Barrington H, Brodbelt AR, Bulbeck H, Byrne A, et al. Opportunities and challenges for the development of "core outcome sets" in neuro-oncology. Neuro Oncol. 2022;24(7):1048-55.

375. Mangwani J, Hau M, Thomson L. Research priorities in foot and ankle conditions: results of a UK priority setting partnership with the James Lind Alliance. BMJ Open. 2023;13(5):e070641.

376. Toupin-April K, Barton J, Fraenkel L, Li LC, Brooks P, De Wit M, et al. Toward the Development of a core set of outcome domains to assess shared decision-making interventions in rheumatology: Results from an OMERACT Delphi survey and consensus Meeting. J Rheumatol. 2017;44(10):1544-50.

377. Singer E, Ye C. The Use and Effects of Incentives in Surveys. Ann Am Acad Pol. 2013;645(1):112-41.

378. Alkhaffaf B, Blazeby JM, Metryka A, Glenny AM, Adeyeye A, Costa PM, et al. Methods for conducting international Delphi surveys to optimise global participation in core outcome set development: a case study in gastric cancer informed by a comprehensive literature review. Trials. 2021;22(1):410.

379. Alanazi SA, Vicenzino B, van Bergen CJA, Hunter DJ, Wikstrom EA, Menz HB, et al. Development of a core domain set for ankle osteoarthritis: An international consensus study of patients and health professionals. Osteoarthrit. Cartil. 2024;32(11):1481-91.

380. Maxwell LJ, Jones C, Bingham CO, Boers M, Boonen A, Choy E, et al. Defining domains: developing consensus-based definitions for foundational domains in OMERACT core outcome sets. Semin Arthritis Rheum. 2024;66:152423.

381. Gladman DD, Mease PJ, Strand V, Healy P, Helliwell PS, Fitzgerald O, et al. Consensus on a core set of domains for psoriatic arthritis. J Rheumatol. 2007;34(5):1167-70.

382. Hughes M, Maltez N, Brown E, Hickey V, Shea B, Pauling JD, Proudman S, Merkel PA, Herrick AL. Domain reporting in Systemic Sclerosis-Related Digital Ulcers: An OMERACT Scoping Review. Semin Arthritis Rheum. 2023 Aug;61:152220

383. Sabet TS, Anderson DB, Stubbs PW, Buchbinder R, Terwee CB, Chiarotto A, et al. Pain and physical function are common core domains across 40 core outcome sets

of musculoskeletal conditions: a systematic review. Journal of Clinical Epidemiology. 2025;180:111687.

384. Barbosa CD, Balp MM, Kulich K, Germain N, Rofail D. A literature review to explore the link between treatment satisfaction and adherence, compliance, and persistence. Patient Prefer Adherence. 2012;6:39-48.

385. Bova G, Domenichiello A, Letzen JE, Rosenberger DC, Siddons A, Kaiser U, et al. Developing consensus on core outcome sets of domains for acute, the transition from acute to chronic, recurrent/episodic, and chronic pain: results of the INTEGRATE-pain Delphi process. EClinicalMedicine. 2023;66:102340.

386. Roddy E, Thomas M, Marshall M, Rathod T, Myers H, Menz H, et al. Population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: The clinical assessment study of the foot (casf). Ann Rheum Dis. 2013;72(Suppl. 3).

387. Pereira D, Peleteiro B, Araújo J, Branco J, Santos RA, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. Osteoarthrit Carti. 2011;19(11):1270-85.

388. Arnold JB, Bowen CJ, Chapman LS, Gates LS, Golightly YM, Halstead J, et al. International Foot and Ankle Osteoarthritis Consortium review and research agenda for diagnosis, epidemiology, burden, outcome assessment and treatment. Osteoarthr. Cartil. 2022;30(7):945-55.

389. Tugwell P, Boers M. OMERACT conference on outcome measures in rheumatoid arthritis clinical trials: introduction. J Rheumatol. 1993;20(3):528-30.

390. OMERACT. Downloadable forms: Core domain set development. 3.4 Delphi protocol template (Appendix F). 2024. https://omeract.org/resources/downloadable-forms/. Accessed on 17th October 2024.

391. Wang L-W, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: Application, results, and recommendations. Res Soc Adm Pharm. 2013;9(5):503-16.

392. De Loof A. Only two sex forms but multiple gender variants: How to explain? Commun Integr Biol. 2018;11(1):e1427399.

393. Aarhus Gender Musem. In-person visit. https://konmuseum.dk/english/.

394. Maxwell LJ, Wright GC, Schultz G, Grosskleg S, Barton JL, Campbell W, et al. Embracing unity at OMERACT: Valuing equity, promoting diversity, fostering inclusivity. Semin Arthritis Rheum. 2024;66:152422.

395. Mortada M, Abdul-Sattar A, Gossec L. Fatigue in Egyptian patients with rheumatic diseases: a qualitative study. Health and Quality of Life Outcomes. 2015;13(1):134.

396. Genovese A, Borna S, Gomez-Cabello CA, Haider SA, Prabha S, Forte AJ, et al. Artificial intelligence in clinical settings: a systematic review of its role in language translation and interpretation. Ann Transl Med. 2024;12(6):117.

397. Matvienko-Sikar K, Hussey S, Mellor K, Byrne M, Clarke M, Kirkham JJ, et al. Using behavioral science to increase core outcome set use in trials. J Clin Epidemiol. 2024;168:111285.

398. OMERACT. Foot and ankle disorders. 2025. https://omeract.org/working-groups/foot-and-ankle-disorders/ Accessed on 4th February 2025.

399. Shea B, Pardo JP, Grosskleg S, Beaton DE, Conaghan P, Goettsch W, et al. Increasing uptake through collaboration in the development of core outcome sets: Lessons learned at OMERACT 2023. Semin Arthritis Rheum. 2024;66:152438.

400. European Alliance of Associations for Rheumatology (EULAR). EULAR PARE Representation Committee. https://www.eular.org/eular-pare-committee Accessed 4th November 2025.

401. American College of Rheumatology (ACR). Annual Meeting Patient Perspectives. 2025. https://rheumatology.org/annual-meeting-patient-perspectives Accessed on 4th November 2025.

402. Hughes KL, Kirkham JJ, Clarke M, Williamson PR. Assessing the impact of a research funder's recommendation to consider core outcome sets. PLoS One. 2019;14(9):e0222418.

403. Hróbjartsson A, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 explanation and elaboration: updated guideline for protocols of randomised trials. BMJ. 2025;389:e081660.

404. Rhodes S, Dodd S, Deckert S, Vasanthan L, Qiu R, Rohde JF, et al. Representation of published core outcome sets in practice guidelines. J Clin Epidemiol. 2024;169:111311.

405. Weinbrecht-Mischkewitz M, Kamal M, Asim F, Guillemin F, Goel N, Voshaar M, et al. Generating a list of potentially important contextual factors covering randomized trials, cohorts, and measurement property studies: An OMERACT initiative. Semin Arthritis Rheum. 2024;65:152370.

406. Wells GA, Tugwell P, Tomasson G, Guillemin F, Maxwell LJ, Shea BJ, et al. Composite outcomes at OMERACT: Multi-outcome domains and composite outcome domains. Semin Arthritis Rheum. 2021;51(6):1370-7.

407. Treweek S, Bevan S, Bower P, Campbell M, Christie J, Clarke M, et al. Trial Forge Guidance 1: what is a Study Within A Trial (SWAT)? Trials. 2018;19(1):139.

408. Trial Forge. SWAT resources. https://www.trialforge.org/swat-resources/ Accessed on 6th July 2024.

409. Spraul G, Koenning G. A descriptive study of foot problems in children with juvenile rheumatoid arthritis (JRA). Arthritis Care Res. 1994;7(3):144-50.

410. Ferrari J. A review of the foot deformities seen in juvenile chronic arthritis. The Foot. 1998;8(4):193-6.

411. Hendry G, Gardner-Medwin J, Watt GF, Woodburn J. A survey of foot problems in juvenile idiopathic arthritis. Musculoskelet. Care. 2008;6(4):221-32.

412. Chapman LS, Duquenne L, Garcia-Montoya L, Nam JL, Wilson M, Hensor EMA, et al. OA09 Progression of foot involvement in individuals at risk of rheumatoid arthritis: findings from a UK cohort study. Rheumatology. 2025;64(Supplement_3).

413. Mankia K, Siddle H, Di Matteo A, Alpízar-Rodríguez D, Kerry J, Kerschbaumer A, et al. A core set of risk factors in individuals at risk of rheumatoid arthritis: a systematic literature review informing the EULAR points to consider for conducting clinical trials and observational studies in individuals at risk of rheumatoid arthritis. RMD Open. 2021;7(3).

414. Chiarotto A, Ostelo RW, Turk DC, Buchbinder R, Boers M. Core outcome sets for research and clinical practice. Braz J Phys Ther. 2017;21(2):77-84.

415. Kearney A, Gargon E, Mitchell JW, Callaghan S, Yameen F, Williamson PR, et al. A systematic review of studies reporting the development of core outcome sets for use in routine care. J Clin Epidemiol. 2023;158:34-43.

416. Gargon E, Gorst SL, Matvienko-Sikar K, Williamson PR. Choosing important health outcomes for comparative effectiveness research: 6th annual update to a systematic review of core outcome sets for research. PLoS One. 2021;16(1):e0244878.

417. Kampstra NA, Zipfel N, van der Nat PB, Westert GP, van der Wees PJ, Groenewoud AS. Health outcomes measurement and organizational readiness support quality improvement: a systematic review. BMC Health Serv Res. 2018;18(1):1005.

418. Drake C, Chapman L, Burnett C, Halstead J, Redmond A, Roddy E, et al. Plantar Heel Pain Management: A Survey of UK Registered Healthcare Professionals. J Foot Ankle Res. 2025;18(4):e70087.

419. Hurst L, Mahtani K, Pluddemann A, Lewis S, Harvey K, Briggs A, Boylan A-M, Bajwa R, Haire K, Entwistle A, Handa A and Heneghan C. Defining value-based healthcare in the NHS: CEBM report May 2019.

420. Office for Natuonal Statistics. Census 2021 statistics on ethnic minority groups. 2022. https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/census2021statisticsonethnicminoritygroups#:~:text=While%207.7%20per%20cent%20of%20the%20population,they%20could%20not%20speak%20English%20at%20all2. Accessed on 4th September 2025.

421. NHS England. Improvement framework: community language translation and interpreting services. https://www.england.nhs.uk/long-read/improvement-framework-community-language-translation-and-interpreting-services/ Accessed on 19th September 2025.

422. Nuffield Foundation. Ethnic health inequalities in later life. 2019. https://www.nuffieldfoundation.org/wp-content/uploads/2019/11/Ethnic-health-inequalities-in-later-life.pdf. Accessed on 4th August 2025.

423. Adas MA, Norton S, Balachandran S, Alveyn E, Russell MD, Esterine T, et al. Worse outcomes linked to ethnicity for early inflammatory arthritis in England and Wales: a national cohort study. Rheumatology. 2022;62(1):169-80.

424. Russell MD, Gibson M, Zuckerman B, Kumar K, Dubey S, Adas MA, et al. Factors associated with biological and targeted synthetic disease-modifying antirheumatic drug initiation for rheumatoid arthritis in underserved patient groups in England and Wales, UK: a national cohort study. Lancet Rheumatol. 2025;7(1):e44-e54.

425. Allison TR, Symmons DP, Brammah T, Haynes P, Rogers A, Roxby M, et al. Musculoskeletal pain is more generalised among people from ethnic minorities than among white people in Greater Manchester. Ann Rheum Dis. 2002;61(2):151-6.

426. Ekezie W, Cassambai S, Czyznikowska B, Curtis F, O'Mahoney LL, Willis A, et al. Health and social care experience and research perception of different ethnic minority populations in the East Midlands, United Kingdom (REPRESENT study). Health Expect. 2024;27(1):e13944.

427. Pardhan S, Sehmbi T, Wijewickrama R, Onumajuru H, Piyasena MP. Barriers and facilitators for engaging underrepresented ethnic minority populations in healthcare research: an umbrella review. Int J Equity Health. 2025;24(1):70.

Appendices

Appendix A: Overview of OMIs included in relevant systematic reviews of measurement properties

Lead author (year)	Condition	Domains of interest	No. OMIs included	OMIs
Ortega-Avila (2019)	RA	Pain, perceived health status, quality of life and disability	14	AAOS Foot, AOS, BFS, FAAM, FAOS, FFI, FHSQ, FIS, MFPDQ, PHQ, RAOS, ROFPAQ, SAFE, SEFAS
Jia (2017)	Any disease/ pathology of the foot or ankle	Not specified	50	ACFAS, AOFAS, AOFAS Midfoot, AOFAS Hallux MTP–IP, AOFAS Lesser Toe MTP–IP, AOFAS-DFQ, AOS, CAIS, CAIT, CWIS, DFSCBS, DFS, DFSQUMA, FAAM, FADI, FAOS, FCCSFCB, FFI, FHSQ, FISRA, HAMSIQ, HFQoL, HFS14, JAFD, JSSF, MFA, SMFAQ-DI, SMFAQ-BI, MFPDI, MOXFQ, NeuroQoL, ODI, OIFA, OIFA-2, OMAS, OxAFQ, PROMISPF, QDFD, QUEOS, ROFPAQ, SAFAS, SAFEQ, SAFE, SAFE-Part A, SEFAS, TelephoneQ, VAS, VAS-FA, WOMUOI, WOMAC
Martin (2007)	Foot and ankle-related pathologic conditions in an orthopaedic	Not specified	14	AAOS, AOS, FAAM, FADI, FAOS, FFI, FHSQ, JAFD, KAFS, LEFS, MFS, OSS, ROFPAQ, Sports Ankle Rating System quality of life

	physical therapy setting			
van der Leeden (2008)	RA	Foot function, foot pain, foot-related disability	16	3D movement analysis, 50 foot walking time, AOFAS, FFI, gait analysis, FIS, MFPDQ, Moeckel forefoot scoring system, pressure measurement, RBFA, SIP, WOMAC
Prudenci (2019)	Ankle conditions	Function	9	AOFAS, CAIT, FAAM, FAOS, FFI, FFI-R, IdFAi, SAFAS, VISA
Naal (2009)	Ankle arthroplasty	Not specified	15	AOFAS Ankle-Hindfoot, FFI, AOS, Kofoed, Evanski and Waugh, NJOH, Takura et al. scoring system, Mazur et al. scoring system, Oxford ankle score
Schrier (2015)	Hallux valgus	Not specified	7	EQ-5D, FAOS, MOXFQ, NRS, SEFAS, SF-36, VAS
Murphy (2018)	AT	Pain, function	13	Likert scale of difficulty in sport, AAOFAS Ankle Hindfoot, FAS-AH, FOAS, FILLA, NRS, PGIC, SF-36, RAND, SF-36, VR-36, treatment satisfaction, VAS, VISA

AOFAS, American Academy of Orthopaedic Foot and Ankle Society; DFQ, Diabetic Foot Questionnaire; AOFAS-M, Midfoot; AOFAS-HJ, American Academy of Orthopaedic Foot and Ankle Society-Hallux Metatarsal-Interphalangeal Joint; AOFAS-LJ, American Academy of Orthopaedic Foot and Ankle Society-Lesser Toe Metatarsal-Interphalangeal Joint; AAOS, American Academy of Orthopaedic Surgeons; ACFAS, American College of Foot and Ankle Surgeons; AOS, Ankle Osteoarthritis Scale; BFS, Bristol Foot Score; CAIT, Cumberland Ankle Instability Tool; CAIS, Cumberland Ankle Instability Scale; CWIS, Cardiff Wound Impact Schedule; DFSCBS, Diabetic Foot Self-Care Behavior Scale; DFS, Diabetic Foot Scale; DFSQUMA, Diabetic Foot Self-Care Questionnaire; EQ-5D, EuroQol 5-Dimension Questionnaire; FAAM, Foot and Ankle Ability Measure; FADI, Foot and Ankle

Disability Index; FAOS, Foot and Ankle Outcome Score; FAS-AH, Foot and Ankle Stability-Ankle Hindfoot; FCCSFCB, Foot Care Confidence Scale / Foot Care Behaviour; FFI, Foot Function Index; FFI-R, Foot Function Index Revised; FHSQ, Foot Health Status Questionnaire; FILLA, Foot Injury and Lower Leg Assessment; FIS, Foot Impact Scale; FISRA, Foot and Ankle Injury Score for Rheumatoid Arthritis; FOAS, Foot and Ankle Severity Score; HAMSIQ, Hallux Metatarsophalangeal-Interphalangeal Scale Questionnaire; HFQoL, Health-Related Foot Quality of Life; HFS-14, Hindfoot Score-14 or Health Foot Score-14; IdFAI, Identification of Functional Ankle Instability; JAFD, Japanese Society for Surgery of the Foot Ankle-Foot Disability Index; JSSF, Japanese Society for Surgery of the Foot Score; KAFS, Korean Foot and Ankle Score; LEFS, Lower Extremity Functional Scale; MFA, Musculoskeletal Function Assessment; MFPDI, Manchester Foot Pain and Disability Index; MFPDQ, Manchester Foot Pain and Disability Questionnaire; MFS, Maryland Foot Score; MOXFQ, Manchester-Oxford Foot Questionnaire; NeuroQoL, Quality of Life in Neurological Disorders Measure; NJOH, New Jersey Orthopaedic Hospital Score; NRS, Numeric Rating Scale; ODI, Oswestry Disability Index; OIFA, Oxford Instability Foot and Ankle Score; OMAS, Olerud-Molander Ankle Score; OxAFQ, Oxford Ankle Foot Questionnaire; QDFD, Questionnaire for the Diagnosis of Diabetic Foot; PGIC, Patient Global Impression of Change; PHQ, Patient Health Questionnaire; PROMIS-PF, Patient-Reported Outcomes Measurement Information System-Physical Function; QUEOS, Questionário de Qualidade de Vida para Estruturas do Pé (Portuguese); RAOS, Rheumatoid Arthritis Outcome Score; RAND, Rand Corporation; RBFA, Rearfoot-Forefoot Angle; ROFPAQ, Rowan Foot Pain Assessment Questionnaire; SAFAS, Self-Administered Foot and Ankle Score; SAFE, Self-Administered Foot Evaluation; SAFEQ, Self-Administered Foot Evaluation Questionnaire; SEFAS, Self-Reported Foot and Ankle Score; SF-36, 36-Item Short Form Health Survey; SIP, Sickness Impact Profile; SMFAQ-DI, Stanford Medicine Foot and Ankle Questionnaire-Disability Index; SMFAQ-BI, Stanford Medicine Foot and Ankle Questionnaire-Burden Index; Sports, Ankle Rating System QoL, Quality of Life Section of the Sports Ankle Rating System; Telephone-Q, Telephone Questionnaire; VAS, Visual Analog Scale; VAS-FA, Visual Analog Scale-Foot and Ankle; VISA, Victorian Institute of Sport Assessment; VR-36, Veterans RAND 36-Item Health Survey; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index; WOMUOI, Western Ontario McMaster University Orthopaedic Index

Appendix B: Qualitative synthesis coding example

ID	Paper	Theme	Data	Code	Categories		
565	Dismore 2022	Theme 2: The psychosocial impact of surgical recovery	Pain was exacerbated during the evening meaning patients were fatigued and frustrated the following day.	Pain in evening, fatigue, frustration	**Fatigue**	**Evening pain**	**Emotional impact**
810	Hendry 2015	Theme 1: impact of disease-related foot symptoms	"I tried to push myself back then but it got to a point where you know you just get too fatigued." (participant 1).	Too fatigued	**Fatigue**		
1067	Ramos-Petersen	Theme 1: Improvement in physical activity	Prior to wearing foot orthoses participants not only commented about the positive mental impact of participating in sport but also how tired this made their feet and legs feel.	Tired after sport	**Fatigue**	**Activities - mental health**	
1071	Ramos-Petersen	Theme 1: Improvement in physical activity	"My legs … I feel so tired after I dance." (Participant number 1)	Tired after activity	**Fatigue**	**Activities - dancing**	
1072	Ramos-Petersen	Theme 1: Improvement in physical activity	"I went hiking with my husband and my feet were shattered." (Participant number 2)	Shattered	**Fatigue**	**Activities - hiking**	
76	Bjork 2018	Theme 3: Foot impairment:	The patients said that the tiredness and pain in the feet when working forced them to rest, slow down, and take breaks, and maybe avoid difficult work tasks.	Tiredness/pain - impact on work	**Fatigue**	**Pain**	**Work changes**
79	Bjork 2018	Theme 3: Foot impairment:	Moreover, the interviewed patients dealt with their foot pain and tiredness at work by alternating between sitting, standing, and walking at work.	Foot pain, tiredness at work	**Fatigue**	**Pain**	**Work changes**

Main coding | Pain | Activity | Emotional | Social | Occupation | Stiffness | Deformity | Sleep | **Fatigue** | Function | Muscle strength | Swelling

Appendix C: CASP

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Campbell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
1. Was there a clear statement of the aims of the research?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2. Is a qualitative methodology appropriate	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
3. Was the research design appropriate to address the aims of the research?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
4a. Has the researcher explained how participants	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
were selected?																																		
4b. Has the researcher explained why the participants they selected were the most appropriate to provide access to the type of	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
knowledge sought by the study?																																		
4c. Was there are any discussions around recruitment (e.g. why some people chose not to take part)?	Y	Y	N	N	N	N	Y	Y	N	N	N	N	N	N	N	N	Y	Y	N	N	Y	N	Y	N	Y	N	N	Y	N	N	N	Y	Y	N

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
5a. Was the setting for data collection justified?	N	N	N	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	N	N	N	N	N	N	N
5b. Is it clear how data were collected (e.g. focus group, semi-structured interview)?	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
5c. Has the researcher justified the methods chosen?	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N
5d. Has the researcher made the methods explicit (e.g. for interview method, is there an	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
indication of how interviews were conducted, or did they use a topic guide)?																																		
5e. If methods were modified during the study, has the researcher	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
explained how and why?																																		
5f. Is the form of data clear (e.g. tape recordings, video material, notes etc)?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
5g. Has the researcher discussed	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
saturation of data?																																		
6a. Has the researcher stated that they critically examined their own role, potential bias and influence during formulation of the research	N	N	Y	N	N	N	Y	N	N	N	Y	N	N	N	N	N	N	N	Y	N	N	N	N	N	N	Y	Y	Y	Y	N	N	Y	N

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
question, and data collection, including sample recruitment and choice of location?																																		
6b. Has the researcher stated whether they considered the	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

CASP item	Backhouse 2016	Bjork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
implications of any changes in the research design?																																		
7a. Are there sufficient details of how the research was explained to participants for the reader to assess whether	N	Y	Y	Y	N	N	N	N	Y	N	Y	N	N	N	N	N	N	N	Y	N	N	N	Y	N	Y	N	N	Y	Y	Y	N	N	N	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
ethical standards were maintained?																																		
7b. Has the researcher discussed issues raised by the study (e.g. issues around informed consent or	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
confidentiality or how they have handled the effects of the study on the participants during and after the study)?																																		
7c. Was approval sought from	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
an ethics committee?																																		
8a. Is there an in-depth description of the analysis process?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
8b. If thematic analysis is used, is it clear how the categories/the	Y	-	-	-	-	-	-	-	Y	Y	Y	Y	-	-	Y	Y	Y	-	-	Y	-	-	Y	Y	Y	Y	-	-	Y	-	-	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
mes were derived from the data?																																		
8c. Has the researcher explained how the data presented were selected from the original sample to demonstrate	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	Y	N	N	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
the analysis process?																																		
8d. Are sufficient data presented to support the findings?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
8e. Has the researcher taken contradictory	N	N	N	N	N	Y	Y	N	Y	Y	N	N	Y	N	Y	N	Y	N	N	Y	Y	Y	N	N	N	N	N	Y	Y	Y	N	Y	N	N

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
data into account?																																		
8f. Has the researcher stated that they critically examined their own role, potential bias and influence during analysis and selection of	N	N	Y	Y	N	N	Y	N	N	Y	Y	N	N	N	N	Y	N	N	Y	N	Y	N	N	N	N	N	Y	Y	Y	Y	N	N	Y	N

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
data for presentation?																																		
9a. Are the findings explicit?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
9b. Is there adequate discussion of the evidence both for and against the	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
researcher's arguments?																																		
9c. Has the researcher discussed the credibility of their findings (e.g. triangulation, respondent validation,	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
more than one analyst)?																																		
9d. Are the findings discussed in relation to the original research question?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10a. Has the researcher discussed the	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
contribution the study makes to existing knowledge or understanding e.g. do they consider the findings in relation to current practice or policy?, or																																		

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Camobell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
relevant research-based literature?																																		
10b. Has the researcher identified new areas where research is necessary?	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10c. Has the researcher	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

CASP item	Backhouse 2016	Biork 2018	Blake 2013	Cambell 2019	Carter 2018	Ceravolo 2020	Conlin 2021	Cotchett 2020	Dando 2020	de Souza 2016	Dismore 2021	Dismore 2022	Firth 2011	Firth 2013	Frecklington 2019	Goodacre 2011	Hendry 2013	Hoque 2022	Laitinen 2022	Liddle 2015	Naidoo 2011	Pinsker 2020	Ramos-Petersen 2021	Richardson 2015	Sanders 2017	Tehan 2019	Thomas 2013	Williams 2007	Williams 2010	Williams 2012	Williams 2017	Wilson 2017	Yeowell 2021	Zaidi 2013
discussed whether or how the findings can be transferred to other populations or considered other ways the research may be used?																																		

Appendix D: List of domains from all sources

Pathophysiological manifestation domains				
Signs				
Scoping review	**Qualitative synthesis (Chapter 3)**	**Secondary analysis (Chapter 4)**	**Qualitative interviews (Chapter 5)**	**Delphi (Chapter 6)**
Joint range of motion	Lack of movement in the foot and ankle joints	-	Reduced range of movement	**Joint range of motion**
Alignment	-	-	-	-
Global disease activity • Assessment of overall condition by physician	-	-	-	**Physician global change in foot symptoms**
Crepitus	-	-	-	-
Joint swelling	Swelling • Swelling at rest • Swelling during activity • Swelling after activity	-	Swelling • Swelling at rest • Swelling after weightbearing	-

Foot/ankle disease activity • Inflammation	-	-	-	**Foot/ankle inflammation**
Presence of deformity	Deformity/change in appearance of the foot/ankle	-	Deformity/change in appearance of the foot/ankle	**Joint damage/deformity**
Presence of callosities	Skin pathologies • Callus • Dry skin	Skin pathologies • Cracking skin • Bleeding skin	Skin condition • Callus • Blisters • Dry skin • Cracking skin	**Presence of callosities**
-	Nail pathologies	-	Nail condition • Ingrowing, breaking, and thickened nails	**Nail abnormalities**
-	Ulceration • Presence of wounds • Recurrence of wounds • Wound healing time	Ulceration	Ulceration Infection	**Foot/ankle ulceration**

	• Infection			
Pain upon palpation	-	-	-	-
Joint tenderness on palpation	-	-	-	-
Pressure-pain threshold	-	-	-	-
Muscle strength	Muscle strength	-	Muscle weakness	-
Joint stability	Balance issues/feeling unstable	-	Instability Balance issues	**Joint instability**
Muscle activity	-	-	-	-
Joint girth	-	-	-	-
Joint temperature	-	-		-
Clinician-assessment of gait	-	-	-	**Normality of gait**
-	-	-	Poor circulation	**Sufficiency of blood supply**
-	-	Colour changes	Colour changes	**Colour changes**
-	-	Cramping	Cramping Cramping during weightbearing	**Cramping**

			Cramping at rest Cramping at night	
Symptoms				
Foot/ankle pain • Pain during weightbearing • Pain during non-weightbearing • Pain intensity/severity • Pain on provocation • Pain at night	Foot/ankle pain • Pain during or after activity • Pain on first step • Pain at rest • Pain intensity/severity • Pain in the morning • Pain at night • Pain during ascending stairs • Pain during descending stairs • Pain on weightbearing	Foot/ankle pain • Pain during walking • Pain at rest • Pain on pressure	Foot/ankle pain • Pain severity • Pain during weightbearing • Pain after weightbearing • Pain at rest • Pain on movement • Pain in the morning • Pain at night	**Foot/ankle pain intensity** **Foot/ankle pain during weightbearing** **Foot/ankle pain at rest**
General pain		-	-	-

Joint stiffness • Stiffness after rest • Stiffness severity upon walking • Stiffness duration upon walking • Degree of morning stiffness	Stiffness • Stiffness after activity	Stiffness	Stiffness	**Foot/ankle stiffness** **Foot/ankle morning stiffness** (based on PPI input)
Fatigue	Fatigue	-	Fatigue • Overall fatigue • Foot/ankle-specific fatigue • Cognitive fatigue	**Overall fatigue** **Foot/ankle-specific fatigue**
Patient global change in foot/ankle symptoms / self-reported magnitude of symptom change	-	-	-	**Patient global change in foot symptoms**
Joint catching	-	-	-	-
Joint grinding	-	-	-	-
Patient-reported disease severity/assessment of	-	-	-	**Patient global change in foot symptoms**

overall condition by patient				
-	-	Temperature changes		Not included – limited evidence of importance/relevance
-	-	Numbness	Numbness	**Numbness**
Biomarkers				
Disease progression/deformity on imaging Joint structure	-	-	-	**Joint damage/deformity**
Gait abnormality Step length Stride length Heel to heel base of support Toeing out angle Cadence	Gait • Step length • Walking speed • Walking distance • Walking on different terrains • Ability to climb up/down stairs or slopes	Gait • Reduced walking distance	Gait • Altered gait/walking • Impact on other joints	**Normality of gait** to encompass gait from a HCP/researcher perspective and a patient perspective (e.g. “walking normally”)

Walking speed/velocity Step time Duration of stance phase Duration of swing phase Double support time Cycle time Contact area Contact time Peak pressure Mean pressure Mean peak pressure Pressure time integral Peak force Coronal range of motion Sagittal range of motion Transverse range of motion	• Normality of walking			

Moment Power Centre of pressure				
Disease activity on imaging	-	-	-	**Joint inflammation**
Disease activity on laboratory markers		-		
Life impact domains				
Foot/ankle function or disability	Ability to walk/mobility Activity limitations	Ability to walk/mobility	Obligatory activities/participation Committed activities/participation Discretionary activities/participation	**Physical function (necessary activities and participation)** **Physical function (discretionary activities and participation)**
Global function or disability General arthritis function	-	-	-	-

Overall quality of life/health status	Quality of life	-	Quality of life	-
Social function	• Social activity limitations • Social isolation • Changes in family roles/friendships	-	Social impact	**Physical function (activities and participation)**
Emotional status	Emotional distress • Anxiety • Depression • Anger • Sadness • Hopelessness • Frustration • Shame • Paranoia • Self-consciousness • Embarrassment • Fear of falling • Body image	Emotional distress • Fear of ulceration • Body image	Psychological impact • Changes in mood • Fear of falling • Perception of appearance • Self-confidence	**Emotional wellbeing** **Body image**

Sports participation	Physical activity limitations (including sports)	-	Physical activity limitations (including sports)	**Physical function (discretionary activities and participation)**
Footwear requirements	Footwear choice/limitations Footwear comfort Footwear appearance	Footwear choice limitations	Footwear comfort Footwear aesthetics Footwear choice	**Footwear comfort** **Footwear acceptability**
Foot/ankle related quality of life/health status	-	-	-	-
Pain interference	Pain interference	Pain interference	Pain interference	-
Treatment satisfaction	Treatment dissatisfaction	-	Treatment experiences/treatment dissatisfaction	**Treatment satisfaction**
Treatment acceptability	Treatment dissatisfaction	-	Treatment experience/treatment dissatisfaction	**Treatment satisfaction**
-	Sleep	Sleep	Sleep	**Sleep**
-	Clothing options	Impact on clothing	Impact on clothing	-

-	Impact on driving	Impact on driving	Impact on driving	**Physical function (activities and participation)**
-	Work disruption	-	Work disruption	**Physical function (necessary activities and participation)**
Adverse events/death				
Survival	-	-	-	Mandatory for all OMERACT core domain sets, thus not included in the Delphi study
Side effects	-	-	-	
Complications	-	-	-	-
Societal/resource use				

Healthcare utilisation • Need for revision/re-operation rate/ incidence of device failure/implant survivorship • Length of hospital stay • A&E/ED attendances • Hospital readmissions • Investigative procedures • Primary care contacts • Medication usage/use of rescue medication	Financial burden (personal) • Cost of treatments • Extra travel expenses • Loss of work income	- -	Healthcare utilisation • Appointments/ consultations • Hospitalisation • Medication use Personal expenses • Appointments/consultations • Medication use • Work impact • Transport costs • Device costs	**Direct costs** **Cost-effectiveness** **Personal expenses**

• Need for secondary interventions • Health service use • Days unable to work • Return to work time • Cost-effectiveness				

Appendix E: GRADE-CERQual Qualitative Evidence Profile

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
Descriptive theme 1: Pain							
Patients with foot and ankle disorders in RMDs experience pain.	All studies	Moderate concerns: 22 studies with minor concerns, 12 studies with moderate concerns.	No or very minor concerns.	No or very minor concerns.	Minor concerns: Only one study included participants with SpA, and one study included participants with CTDs. All studies were from high-income countries	Moderate confidence	Moderate concerns about methodology and minor concerns about relevance
Descriptive theme 2: Change in appearance							
Patients with foot and ankle disorders in RMDs experience joint deformity.	Backhouse 2016 Blake 2013 Bjork 2018 Carter 2018 Dando 2020	Moderate concerns: 15 studies with minor concerns, 9 studies with moderate concerns.	No or very minor concerns.	Minor concerns: 24 studies contributed to the finding, one study provided superficial data	Moderate concerns: No studies included participants with CTDs. Only one study included participants with SpA. All studies	Moderate confidence	Moderate concerns about methodology and relevance, minor concerns about adequacy

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	de Souza 2016 Dismore 2021 Dismore 2022 Goodacre 2011 Frecklington 2019 Hendry 2013 Hoque 2022 Laitinen 2022 Liddle 2015 Naidoo 2011 Ramos-Petersen 2021 Sanders 2017 Tehan 2019 Thomas 2013 Williams 2012 Williams 2007			only (Williams 2012).	were from high-income countries.		

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Williams 2010 Wilson 2017 Williams 2017						
Patients with foot and ankle disorders in RMDs experience joint swelling.	Bjork 2018 Carter 2018 de Souza 2016 Dismore 2022 Goodacre 2011 Laitinen 2022 Liddle 2015 Hoque 2022 Wilson 2017 Yeowell 2021	Moderate concerns: 6 studies with minor concerns, 4 studies with moderate concerns.	No or very minor concerns.	Moderate concerns: Only 10 studies contributed to the finding, and 3 studies (Hoque 2022, Goodacre 2011, Laitinen 2022) provided superficial data only.	Moderate concerns: Only one study included participants with MSK disorders. No studies included participants with SpA or CTDs. All studies were from high-income countries.	Low confidence	Moderate concerns about adequacy and relevance, minor concerns about methodology.
Patients with foot and ankle disorders in RMDs	Backhouse 2016 Blake 2013	Minor concerns:	No or very minor concerns.	Moderate concerns:	Moderate concerns:	Moderate confidence	Moderate concerns about adequacy and

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
experience skin and nail complaints, including wounds.	Carter 2018 Firth 2011 Firth 2012 Hendry 2013 Hoque 2022 Laitinen 2022 Williams 2012 Williams 2017 Wilson 2017	8 studies with minor concerns, 3 studies with moderate concerns.		Only 11 studies contributed to the finding and 3 studies (Hendry 2013, Hoque 2022, Laitinen 2022) provided superficial data only.	Only one study included participants with SpA, and one study included participants with CTDs. No studies involved participants with gout or MSK disorders. All studies were from high-income countries.		relevance, minor concerns about methodology.
Descriptive theme 3: Function							
Patients with foot and ankle disorders in RMDs experience limitations in walking and changes in gait.	Backhouse 2016 Bjork 2018 Campbell 2019 Carter 2018 Conlin 2021 Cotchett 2020	Moderate concerns: 20 studies with minor concerns, 10 studies with moderate concerns.	No or very minor concerns.	Minor concerns: 30 studies contributed to the finding; 9 studies (Campbell 2019, Goodacre 2011, Frecklington 2019,	Minor concerns: Only one study included participants with SpA, and only one study included participants CTDs. All	Moderate confidence	Moderate concerns about methodology, minor concerns about adequacy and relevance.

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Dando 2020 de Souza 2016 Dismore 2021 Dismore 2022 Firth 2011 Frecklington 2019 Goodacre 2011 Hendry 2013 Hoque 2022 Laitinen 2022 Liddle 2015 Naidoo 2011 Pinsker 2020 Ramos-Petersen 2021 Richardson 2015 Sanders 2017			Hendry 2013, Liddle 2015, Naidoo 2011, Richardson 2015, Tehan 2019, Williams 2012) provided superficial data only.	studies were from high-income countries.		

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Tehan 2019 Thomas 2013 Williams 2017 Williams 2007 Williams 2010 Williams 2012 Wilson 2017 Yeowell 2021						
Patients with foot and ankle disorders experience a reduction in physical and domestic activities.	Backhouse 2016 Bjork 2018 Carter 2018 Campbell 2019 Ceravolo 2020 Conlin 2021 Cotchett 2020 Dando 2020	Moderate concerns: 14 studies with minor concerns, 7 studies with moderate concerns.	No or very minor concerns	Minor concerns: 21 studies contributed to the finding; 1 study (Firth 2013) provided superficial data only.	Minor concerns: Only one study included participants with SpA, and one study included participants with CTDs. All studies were from high-income countries.	Moderate confidence	Moderate concerns about methodology, minor concerns about adequacy and relevance.

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Dismore 2021 Dismore 2022 Firth 2011 Firth 2013 Hendry 2013 Laitinen 2022 Pinsker 2020 Ramos-Petersen 2021 Thomas 2013 Williams 2010 Williams 2017 Wilson 2017 Zaidi 2013						

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
Patients with foot and ankle disorders in RMDs experience joint stiffness.	Bjork 2018 Carter 2018 Conlin 2021 Laitinen 2022 Hoque 2022 Ramos-Petersen 2021 Thomas 2013 Wilson 2017 Yeowell 2021	Minor concerns: 8 studies with minor concerns, 1 study with moderate concerns.	No or very minor concerns.	Moderate concerns: Only 9 studies contributed to the finding, one study (Hoque 2022) provided superficial data only.	Moderate concerns: Only one study included participants with SpA. No studies included participants with gout, CTDs or MSK disorders. All studies were from high-income countries.	Moderate confidence	Moderate concerns about adequacy and relevance, minor concerns about methodology.
Patients with foot and ankle disorders in RMDs experience a reduction in social participation.	Backhouse 2016 Bjork 2018 Campbell 2019 Carter 2018 Ceravolo 2020 Cotchett 2020	Moderate concerns: 15 studies with minor concerns, 8 studies with moderate concerns.	No or very minor concerns.	Minor concerns: 23 studies contributed to the finding, 1 study (Pinsker 2020) provided	Moderate concerns: Only one study included participants with SpA, and only one study included participants with CTDs. No studies	Moderate confidence	Moderate concerns about methodology and relevance, minor concerns about adequacy.

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Dismore 2021 Dismore 2022 Firth 2011 Goodacre 2011 Hendry 2013 Hoque 2022 Laitinen 2022 Naidoo 2011 Pinsker 2020 Ramos 2021 Tehan 2019 Thomas 2013 Williams 2007 Williams 2010 Williams 2017 Wilson 2017 Yeowell 2021			superficial data only.	included participants with gout.. All studies were from high-income countries.		

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
Descriptive theme 4: Footwear requirements							
Patients with foot and ankle disorders experience a dissatisfaction with footwear.	Backhouse 2016 Bjork 2018 Blake 2013 Campbell 2019 Carter 2018 Dando 2020 de Souza 2016 Dismore 2021 Dismore 2022 Firth 2011 Frecklington 2019 Goodacre 2011 Hendry 2013 Laitinen 2022 Naidoo 2011 Ramos 2021	Moderate concerns: 18 studies with minor concerns, 8 studies with moderate concerns.	Minor concerns: Some concerns about the fit between the data from primary studies and the review finding. Some participants in one study highlighted the importance of wearing supportive shoes to alleviate pain (Cotchett 2020), whilst some participants in another study (Frecklington 2019) noted they had found comfortable and acceptable footwear and didn't	Minor concerns: 26 studies contributed to the finding, one study (Blake 2013) provided superficial data only.	Minor concerns: Only one study included participants with SpA, and one study included participants with CTDs. All studies were from high-income countries.	Moderate confidence	Moderate concerns about methodology, minor concerns about coherence, adequacy and relevance.

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Richardson 2015 Sanders 2017 Tehan 2019 Thomas 2013 Williams 2007 Williams 2010 Williams 2012 Williams 2017 Wilson 2017 Yeowell 2021		think about shoes. Male participants in two studies (Williams 2007; Richardson 2015) also had no issues with footwear requirements.				
Descriptive theme 5: Work disruption							
Patients with foot and ankle disorders in RMDs experience disruptions in work.	Bjork 2018 Blake 2013 Campbell 2019 Carter 2018 Dando 2020	Moderate concerns: 11 studies with minor concerns, 7 studies with moderate concerns.	No or very minor concerns.	Moderate concerns: Only 18 studies contributed to the finding and 6	Moderate concerns: Only one study included participants with SpA. No studies included participants	Low confidence	Moderate concerns about methodology, adequacy and relevance.

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Dismore 2021 Dismore 2022 Firth 2011 Frecklington 2019 Hendry 2013 Hoque 2022 Liddle 2015 Pinkser 2020 Richardson 2015 Tehan 2019 Thomas 2013 Wilson 2017 Zaidi 2013			studies (Blake 2018, Dando 2020, Campbell 2019, Liddle 2015, Pinsker 2020, Thomas 2013) provided superficial data only.	with CTDs. All studies were from high-income countries.		
Descriptive theme 6: Financial burden							
Patients with foot and ankle disorders	Blake 2013 Carter 2018 Firth 2011	Moderate concerns	No or very minor concerns.	Moderate concerns:	Moderate concerns: No studies included participants with MSK	Low confidence	Moderate concerns about methodology,

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
experience financial impact.	Frecklington 2019 Goodacre 2011 Hendry 2015 Hoque 2022 Laitinen 2022 Liddle 2015 Ramos-Petersen 2021 Tehan 2019 Williams 2017 Wilson 2017 Zaidi 2013	7 studies with minor concerns, 7 studies with moderate concerns		Only 14 studies contributed to the finding, two studies (Blake 2013; Wilson 2017) provided superficial data only.	disorders. Only one study included participants with SpA, one study included participants with OA, and one study included participants with CTDs. All studies were from high-income countries.		adequacy and relevance.
Descriptive theme 6: Emotional status							
Patients with foot and ankle disorders experience emotional distress.	Backhouse 2016 Bjork 2018 Campbell 2019 Carter 2018	Moderate concerns: 17 studies with minor concerns, 11 studies with moderate concerns.	Minor concerns: Some concerns about the fit between the data	Minor concerns: 28 studies contributed to the finding, 4 studies	Minor concerns: Only one study included participants with SpA, and one	Moderate confidence	Moderate concerns about methodology, minor concerns about coherence,

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Ceravolo 2020 Cotchett 2020 Dando 2020 Dismore 2021 Dismore 2022 Firth 2011 Firth 2012 Frecklington 2019 Goodacre 2011 Hendry 2013 Hoque 2022 Laitinen 2022 Naidoo 2011 Pinsker 2020 Ramos-Petersen 2021 Richardson 2015		from primary studies and the review finding. Participants in one study attached less emotional significance to their physical capabilities (Ceravolo 2020), some described coping with anxiety and optimism about the future (Firth 2012), some were unaffected emotionally by interactions with practitioners about foot deformity (Williams 2007).	(Backhouse 2016, Dando 2020, Wilson 2017, Frecklington 2019) provided superficial data only.	included participants with CTDs. All studies were from high-income countries.		adequacy and relevance

Summary of review finding	Studies contributing to the finding	Methodological limitations	Coherence	Adequacy	Relevance	CERQual assessment of confidence in the evidence	Explanation of CERQual assessment
	Tehan 2019 Thomas 2013 Williams 2007 Williams 2010 Williams 2012 Williams 2017 Wilson 2017 Yeowell 2021						

Appendix F: HRA ethical approval (interviews)

Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

Lara Chapman
Leeds Institute of Rheumatic and Musculoskeletal Medicine
Chapel Allerton Hospital
Leeds
LS7 4SA

Email: approvals@hra.nhs.uk
HCRW.approvals@wales.nhs.uk

07 November 2022

Dear Chapman,

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**Patient priorities in relation to foot and ankle disorders in rheumatic and musculoskeletal diseases: a qualitative study**
IRAS project ID:	**318951**
Protocol number:	**N/A**
REC reference:	**22/NE/0226**
Sponsor	**University of Leeds**

I am pleased to confirm that **HRA and Health and Care Research Wales (HCRW) Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?
HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The standard conditions document "*After Ethical Review – guidance for sponsors and investigators*", issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?
Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **318951**. Please quote this on all correspondence.

Yours sincerely,
Christopher Cole
HRA Approvals Specialist
Email: approvals@hra.nhs.uk

Copy to: *Ms Jean Uniacke*

Appendix G: Consent to contact form (interviews)

CONSENT TO CONTACT FORM

Patient priorities in relation to foot and ankle disorders in rheumatic and musculoskeletal diseases: a qualitative study

I have read the participant information sheet (version _____, date _____) about the current study. I am interested in being contacted by a member of the research team regarding the possibility of being involved in the above study.

My contact details are:

Name _____

Telephone Number _____

Email: _____

Please return this form to your treating clinician, or email your contact details to Lara Chapman via L.Chapman1@leeds.ac.uk

A member of the research team will contact you to discuss the study further and answer any questions you may have.

Thank you.

Appendix H: Qualitative study recruitment pack

INVITATION LETTER

Patient priorities in relation to foot and ankle disorders in rheumatic and musculoskeletal diseases: a qualitative study

I enclose an information sheet for a study we are undertaking. You are being invited to participate because you have a rheumatic or musculoskeletal condition that is affecting your feet and/or ankles. We would like to understand more about what is most important to people living with foot and ankle problems and why.

If you decide to participate, you will be asked to attend a one-off interview to talk about your experiences of living with foot and/or ankle problems. This can be online, by telephone, or face-to-face at the clinic you usually attend. The interview is unlikely to take more than one hour.

This study is being sponsored by the University of Leeds. It is up to you to decide whether to take part or not after you have read the enclosed information sheet. If you decide you would like to take part, or if you have any questions, you can fill out the Consent to Contact form and leave it with your treating clinician, or email me using the contact details below. If you have indicated that you are happy to contacted about this study but we do not hear from you, a member of the research team will contact you. If you do not wish to take part in this study, this will not affect the care you receive.

We appreciate your time and help with this study. Thank you for taking the time to read this letter.

Yours sincerely,

Lara Chapman

Leeds Institute of Rheumatic and Musculoskeletal Medicine

2nd Floor, Chapel Allerton Hospital

Harehills Lane, Leeds, LS7 4SA

Email: L.Chapman1@leeds.ac.uk

PARTICIPANT INFORMATION SHEET

Patient priorities in relation to foot and ankle disorders in rheumatic and musculoskeletal diseases: a qualitative study

Invitation

You are being invited to take part in a research study. Before you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve. Please take as much time as you need to read the following information, and discuss it with others if you wish. Please ask us if there is anything that is not clear, or if you would like more information.

What is the purpose of the study?

Foot and ankle disorders are common in rheumatic and musculoskeletal diseases, for example in different types of arthritis. They can be painful and distressing. There are lots of different treatments for foot and ankle problems. To help patients and health professionals make decisions about treatments, we need evidence about which treatments work best. We find this out by measuring the effects of treatments on outcomes. Examples of outcomes include pain, or how hard it is to walk. At the moment, the evidence for treatments for foot and ankle problems is not very good. One main reason for this is a lack of agreement on what outcomes we should be measuring. This research aims to find out what outcomes are important to people with foot and ankle disorders and why. This will be the first step towards agreeing on a set of outcomes that can be used in all future foot and ankle research.

Why have I been chosen?

You have been chosen because you have a rheumatic or musculoskeletal condition affecting your feet and/or ankles that you are receiving treatment for.

Do I have to take part?

No. It is up to you to decide whether or not to take part. If you decide to take part you are free to leave the study at any time and without giving a reason. Your decision about whether to take part in the study, or to leave the study after you have started taking part, will not affect the routine care you receive.

What will happen to me if I take part?

If you choose to take part, you will be invited to arrange a video call via Microsoft Teams or Zoom, a telephone call, or a face-to-face meeting at your local hospital, to discuss your experiences of living with foot and/or ankle problems. You can choose how you would like the interview to be conducted. The interview should take no longer than one hour, but there will be no set time limit. With your consent, we will audio (voice) record your interview. After this, it will be transcribed by an external transcription company. This company will have signed a confidentiality agreement.

Before the interview, you will be asked to provide some personal details, such as your age, gender and ethnicity. It is important for us to collect this information so that we know whether the people who take part in this study have similar characteristics to other people with foot and ankle disorders in rheumatic and musculoskeletal diseases. You will also be asked to complete a short consent form.

What will happen to the findings of this study?

The results of the study will be available after it finishes. They will likely be published in a medical journal and be presented at scientific conferences. Anonymous quotes from the interviews may appear in the publication and presentations. None of the participants involved in the study will be identifiable. If you would like to see the results, or the publication, please ask your researcher.

What are the possible benefits of taking part?

There are no direct benefits to taking part, although this research will help our understanding of what is most important to patients who live with foot and ankle conditions. This will help us to decide what outcomes we should be measuring in other research studies to work out which treatments for foot and ankle problems are most effective.

What are the possible disadvantages and risks of taking part?

It is possible that answering questions about your foot problems or medical condition may be upsetting. You will not be required to answer any questions you do not feel comfortable answering and you will be able to take a break or end the interview at any time. Any answers you give will remain anonymous. If you decide to take part in a face-

to-face interview, you will need to travel to your local hospital; however, we can reimburse your travel expenses up to the value of £50. Additionally, we can reimburse childcare or personal assistant costs up to the value of £30 per hour if required, regardless of where the interview is conducted.

How long will the data be retained?

The recording from your interview will be destroyed after the research team have confirmed they are happy with the quality of the transcription. Any personal data and transcripts from the interviews will be stored on password protected documents on secure University of Leeds servers and destroyed five years after the study ends.

What if there is a problem?

This study requires no medical procedures or assessments, so no significant risks are anticipated. If you wish to complain or have any concerns about any aspect of the way you have been approached or treated during the course of this study, the usual National Health Service complaints mechanism is available to you.

How will information about me be used?

The University of Leeds is the sponsor for this study. We will need to use information from you and your medical records for this research project. Your medical records will only be accessed by your clinical team, not by the research team. This information will include your name, contact details, age, gender, and medical condition. People will use this information to do the research or to check your records to make sure that the research is being done properly. People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead. We will keep all information about you safe and secure. Once we have finished the study, we will keep some of the data so we can check the results, and we may use it in future research or share it with other research teams. We will write our reports in a way that no-one can work out that you took part in the study.

What are my choices about how my information is used?

You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have. We need to manage your records in

specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can I find out more about how my information is used?

You can find out more about how we use your information:

- At www.hra.nhs.uk/information-about-patients/
- In our leaflet available from https://dataprotection.leeds.ac.uk/wp-content/uploads/sites/48/2019/02/Research-Privacy-Notice.pdf
- By asking one of the research team (contact information at the end of this letter)
- By contacting the University's Data Protection officer dpo@leeds.ac.uk

What should I do now?

If you are interested in taking part in this study, please complete the Consent to Contact Form and hand this back to the clinician who told you about the study, or contact Lara Chapman (email L.Chapman1@leeds.ac.uk).

Who can I contact for further information?

If you require any further information or have any concerns while taking part in the study please contact one of the following people:

Lara Chapman (Chief Investigator)

HEE/NIHR Clinical Doctoral Research Fellow

Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds

Email: L.Chapman1@leeds.ac.uk

Dr Heidi Siddle (Academic Supervisor)

HEE/NIHR Senior Clinical Lecturer

Associate Professor, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds

Consultant Podiatrist, Leeds Teaching Hospitals NHS Trust

Email: H.Siddle@leeds.ac.uk

Professor Anthony Redmond (Academic Supervisor)

Professor of Clinical Biomechanics, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds

Email: A.Redmond@leeds.ac.uk

Alternatively, if you have any questions about this study you may wish to contact an organisation that is independent of the hospital at which you are being treated. The Patient Advice and Liaison Service, known as PALS, has been introduced to ensure that the NHS listens to patients, their relatives, carers and friends, and answers their questions and resolves their concerns as quickly as possible. For comments, suggestions or questions, please contact your local PALS.

Thank you for taking the time to read this information sheet and to consider this study.

Consent Form (interviews)

Patient priorities in relation to foot and ankle disorders in rheumatic and musculoskeletal diseases: a qualitative study

Participant ID:

1) I confirm that I have read the information sheet dated…… (version……..) for this study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2) I agree to participate in this study, which will involve taking part in an interview.

3) I understand that my participation in this study is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.

4) I understand that if I withdraw from the study after giving consent, any information held about me will be retained.

5) I consent to my interview being audio recorded.

6) I agree to direct quotes from my interview being used in reports, publications and presentations. I understand that these quotes will be anonymised.

7) I understand that relevant individuals from the University of Leeds and regulatory organisations may have access to my personal data to check the accuracy of the study.

8) I understand that the information collected during this study may be used to support other research in the future and may be shared anonymously with other researchers.

9) I agree to my personal information being stored for the purposes of this study. I understand that any information which could identify me will be

kept strictly confidential, unless I indicate that I have broken the law, am planning to break the law or am at risk of harming myself or others.

Participant Name	Date	Signature

Name of person taking consent	Date	Signature

Appendix I: Interview schedule

Symptoms

- Can you tell me about how your condition(s) affect your feet and ankles?
  Prompts: Pain, stiffness, lack of movement, swelling, change in joint shape, skin/nail complaints, ulcers, balance, muscle weakness
- How long have you have had [condition]?
- And have you been diagnosed with anything else as a result of your [condition]?
- Do you get any fatigue?
- When did you first start getting symptoms in your feet/ankles?
- How often do you have problems?
  Prompts: All day, part of the day, every day?
- When it comes to your feet, what bothers you the most?

Work/finances

- Have your foot/ankle problems limited your ability to work?
  Prompts: Paid work, voluntary work, if retired – problems before retiring?
  Prompts: Is this because of pain, not being able to stand/walk, footwear?
- Have these problems affected your finances in any way?
  Prompts: Have you had to reduce your hours? Give up work? Has this affected your income?
- Have you had any extra expenses because of pain or discomfort or damage to your feet?
  Prompts: Footwear, private podiatry, extra transport costs due to not being able to walk?

Social participation

- Have your foot/ankle problems affected the time you have been able to spend with family or friends?
- Are there any activities would have done before that you can't do now
  Prompts: sports, special occasions, leisure activities? Holidays? Driving?
  Prompts: Is that because of pain, lack of movement in your feet, footwear?

- Do you need help with any everyday activities because of your foot/ankle problems?

Emotional impact

- How does [all of this] affect your emotional wellbeing?

  Prompt: Have there been times when your foot and ankle problems have really got you down?

  Prompt: What aspect has made you feel like that? E.g. impact on work, social, family, leisure activities, self of self, footwear, clothing)?

Treatments

- What treatments have you had for your foot/ankle problems (e.g. insoles, footwear, hard skin, corns, ulcers, steroid injections, surgery, change in medication)?
- What has been successful about those treatments? Is there a reason why they weren't successful? Are you still having treatment? What has treatment allowed you to do that you couldn't do before?
- For you, what are the most important things that any treatment for foot and ankle problems should achieve?

  Prompts: reducing pain, walking more?
- Is there anything you wish your team had considered more when treating you for your foot/ankle problems? Is there anything you feel that you should have considered more?

Close

- Is there anything else would like to add about your foot/ankle problems that we have not already talked about?
- If we could do anything, what is the one thing we could do for you when it comes to your foot/ankle problems? Or what do you wish had been done?

Appendix J Data management plan (interviews)

Brief overview of project including proposed research methods

This study aims to explore the perceptions and experiences of patients living with foot and ankle disorders in rheumatic and musculoskeletal diseases. It will involve online, telephone and face-to-face interviews with participants.

1. **What data will be produced? What data will be used from other sources?**

   Participant names and contact details (email address and/or telephone number).

   Audio recordings of interviews, transcripts of these recordings.

2. **Where will data be stored? How will data be structured? Include file formats and approximate volume.**

   Participant contact details

   Once a potential participant has provided informed written or electronic consent (hosted on *OnlineSurveys*), their name and contact details will be entered into a participant identification list and they will be allocated a unique identification (ID) code. The participant identification list will only be accessible to LC, primary supervisor (Dr Heidi Siddle) and authorised representatives of the Sponsor and regulatory authorities.

   Once a participant has completed the eConsent form, it will be exported and deleted from *Online Surveys* at the earliest opportunity. A digital copy of the eConsent form will be stored within the encrypted folder on the University of Leeds secure servers and emailed to the participant. A paper copy of the eConsent form will be printed, signed and dated by me and stored in a locked filing cabinet in Dr Heidi Siddle's office at Chapel Allerton Hospital.

   Digital recordings/transcripts

   Interviews will be recorded using an encrypted digital voice recording device (e.g. a University-owned encrypted laptop, mobile phone or Dictaphone). After a recording/transcript has been created, it will be transferred to the

secure University of Leeds servers and deleted from the initial recording location at the earliest opportunity. Correspondingly, a recording will only be held on an encrypted digital voice recording device for a very brief time period. During this time period, LC will keep the device about her person or lock it in a secure physical location.

A recording of each interview/focus group will be sent for transcription to 1st Class Secretarial Services (trading name of Lawson Hardwick Limited). 1st Class Secretarial Services is a University of Leeds-approved supplier and provides a secure online system for uploading files, which involves an encrypted channel. When 1st Class Secretarial Services return a transcript, it will be verified and any corresponding recordings and automatically generated transcripts will be deleted at the earliest opportunity. The files will be Word documents.

3. **Access to data during the project. Give details of collaborators and any controls.**

   All digital files will be archived within password-protected folders on the University of Leeds secure servers and will only be accessible to LC and primary supervisor, Dr Heidi Siddle, as well as authorised representatives of the Sponsor and regulatory authorities. Two members of the research team outside of the University of Leeds (co-supervisors – Dr Caroline Flurey, University of the West of England, and Pamela Richards, OMERACT) will access the transcripts (through secure University email) to undertake coding. These will contain no identifiable data. Additionally, the third-party transcription service will need access to the digital recordings, as described in section

4. **Ethics and legal compliance: are there any 'special' requirements for your data? Any contractual or consent issues? Key policies (internal and external)**

   No.

5. **How will data be documented and described? Methodologies and protocols.**

   The data will consist of anonymised quotes from transcripts. There are no table and spreadsheet values. It is qualitative research, thus by its very nature does not need to be reproducible.

6. Training and support LC will undertake the Qualitative Research Methods course at King's College London to support data collection and analysis. LC will also undertake the Data Management Essentials course at the University of Leeds when a slot becomes available.
7. What are the plans for data sharing beyond project partners? Include justification if some of your data needs to be restricted. Include data and code. Include repository. Data will not be shared beyond project partners. The study data will include in-depth information about participants' experiences and perspectives; therefore, it will not be deposited in a repository to help ensure that participants' identities remain confidential.
8. What Intellectual Property will be generated? How will IP be protected and exploited? No IP will be generated from this qualitative focus group/interview study.
9. Who is responsible for managing the data? What resources will you need? LC will be responsible for managing the data. This is a qualitative study that does not need any major resources.

Appendix K: Domain binning and winnowing process

FIRST SOURCE	DOMAIN/CONCEPT/THEME	REASON FOR BIN/WINNOW	FINAL DELPHI TARGET DOMAIN	NAME OF BROAD DOMAIN
	This is specific domains, items, topics or elements generated from scoping reviews or literature searches.			
Scoping review	Inflammation	Included	**Foot/ankle inflammation**	Disease activity
Scoping review	Joint range of motion	Included	**Joint range of motion**	Disease activity
Scoping review	Alignment	Low importance/relevance in all RMDs	N/A	N/A
Scoping review	Global disease activity/assessment of overall condition by clinician	Included	**Physician global change in foot/ankle symptoms**	Disease activity
Scoping review	Joint swelling	Overlap with inflammation	**Foot/ankle inflammation**	Disease activity
Scoping review	Foot/ankle disease activity	Overlap with inflammation	**Foot/ankle inflammation**	Disease activity
Scoping review	Presence of deformity	Included	**Joint damage/deformity**	Disease activity
Scoping review	Presence of callosities	Included	**Presence of callosities**	Presence of callosities

Scoping review	Pain upon palpation	Low importance/relevance in all RMDs	N/A	N/A
Scoping review	Joint tenderness	Overlap with inflammation	**Foot/ankle inflammation**	Disease activity
Scoping review	Pressure-pain threshold	Lack of PAG consensus	N/A	N/A
Scoping review	Muscle strength	Lack of PAG consensus	N/A	N/A
Scoping review	Joint stability	Included	**Joint instability**	Joint instability
Scoping review	Muscle activity	Low importance/relevance to patients	N/A	N/A
Scoping review	Joint girth	Low importance/relevance to patients	N/A	N/A
Scoping review	Joint temperature	Low importance/relevance in all RMDs	N/A	N/A
Scoping review	Clinician-assessment of gait	Overlap with gait	**Normality of gait**	Gait
Scoping review	Foot/ankle pain	Overlap with other pain domains	**Pain intensity, pain during weightbearing, pain during non-weightbearing**	Pain
Scoping review	Pain during weightbearing	Included	**Pain during weightbearing**	Pain
Scoping review	Pain during non-weightbearing	Included	**Pain during non-weightbearing/at rest**	Pain
Scoping review	Foot/ankle pain severity/intensity	Included	**Pain severity/intensity**	Pain

Scoping review	Pain at night	Overlap with pain during non-weightbearing/at rest	**Pain during non-weightbearing/at rest**	Pain
Scoping review	Pain on provocation	Low importance/relevance to patients	N/A	N/A
Scoping review	General pain	Too broad	N/A	N/A
Scoping review	Joint stiffness	Included	**Joint stiffness**	Stiffness
Scoping review	Stiffness after rest	Too narrow - overlap with stiffness	N/A	N/A
Scoping review	Stiffness during weightbearing	Too narrow - overlap with stiffness	N/A	N/A
Scoping review	Fatigue	Included	**Fatigue**	Fatigue
Scoping review	Patient global change in foot/ankle symptoms	Included	**Patient global change in foot symptoms**	Disease activity
Scoping review	Joint catching	Low importance/relevance to patients	N/A	N/A
Scoping review	Joint grinding	Low importance/relevance to patients	N/A	N/A
Scoping review	Patient-reported disease severity/assessment of overall condition by patient	Overlap with patient global	**Patient global change in foot symptoms**	Disease activity
Scoping review	Disease progression/deformity on imaging	Included	**Joint damage/deformity**	Disease activity

Scoping review	Gait	Included	**Normality of gait**	Gait
Scoping review	Temporal spatial parameters	Too narrow, overlap with gait	**Normality of gait**	Gait
Scoping review	Plantar pressure	Too narrow, overlap with gait	**Normality of gait**	Gait
Scoping review	Kinematics	Too narrow, overlap with gait	**Normality of gait**	Gait
Scoping review	Kinetics	Too narrow, overlap with gait	**Normality of gait**	Gait
Scoping review	Disease activity on imaging	Overlap with inflammation	**Inflammation**	Disease activity
Scoping review	Disease activity on laboratory markers	Overlap with inflammation	**Inflammation**	Disease activity
Scoping review	Foot/ankle function or disability	Included	**Physical function (activities and participation)**	Physical function
Scoping review	Global function or disability	Too broad	**Physical function (activities and participation)**	Physical function
Scoping review	Overall quality of life/health status	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Scoping review	Social function	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Scoping review	Emotional status	Overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing

Scoping review	Sports participation	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Scoping review	Footwear requirements	Overlap with footwear acceptability/comfort	**Footwear acceptability and footwear comfort**	Footwear
Scoping review	Foot/ankle related quality of life/health status	Too broad, overlap with function	**Physical function (activities and participation)**	Physical function
Scoping review	Pain interference	Too broad, overlap with function	**Physical function (activities and participation)**	Physical function
Scoping review	Treatment satisfaction	Included	**Treatment satisfaction**	Treatment satisfaction
Scoping review	Survival	Mandatory domain	**Adverse events including death**	Death
Scoping review	Adverse events	Mandatory domain	**Adverse events including death**	Death
Scoping review	Need for revision/incidence of device failure/implant survivorship	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Hospitalisation/investigative procedures	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Length of hospital stay	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Hospital readmissions	Too narrow - part of costs	**Direct costs**	Health service use

Scoping review	Visits to healthcare providers	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	A&E attendances	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Medication/rescue medication use/need for secondary interventions	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Healthcare utilisation	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Direct/indirect costs	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Return to work time	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Days unable to work	Too narrow - part of costs	**Direct costs**	Health service use
Scoping review	Cost-effectiveness	Included	**Cost-effectiveness**	Health service use
Qualitative synthesis	Lack of movement in joints	Overlap with joint range of motion	**Joint range of motion**	Joint range of motion
Qualitative synthesis	Swelling at rest	Too narrow, overlap with swelling	**Joint swelling**	Disease activity
Qualitative synthesis	Swelling during activity	Too narrow, overlap with swelling	**Joint swelling**	Disease activity
Qualitative synthesis	Swelling after activity	Too narrow, overlap with swelling	**Joint swelling**	Disease activity
Qualitative synthesis	Change in appearance of foot/ankle	Overlap with joint deformity	**Joint damage/deformity**	Disease activity
Qualitative synthesis	Callus	Included	**Presence of callosities**	Presence of callosities

Qualitative synthesis	Dry skin	Low importance/relevance to all RMDs	N/A	N/A
Qualitative synthesis	Ingrowing toenails, thickening of toenails	Included	**Nail abnormalities**	Nail abnormalities
Qualitative synthesis	Presence of wound	Included	**Foot/ankle ulceration**	Ulceration
Qualitative synthesis	Recurrence of wounds	Recurrence is the measurement	**Foot/ankle ulceration**	Ulceration
Qualitative synthesis	Wound healing time	Healing time is the measurement	**Foot/ankle ulceration**	Ulceration
Qualitative synthesis	Infection	Low importance/relevance to all RMDs	N/A	N/A
Qualitative synthesis	Balance issues/feeling unstable	Overlap with joint instability	**Joint instability**	Joint instability
Qualitative synthesis	Circulation	Included	**Sufficiency of blood supply**	Circulation
Qualitative synthesis	Pain during/after activity	Overlap with pain during weightbearing	**Pain during weightbearing**	Pain
Qualitative synthesis	Pain at rest	Overlap with pain during non-weightbearing	**Pain during non-weightbearing/at rest**	Pain
Qualitative synthesis	Pain in the morning	Overlap with pain intensity	**Pain intensity**	Pain
Qualitative synthesis	Stiffness after activity	Too narrow, overlap with stiffness	**Joint stiffness**	Stiffness
Qualitative synthesis	Step length	Too narrow, overlap with gait	**Normality of gait**	Gait
Qualitative synthesis	Walking speed	Too narrow, overlap with gait	**Normality of gait**	Gait

Qualitative synthesis	Walking distance	Too narrow, overlap with gait	**Normality of gait**	Gait
Qualitative synthesis	Walking on different terrains	Too narrow, overlap with gait	**Normality of gait**	Gait
Qualitative synthesis	Ability to climb up/down stairs or slopes	Too narrow, overlap with gait	**Normality of gait**	Gait
Qualitative synthesis	Normality of walking	Overlap with gait	**Normality of gait**	Gait
Qualitative synthesis	Ability to walk/mobility	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Activity limitations	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Social activity limitations	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Social isolation	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Changes in family roles/friendships	Too narrow, overlap with function	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Anxiety	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing

Qualitative synthesis	Depression	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Anger	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Sadness	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Hopelessness	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Frustration	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Shame	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Paranoia	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Self-consciousness	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Embarrassment	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing

Qualitative synthesis	Fear of falling	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Body image	Too narrow, overlap with emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative synthesis	Physical activity limitations	Part of broader function domain	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Footwear choice/limitations	Included	**Footwear acceptability**	Footwear
Qualitative synthesis	Footwear comfort	Included	**Footwear appearance**	Footwear
Qualitative synthesis	Footwear appearance	Overlap with footwear choice/limitations	**Footwear acceptability**	Footwear
Qualitative synthesis	Sleep	Included	**Sleep**	Sleep
Qualitative synthesis	Clothing options	Overlap with footwear	**Footwear acceptability**	Footwear
Qualitative synthesis	Impact on driving	Too narrow, overlap with costs	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Work disruption	Too narrow, overlap with costs	**Physical function (activities and participation)**	Physical function
Qualitative synthesis	Cost of treatments	Too narrow, overlap with costs	**Personal expenses**	Health service use
Qualitative synthesis	Extra travel expenses	Too narrow, overlap with costs	**Personal expenses**	Health service use

Qualitative synthesis	Loss of work income	Too narrow, overlap with costs	**Personal expenses**	Health service use
Secondary analysis	Bleeding skin	Low importance/relevance to all RMDs	N/A	N/A
Secondary analysis	Cracking skin	Overlap with dry skin	N/A	N/A
Secondary analysis	Temperature changes	Low importance/relevance to all RMDs	N/A	N/A
Secondary analysis	Colour changes	Included	**Colour changes**	Disease activity
Secondary analysis	Cramping	Included	**Cramping**	Cramping
Secondary analysis	Pain on pressure	Low importance/relevance to patients	N/A	N/A
Secondary analysis	Numbness	Included	**Numbness**	Numbness
Secondary analysis	Fear of ulceration	Too narrow - part of emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative interviews	Swelling after weightbearing	Too narrow - part of swelling	**Joint swelling**	Disease activity
Qualitative interviews	Cramping during weightbearing	Too narrow - part of cramping	**Cramping**	Cramping
Qualitative interviews	Cramping at rest	Too narrow - part of cramping	**Cramping**	Cramping
Qualitative interviews	Cramping at night	Too narrow - part of cramping	**Cramping**	Cramping
Qualitative interviews	Overall fatigue	Included	**Overall fatigue**	Fatigue
Qualitative interviews	Cognitive fatigue	Overlap with overall fatigue	**Overall fatigue**	Fatigue
Qualitative interviews	Foot/ankle-specific fatigue	Included	**Foot/ankle specific fatigue**	Fatigue

Qualitative interviews	Altered gait/walking	Included	**Normality of gait**	Gait
Qualitative interviews	Impact on other joints	Overlap with gait	**Normality of gait**	Gait
Qualitative interviews	Changes in mood	Too narrow - part of emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative interviews	Perception of appearance	Too narrow - part of emotional wellbeing	**Emotional wellbeing**	Emotional wellbeing
Qualitative interviews	Footwear comfort	Included	**Footwear comfort**	Footwear
Qualitative interviews	Appointments/consultations	Too narrow - part of costs	**Direct costs**	Health service use

Appendix L: Ethical approval (Delphi)

UNIVERSITY OF LEEDS

26 March 2025

Dear Lara

Your research ethics application reference: 2029

Your research project: Consensus on a core set of domains for foot and ankle disorders in rheumatic and musculoskeletal diseases: An international, online, modified Delphi study

I am pleased to inform you that the above research ethics application has been reviewed by the Research Ethics Committee for School of Medicine which has issued a favourable ethical opinion based on the application submitted. **Please retain this email in your project file as it is evidence of the Committee's approval.**

Matters you should note:

- Ethics approval does not infer you have the right of access to any member of staff or student or documents and the premises of the University of Leeds. Nor does it imply any right of access to the premises of any other organisation, including clinical areas. The Committee takes no responsibility for you gaining access to staff, students and/or premises prior to, during or following your research activities.
- It is your responsibility to comply with all relevant Health and Safety, Data Protection and other legal and professional requirements and guidelines.
- You are expected to keep a record of all your approved documentation, as well as documents such as sample consent forms, risk assessments and other documents relating to the research project. This should be kept in your project file.
- Audits are undertaken on approved ethics applications. Your project could be chosen for such an audit. You should therefore ensure your project files are kept up to date and readily available for audit purposes. You will be given a two week notice period if your project is selected.
- Please always include the above research ethics application reference in any correspondence with the Research Ethics team.

If you need to make **amendments** to the original research project as submitted you are expected to seek approval from the Committee before taking any further action. Changes could include (but are not limited to) the project end date, project design or recruitment methodology, or study documentation. Please go to https://secretariat.leeds.ac.uk/research-ethics/how-to-apply-for-research-ethics-amendment/ or contact the Research Ethics team for further information at Research Ethics.

I hope your research project goes well.

Yours sincerely,

Ms Sou Sit Chung, Research Ethics, Governance & Compliance (formerly Secretariat), University of Leeds

On behalf of Dr Klaus Witte, Chair, SoMREC

Appendix M: Participant information sheet (Delphi)

COMFORT-Delphi: Participant Information Sheet

Consensus on a core set of domains for foot and ankle disorders in rheumatic and musculoskeletal diseases: An international, online, modified Delphi study

This document will explain the COMFORT-Delphi study in more detail. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide if you wish to take part.

What is the purpose of this project?

At the moment, the evidence for treatments for foot and ankle problems in rheumatic and musculoskeletal diseases (e.g. different types of arthritis, gout, heel pain) is not very good. One major reason for this is a lack of agreement on what domains we should be measuring in research studies. Examples of domains include pain, inflammation, how hard it is to walk, and quality of life. This study involves rating different domains in terms of how important you think they are. We will use this research to reach consensus on what domains should be measured in future studies.

By gathering input from diverse collaborators, we aim to narrow down the list of candidate domains to a focused set of 5-7 domains that should be included in all future research studies in this area. We are conducting this work with OMERACT, a global, volunteer-driven, not-for-profit organisation committed to improving outcomes for patients with rheumatic and musculoskeletal diseases. To learn more about OMERACT, please visit https://omeract.org/

Why have I been chosen?

We are interested in the priorities of experts with an interest in foot and ankle treatments and research. You might be an expert because you are a patient with personal experience in this area. If this is the case, we only expect you to be an expert on your condition and do not expect you to speak for all patients.

Alternatively, you could be a health professional or clinical researcher who is interested in foot and ankle disorders. We aim to recruit 400 participants to this international study (including 200 patients).

Do I have to take part?

No. It is up to you to decide whether or not to take part. You are also free to change your mind and stop being part of the study at any point, without giving a reason. This will not affect your clinical care or your work in any way. If you would like to stop being part of the study, please email the Chief Investigator at L.Chapman1@leeds.ac.uk. If you withdraw, we will no longer collect any information about you or from you, but we will keep the information about you that we already have. This information may still be used in the final study analyses.

What will this study involve?

This study will involve completing an online survey up to four times. The survey will contain a list of items, and you will be asked to rate how important you think each item is on a scale. Each item will correspond to a domain that could be measured in future foot and ankle research studies. After each survey 'round', you will receive a summary of all participants' anonymous responses.

What do I have to do?

If you agree to take part, please scroll down to the consent statement. It confirms that you are eligible, have read this information sheet, have had time to decide whether or not to take part, and agree to take part in the Delphi study. You can then complete the first survey. At the start of the survey, you will be asked to provide your name, email address, information about your condition (if you are a patient) or your job role, and demographics. You will be asked to complete a similar survey on 2-3 other occasions during the next 3-4 months, by email. It is important that you complete all surveys so that your opinions count.

What are the possible benefits of taking part?

There are no direct benefits to taking part, although this research will help us to decide what domains we should be measuring in future research studies to work out which treatments for foot and ankle problems are most effective. This could improve patient care in future.

Are there any disadvantages or risks to taking part?

No risks have been identified. Each survey will take around 15 minutes to complete.

What type of information will be sought from me and why?

We will collect your name and email address to provide feedback on survey responses and to invite you to complete the surveys at different time points. We will also collect some demographic information, including gender, ethnicity, first language and age category, to understand how representative our findings are.

What will happen to my personal information?

Your contact details (name, email address) will be kept electronically in password-protected files until the study has finished. If you complete all of the surveys and consent to be named in the study report and any associated publication, you will be recognised as a Delphi panellist. We will not publish your name without explicit consent. Your details will otherwise be securely destroyed. If you have any concerns about data privacy during the study you can email dpo@leeds.ac.uk or visit https://dataprotection.leeds.ac.uk. Please also read the Research Participants Privacy Notice.

What will happen to the results of this study?

All the contact information that we collect about you during the course of the research will be kept strictly confidential and will stored separately from the research data. We will take steps wherever possible to anonymise the research data so that you will not be personally identified in any reports or publications. The results of this study will be shared with people living with foot and ankle disorders, through patient networks. The results will also be shared with clinicians and researchers through professional organisations, scientific journals and conferences.

Who is organising/ funding the research?

This research is sponsored by the University of Leeds and funded by Lara Chapman's Health Education England (HEE)/National Institute of Health and Care Research (NIHR) Clinical Doctoral Research Fellowship. It is being delivered in collaboration with OMERACT.

Who has reviewed this study?

The University of Leeds School of Medicine Research Ethics Committee (reference 2029). This ensures that research is conducted with the least risk of burden possible to participants.

Further information

This study is sponsored by the University of Leeds and is funded by a Health Education England/National Institute of Health and Care Research fellowship. If you wish to discuss any aspect of the research study, or you have any questions or concerns, then you can contact the Chief Investigator, Lara Chapman at L.Chapman1@leeds.ac.uk or the Academic Supervisor, Professor Heidi Siddle at H.Siddle@leeds.ac.uk

Thank you for reading this information sheet and considering this study.
Please keep a copy of this information sheet for your records.

Appendix N: Data management plan (Delphi)

Provide a brief overview of your project including proposed research methods.

1. **What data will be produced? Are you using data from other sources?**

Survey data.

- What physical data will you study? What data will be produced (eg artefacts, samples, paper archives)
- What digital data will you generate? (eg field-notes, images, spreadsheets, audio interviews, survey data, an annotated bibliography)
- What original software will you generate?
- What third party data will you reuse?

2. **Where will data be stored? Include file formats and approximate volume.**

Survey data will be downloaded from Online Surveys into Microsoft Excel, and then deleted from Online Surveys. There will be three rounds of surveys. Access to the survey data in will only be available to the CI via a password protected account. Access to the Microsoft Excel files will only be available to the CI via password protection, and the files will be saved on secure University of Leeds servers (OneDrive).

- How much data you will produce over time to ensure you have enough storage?
- What file formats and software will you use?
- Do you have a logical file naming convention and directory structure?
- How will you use versioning so you can identify the current version of documents/data?
- How will data generated in the field be saved to safe University storage?

3. **Access to data during the project. Give details of collaborators and agreements.**

No data sharing agreement is required. Only the CI requires access.

- Do you need a data sharing agreement?
- Have you discussed data sharing with your research collaborators/ supervisor?
- Who needs to access data during the research? How will they access data?

4. **Ethics and legal compliance: Any ethical, contractual, special, sensitive or consent issues?**

This is an online survey; consent will be implied by completion.

Survey data will be anonymised and kept separately from personal data. Personal data will be destroyed once the study has finished, unless the participant has indicated that they would like to be acknowledged in the publication arising from this work. In this instance, as list of participants (names only) will be stored until publication.

- Have you read the University's Information Protection Policy? Data must be assessed for sensitivity and storage in line with this policy
- Are you familiar with the University's advice on data protection and GDPR (General Data Protection Regulation)?
- Does your research funder have specific data management and sharing requirements?
- Are there other policies and protocols you need to be aware of and observe? For example, NHS codes of practice?
- Will you anonymise your data?
- Should some data be destroyed? When and how?
- How and where will you record any participant consents and/or contractual requirements which impact data management and sharing?

5. **How will data be documented and described?**

Data will be collected through Online Surveys and imported into Microsoft Excel. Spreadsheet values will be clearly labelled. No additional documentation will be required.

- Will others understand your data? Write documentation. Make sure table and spreadsheet values are clearly labelled. Think about your methodologies and protocols.
- What information about data collection methodology will be recorded?
- Is it important for the research to be reproducible? Why/why not?
- What additional documentation will be required?
- Will you write software? Where will this be documented and stored for future use?

6. **Training and support**

Online Surveys is widely used data collection software. Support is provided by the software company if required. The CI has experience using this software.

- What training will you/your team need for data gathering, organisation, analysis or presentation?
- Are there relevant courses available at the University? Online?
- Who can provide support?

7. What are the plans for data sharing beyond project partners? How is this impacted if some of your data needs to be restricted?

Anonymised data (e.g. verbatim free-text comments) will be published in the open access article arising from this research. This will be made clear to participants in the Participant Information Sheet. Participants will be offered acknowledgement (by name) in the final publication arising from this study if they have completed all four rounds of the survey.

- Have you considered reasons for and against sharing data?
- Will data be openly available to everyone, or will there be access restrictions?
- If your research involves people, have you obtained appropriate consent for data sharing?
- Can your data be released immediately, or should you embargo (delay access to) the data?
- Will you use a data repository? Which one? Are there subject specific data repositories in your field?

8. What intellectual property will be generated?

N/A.

- Will you be applying for a patent?
- Will your research have commercial applications?
- Do you need to contact the Commercialisation team in the Research and Innovation Service?
- Have you read the University's Intellectual Property Policy?

9. Who is responsible for managing the data? What resources will you need?

The CI will be responsible for managing the data. The only resources needed are Online Surveys, Microsoft Excel and a secure server, all of which are available through the University of Leeds

- On projects with complex data management requirements, distinct types of roles should be specified.
- Are sufficient resources (skills, people, storage, technology) available to deliver your plan?
- What will you need?

Appendix O: Recruitment materials (Delphi)

Patient recruitment

Sent on behalf of the OMERACT Foot and Ankle Working Group

Can you help us improve care for people with foot and ankle disorders in rheumatic and musculoskeletal diseases?

We are asking patients who have experienced foot and ankle problems to take part in our research study. It is an online consensus study aiming to gain agreement on important domains that should be measured in future foot and ankle research. The study involves completing 3-4 short surveys. It is being undertaken with OMERACT (www.omeract.org). Our study should help to improve the relevance of future research findings for patients and lead to better evidence for foot and ankle treatments.

You can find out more information and take part here:
https://app.onlinesurveys.jisc.ac.uk/s/leeds/comfort-delphi

[QR code here]

Non-patient recruitment

Sent on behalf of the OMERACT Foot and Ankle Working Group

Can you help us improve care for people with foot and ankle disorders in rheumatic and musculoskeletal diseases?

We are asking health professionals, researchers, commissioners and industry representatives to take part in an online consensus study. The study aims to gain agreement on the domains that should be measured in future foot and ankle research studies. It is being undertaken with OMERACT (www.omeract.org) and involves completing 3-4 short surveys. This should help to improve evidence for foot and ankle treatments.

You can find out more information and take part here:
https://app.onlinesurveys.jisc.ac.uk/s/leeds/comfort-delphi

[QR code here]

Appendix P: Example of personalised reminder email (Delphi)

Dear [name of participant],

This is just a reminder that the third round of the COMFORT-Delphi foot and ankle survey closes on [date here]. Please let me know if you need extra time to complete it.

[Survey link here]

At the moment, [current number] out of [total number from previous round] patient participants have completed the survey.

Thank you for all of your help so far in our aim to improve treatments for people with [participant's condition].

Kind regards,

Lara

Appendix Q: Round 1 Delphi survey

COMFORT-Delphi

Hello, welcome to the COMFORT-Delphi study. This study is being conducted by the OMERACT Foot and Ankle Working Group, led by the group's Fellow, Lara Chapman, from the University of Leeds. We would appreciate your help to improve care for people who have foot and ankle problems in rheumatic and musculoskeletal diseases.

At the moment, the combined evidence for treatments for foot and ankle problems in rheumatic and musculoskeletal diseases (RMDs, e.g. different types of arthritis, gout, heel pain) is not high-quality. A major reason for this is a lack of agreement on what domains we should be measuring in research studies. A domain is WHAT we measure - examples include pain, inflammation, how hard it is to walk, and quality of life.

The **purpose of this study** is to rate different domains in terms of how important you think they are to measure, through online surveys. You will be asked to complete 3-4 surveys, about 4 weeks apart. Each survey will take about 15 minutes. We will use this research to reach consensus (agreement) on what domains should be measured in future studies.

The following groups of people are eligible to take part:

1. **Patients** (aged 18 and over) who have sought treatment for a foot or ankle problem in the last 12 months and have a diagnosis of a rheumatic or musculoskeletal condition.

Eligible conditions include:

- rheumatoid arthritis
- spondyloarthritis (e.g. psoriatic arthritis, ankylosing spondylitis, enteropathic arthritis)
- osteoarthritis
- crystal arthropathy (e.g. gout)
- connective tissue disease (e.g. scleroderma, lupus)
- a musculoskeletal condition affecting the foot or ankle in the absence of systemic disease, such as plantar heel pain or tendinopathy).

To be able to take part, you should not have diabetes, an acute injury (e.g. fracture, rupture) or a primary neurological condition (e.g. multiple sclerosis).

2. **Health professionals** who work in a clinical setting (where part of their role is employed by a clinical organisation) managing patients with foot and ankle problems in RMDs and have at least 12 months of experience treating adults with foot and ankle disorders in RMDs.

3. **Researchers, industry representatives, guideline developers** and **clinical commissioners** who have at least 12 months of experience working in an area relevant to foot and ankle disorders in RMDs.

If you are eligible, please read this **information sheet**. If you have any queries about whether or not you are eligible to take part in this study, please contact us using the details below.

Your participation in this study is entirely voluntary. You can withdraw at any time without giving a reason via the contact details below. We will keep the information about you that we already have. This information may still be used in the final study analyses.

We believe there are no known risks associated with this research study; however, as with any online related activity the risk of a breach is always possible. To the best of our ability your participation in this study will remain confidential, and only anonymised data will be published. We will minimise any risks by storing data in password protected files on secure drives. Your personal details will be stored separately to your survey answers. Your details will be destroyed when the study has finished. Further information is available via the **University of Leeds Privacy Notice**.

If you are happy to take part, you may now proceed to the next page of the survey.

Contact details
Lead Researcher: Lara Chapman
Email: L.Chapman1@leeds.ac.uk

Academic Supervisor: Professor Heidi Siddle
Email: H.Siddle@leeds.ac.uk

Demographics

Page 2 of 5

Please provide your name. *

0/32,000 characters

Please provide your email address. *

0/32,000 characters

Which country do you live in? *

0/32,000 characters

Please tell us why you are interested in participating in this Delphi. *

- ◯ I am a patient with foot or ankle problems, a caregiver, or patient representative
- ◯ I am a healthcare provider with experience in treating patients with foot/ankle problems
- ◯ I am a researcher with experience in trials of patients with foot/ankle problems
- ◯ Other (please specify below)

Please select the condition(s) you have been diagnosed with. *

- ☐ Achilles tendinitis/tendinopathy
- ☐ Ankylosing spondylitis
- ☐ Foot or ankle osteoarthritis
- ☐ Gout
- ☐ Juvenile idiopathic arthritis
- ☐ Lupus/systemic lupus eryethmatosus
- ☐ Plantar heel pain/plantar fasciitis
- ☐ Posterior tibial tendon dysfunction (PTTD)
- ☐ Pseudogout
- ☐ Psoriatic arthritis
- ☐ Rheumatoid arthritis
- ☐ Systemic sclerosis/scleroderma
- ☐ Other (please specify below)
- ☐ Not applicable

Please estimate how long you have had problems with your foot/ankle, in number of years.

What is your professional background/research field(s)?

- [ ] Rheumatologist
- [ ] Podiatrist
- [ ] Physiotherapist
- [ ] GP/family medicine
- [ ] Nurse
- [ ] Occupational therapist
- [ ] Orthotist
- [ ] Clinical trialist
- [ ] Pharmaceutical representative
- [ ] Government organisation representative
- [ ] Non-government organisation representative
- [ ] Other (please specify below)
- [ ] Not applicable

How much of your time, if any, do you spend in research? *

- ◯ I do not do any research
- ◯ Less than 10%
- ◯ 11-25%
- ◯ 26-50%
- ◯ Over 50%

We understand that a person's ethnicity or cultural heritage is not necessarily indicative of the race with which they identify. For many, each is a distinct part of a person's identity. Please select options below that best describe you. *

- [ ] East Asian (e.g., Chinese, Korean, Japanese, etc.)
- [ ] South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)
- [ ] Southeast Asian (e.g., Filipino, Vietnamese, Cambodian, etc.)
- [ ] Black African (e.g., Nigerian, Ethiopian, Congolese, etc.)
- [ ] Black Caribbean (e.g., Jamaican, Haitian, Trinidadian/Tobagonian, etc.)
- [ ] First Nations
- [ ] Inuk (Inuit)
- [ ] Métis
- [ ] Indigenous (North American)
- [ ] Latin American (e.g., Colombian, Salvadoran, Peruvian, Mexican, Chilean etc)
- [ ] Middle Eastern, West and Central Asian (e.g., Iranian, Lebanese, Afghan, etc)
- [ ] North African (e.g. Egyptian, Moroccan, Algerian, etc.)
- [ ] Pacific Islander (e.g., Native Hawaiian, Samoan, Tahitian, Polynesian, etc.)
- [ ] South Pacific (e.g., Australian, New Zealanders, etc.)
- [ ] White (e.g., English, French, Italian, Ukrainian, Polish, Swedish, etc.)
- [ ] Multiracial
- [ ] Prefer not to answer
- [ ] Other (please specify below)

What language(s) do you speak most often at home (please select all that apply) *

- [ ] Arabic
- [ ] Bengali
- [ ] Chinese
- [ ] English
- [ ] French
- [ ] Hindi
- [ ] Japanese
- [ ] Punjabi
- [ ] Russian
- [ ] Spanish
- [ ] Urdu
- [ ] Prefer not to answer
- [ ] Other (please specify below)

What is your age? *

- ◯ 18-29
- ◯ 30-39
- ◯ 40-49
- ◯ 50-59
- ◯ 60-69
- ◯ 70-79
- ◯ 80+

What term(s) best describe your current gender? Please select the option(s) that apply to you. *

- [ ] Man
- [ ] Non-binary
- [ ] Transgender
- [ ] Woman
- [ ] Prefer not to answer
- [ ] Other (please specify below)

Page 3 of 5

When we test how effective a foot and ankle treatment is, we measure the effectiveness of the treatment on different domains. On the following three pages, we are going to show you some domains that could be measured in future foot and ankle research (involving patients with rheumatic and musculoskeletal diseases).

We would like to know how important you think each of the domains below are, on a scale of 1-9. We will use your scores to develop a core domain set (a set of 5-7 domains mandatory for each clinical trial in this field). This doesn't mean that other domains cannot be included; it just means that the core domains **must** be used so that findings from different studies can be compared.

Rating a domain between 1-3 means you think the domain is **not important.**

Rating a domain between 4-6 means you think the domain is **important but not critical.**

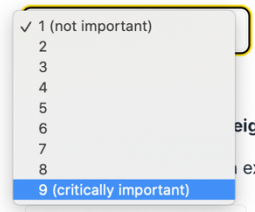
Rating a domain between 7-9 means you think the domain is **critically important.**

Pathophysiological domains

The first set of domains relate to signs, symptoms and biomarkers. If you would like to understand more about these types of domains, please click here.

Foot/ankle pain intensity

Definition: How severe pain in the foot/ankle is. *



eightbearing

experienced whilst weightbearing (e.g. walking, standing). *

Your feedback: Please tell us in the box below if:

- you would like to suggest any additional domains
- you think that some of the domains on this page can be combined
- you feel some domains are only important to measure in certain circumstances (e.g. for specific conditions or when testing specific types of treatment)

0/32,000 characters

Appendix R: Trial examples (Delphi)

Treatments for foot and ankle problems include insoles, footwear, exercises, steroid injections, change in medication and surgery.

When we test how effective a foot and ankle treatment is in a research study we measure the effectiveness of the treatment on different domains.

We are developing a core domain set for foot and ankle disorders in rheumatic and musculoskeletal diseases. This is a set of domains that must be measured in every clinical trial in this area of health. This doesn't mean that other domains cannot be included; it just means that the core domains must be used so that findings from different research studies can be compared.

Following the three previous surveys that you completed, we now have 10 domains that patients and health professionals/researchers agree are critically important for inclusion in a core domain set.

We need your help to narrow this down to 5-7 domains!

On the next page, we are going to show you the 10 domains. Each domain includes an example of how it might be used in a research study.

For each domain, we would like you to decide if it should be IN or OUT of the final core domain set, based on your own experience. Please select IN if you think the domain should be included. This means it will be mandatory to measure in all foot and ankle research studies in this area of health. In the final core domain set, no more than 7 domains can be mandatory.

Remember, any domains selected as OUT are still important and can be measured in research studies, but will be optional for researchers.

You are welcome to comment on the reasons for your decisions at the end of each section.

Section 1: Pathophysiological manifestations

You must select IN for at least one domain in this section. For full definitions, you can click on each domain.

Domain 1: Pain intensity

Example: Before a new treatment, a patient is experiencing severe pain in their foot or ankle. After treatment, the patient reports that the pain is less intense.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on pain intensity.

Should pain intensity be IN or OUT of the core domain set?

Domain 2: Pain during weightbearing

Example: Before a new treatment, a patient is experiencing pain when they walk and stand. After treatment, the patient reports that their pain is less intense during weightbearing.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on pain during weightbearing.

Should pain during weightbearing be IN or OUT of the core domain set?

Domain 3: Joint range of motion

Example: Before a new treatment, a patient has a lack of movement in one or more of their joints. This can make certain activities difficult. After treatment, the clinician reports that there is more movement in the patient's joint(s).

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on joint range of motion.

Should joint range of motion be IN or OUT of the core domain set?

Domain 4: Normality of gait

Example: Before a new treatment, a patient walks differently (e.g. slowly, with shorter steps, or with unsteady gait). After treatment, the patient's gait is more 'normal'.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on normality of gait.

Should normality of gait be IN or OUT of the core domain set?

Please add any comments relating to your decisions for including or excluding the domains in this section. If you think that a domain should only be mandatory in certain circumstances (e.g. for a specific health condition), please let us know here.

Section 2: Life impact

You must select IN for at least one domain in this section. For full definitions, you can click on each domain.

Domain 5: Physical function (activities and participation)

Example: Before a new treatment, a patient has to reduce their daily activities and their participation in life events. After treatment, the patient reports that they are able to do more daily activities and participate in more life events.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on physical function (activities and participation).

Should physical function (activities and participation) be IN or OUT of the core domain set?

Domain 6: Emotional wellbeing

Example: Before a new treatment, a patient feels unhappy and anxious because of their foot or ankle problems. This might be because they are in pain or because of the appearance of their feet. After treatment, the patient reports feeling happier and less worried.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on emotional wellbeing.

Should emotional wellbeing be IN or OUT of the core domain set?

Domain 7: Sleep

Example: Before a new treatment, a patient finds it difficult to fall and stay asleep. This could be because of pain, numbness or temperature of the foot or ankle. After treatment, the patient report that their sleep has improved.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on emotional wellbeing.

Should sleep be IN or OUT of the core domain set?

Domain 8: Footwear comfort

Example: Before a new treatment, a patient finds wearing shoes very uncomfortable. This might be because they have foot pain, swelling, or because of the shape of their feet. After treatment, the patient reports that their shoes are more comfortable.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on footwear comfort.

Should footwear comfort be IN or OUT of the core domain set?

Domain 9: Treatment satisfaction

Example: After a new treatment, a patient reports that the treatment has been good and has worked well.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on satisfaction.

Should treatment satisfaction be IN or OUT of the core domain set?

Please add any comments relating to your decisions for including or excluding the domains in this section. If you think that a domain should only be mandatory in certain circumstances (e.g. for a specific health condition), please let us know here.

Section 3: Resource use

For a full definition, you can click on the domain.

Domain 10. Personal expenses

Example: Before a new treatment, a patient is losing out financially because of their foot or ankle problems. This could be due to extra transport costs or loss of income. After treatment, the patient reports that their personal finances have improved.

When researchers are looking at how effective a treatment is, they could look at the effect of the treatment on personal expenses.

Should personal expenses be IN or OUT of the core domain set?

Please add any comments relating to your decisions for including or excluding the domains in this section. If you think that a domain should only be mandatory in certain circumstances (e.g. for a specific health condition), please let us know here.

Final ranking

Please rank the 10 domains above in order of how important they are for inclusion in the final core domain set, based on your own experience.

The **most important** domain at the top of the list (number 1). The **least important** domain should be at the bottom of the list (number 10).

You can move the domains into your chosen order by dragging them.

Appendix S: Overview of participants with multiple RMDs

Round 1

RMDs	No. participants
RA, OA	5
AS, OA	1
PHP, Sjogren's	1
OA, Sjogren's	1
AT, RA, Sjogren's	1
OA, SLE, RA	1
RA, SSc	2
OA, gout, Sjogren's	1
SSc, OA, PHP, dermatomyositis	1
Pseudogout, enteropathic arthritis	1
RA, SSc	1
OA, PsA	1
OA, AS	1
Gout, enteropathic arthritis	1
PHP, RP	1
OA, JIA, RA	1
Pseudogout, dermatomyositis	1
OA, gout, PsA, RA	1

Round 2

RMDs	**No. participants**
RA, OA	5
RA, SSc	2
OA, gout, Sjogren's	1

RA, OA, SLE	1
RA, PsA, OA, gout	1
IA, OA, SSc	1
OA, AS, gout	1
RA, Sjogren's	1

Round 3

RMDs	No. participants
RA, OA	5
RA, SSc	2
OA, gout, Sjogren's	1
Gout, enteropathic arthritis	1
RA, OA, SLE	1
OA, gout, PsA, RA,	1
IA, OA, SSc	1

Round 4

RMDs	No. participants
RA, OA	4
RA, SSc	2
OA, gout, Sjogren's	1
RA, OA, SLE	1
OA, gout, PsA, RA	1
RA, PsA, OA, gout	1

Appendix T: Overview of candidate OMIs for the COMFORT core domain set

Core domain: Pain intensity

Candidate OMI	Type of OMI	Overview of OMI
AOFAS	Composite, mixed.	Measures pain severity on a Likert scale (none; mild, occasional; moderate, daily; severe, almost always present).
FFI	Composite PROM	Measures pain severity at its worse, and also across different pain domains. Each pain question is scored on a 0-10 scale for foot pain over the last week.
FFI-R	Composite PROM	As per FFI.
McGill Pain Questionnaire	PROM	Measures three pain components: sensory intensity, cognitive evaluation of pain, and the emotional impact of pain. Patients choose words those that best describe their experience of pain. Scores are tabulated by adding values associated with each word, ranging from 0 (no pain) to 78 (severe pain).
VAS-FA	Composite PROM	Measures pain severity on a 0-100 scale.
MFDPI	Composite PROM	Contains 7 items relating to pain intensity. Responses are recorded on a 3-point scale (none of the time, on some days, on most/every day(s).
VAS	PROM	Pain intensity is rated on a 100mm line, with ‘no pain’ at one end and ‘worst pain imaginable’ at the other.

FIS	Composite PROM	Pain intensity is measured through selection of yes/no to statements, including “I cry with pain.”
FADI	Composite PROM	Contains four questions relating to pain, including pain severity rated on a scale of 0 (unbearable pain) to 4 (no pain).
JSSF	Composite, mixed	As per AOFAS.
FAOS	Composite PROM	9/42 items relate to pain, rated as none, mild, moderate, severe or extreme for various pain domains.
AOS	Composite PROM	Measures the severity of ankle pain during various activities. Each item is rated on a 100mm VAS (0 = no pain/disability, 100 = worst imaginable pain/disability).
MOXFQ	Composite PROM	16-item PROM on a five-point Likert scale. Each item is scored from 0 to 4, with 4 denoting ‘most severe’. Covers foot pain (five items).
Beuchel-Pappas	Composite, mixed	As AOFAS.
FHSQ	Composite PROM	Measures overall pain severity (1-5 none to severe).
Chronic Pain Grade	Composite PROM	Measures current pain, worst pain, and average pain in the past 6 months, on a 0–10 scale.

AOFAS, American Orthopaedic Foot and Ankle Society Score; AOS, Ankle Osteoarthritis Scale; FADI, Foot and Ankle Disability Index; FAOS, Foot and Ankle Outcome Score; FFI, Foot Function Index; FFI-R, Foot Function Index - Revised; FHSQ, Foot Health Status Questionnaire; FIS, Foot Impact Scale; JSSF, Japanese Society for Surgery of the Foot Score; MFDPI, Manchester Foot Pain and Disability Index; MOXFQ, Manchester-

Oxford Foot Questionnaire; McGill Pain Questionnaire, McGill Pain Questionnaire; VAS, Visual Analog Scale; VAS-FA, Visual Analog Scale - Foot and Ankle.

Core domain: Pain when weightbearing

Candidate OMI		Brief overview of OMI
AOS	Composite PROM	Measures the severity of ankle pain during various activities. Each item is rated on a 100mm VAS (0 = no pain/disability, 100 = worst imaginable pain/disability). Includes questions relating to walking on uneven ground, level ground, climbing stairs, descending stairs, standing for long periods, and walking long distances.
VAS	PROM	Pain when weightbearing is rated on a 100mm line, with ‘no pain’ at one end and ‘worst pain imaginable’ at the other.
FIS	Composite PROM	Contains the statement “my feet get painful when standing” (rated as true or not true).
EFAS	PROM	Measures pain when walking with the question: Do you have pain in your foot/ankle when walking? This is scored from 0 (always) to 4 (never).
FFI	Composite PROM	As described above. Pain is measured as severity, including: pain in the morning upon taking your first step, pain standing barefoot, pain walking barefoot, pain standing with

		shoes, pain walking with shoes, pain standing with orthotics, and pain walking with orthotics.
FFI-R	Composite PROM	Pain is measured as severity, including: Before you get up in the morning, first pain standing without shoes, first pain walking without shoes, first pain standing with shoes, first pain walking with shoes, pain standing with custom shoe inserts, pain walking with custom shoe inserts
FADI	Composite PROM	Contains four questions relating to pain; one question covers severity of pain during normal activity.
FAOS	Composite PROM	9/42 items relate to pain. Pain is rated as none, mild, moderate, severe or extreme for various pain domains, including pain when walking on a flat surface, going up and down stairs, and standing upright.
MOXFQ	Composite PROM	16-item PROM on a five-point Likert scale. Each item is scored from 0 to 4, with 4 denoting ‘most severe’. Covers foot pain (five items).

EFAS, European Foot and Ankle Society Score

Core domain: Physical function (activities and participation)

Candidate OMI	Type of OMI	Description of domain
ACFAS	Composite PROM	Measures pain during activities (walking, standing, climbing stairs), ability to perform daily living tasks, and sport or higher-level physical activity, on a five-point Likert scale.
AOFAS	Composite, mixed	Measures function in terms of activity limitations and support requirements, maximum walking distance, and difficulty with walking surfaces.
AOS	Composite PROM	Measures difficulty performing daily activities, including walking on flat surfaces, **c**limbing stairs, standing, squatting, and doing household or work tasks.
Buechel-Pappas ankle score	Composite, mixed	Evaluates ability to perform daily activities and the level of functional limitation.
JSSF Ankle-Hindfoot Score	Composite, mixed	As AOFAS Ankle-Hindfoot Score.
JSSF Hallux and Lesser Toe Scales	Composite, mixed	As AOFAS Ankle-Hindfoot Score.
FAAM	Composite PROM	Contains an ADL subscale measuring difficulty performing general functional tasks (e.g. walking, climbing stairs, standing), and a sports subscale measuring difficulty performing higher-level, physically demanding activities (e.g. running, jumping).
FADI	Composite PROM	As FAAM.

FAOS	Composite PROM	Contains a daily living subscale and function, sports and recreational activities subscale.
FFI	Composite PROM	Measures disability and activity limitation. The disability domain covers difficulty performing various functional activities because of foot problems, such as difficulty climbing stairs. The activity limitation domain measures limitations in activities because of foot problems, such as staying indoors or in bed, and limitations in physical activities.
FFI-R	Composite PROM	As FFI.
FFI-R SF	Composite PROM	As FFI.
FHSQ	Composite PROM	Measures the extent to which foot problems interfere with activities performed in a typical day, vigorous activities, and social activities, rated on a 1-5 scale from not at all to extremely.
FIS	Composite PROM	Includes questions relating to walking, activities and social participation. Function is measured through selection of yes/no to statements,
Likert scale	PROM	Generic scale measuring functional limitation using ordinal data.
LLTQ	Composite PROM	Measures difficulty in doing a series of tasks or activities that a patient might perform in daily life or sports, including walking on flat or uneven surfaces, climbing stairs, squatting or bending, running or jumping.
Mazur’s ankle function score	PROM	Measures ability to perform daily activities.

MFPDI	Composite PROM	Measures functional limitation (10/19 items). Responses are recorded on a 3-point scale (none of the time, on some days, on most/every day(s).
MOXFQ	Composite PROM	16-item PROM on a five-point Likert scale. Each item is scored from 0 to 4, with 4 denoting 'most severe'. Covers foot pain (five items).
NRS	PROM	Generic 0-10 or 0-100, scale, with functional limitations rated as a numerical value.
Van Valburg score	Composite, mixed	Evaluates mobility based on the patient's ability to perform daily activities and the range of motion in the ankle joint.
VAS	PROM	Generic 0-10 or 0-100, scale, with functional limitations rated as a numerical value.
VAS-FA	Composite PROM	Contains 11 questions relating to function, including impact on occupation, standing, running, daily activities, and walking on uneven ground.
WOMAC	Composite PROM	Measures difficulty in performing common daily activities, such as using stairs, rising from sitting, walking on flat surfaces, standing, getting in/out of a car, shopping, housekeeping and bending.

ACFAS, American College of Foot and Ankle Surgeons Score; FFI-R SF, Foot Function Index - Revised Short Form; LLTQ, Lower Limb Tasks Questionnaire; MFPDI, Manchester Foot Pain and Disability Index; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index

Core domain: joint movement

Candidate OMI	Type of OMI	Description of OMI
Goniometer (including electrogoniometer)	Clinician-assessed	A goniometer has two arms connected by a hinge that lines up with a joint. The angle between the arms is read in degrees to determine how far a foot or ankle joint can plantarflex or dorsiflex.
Imaging	Clinician-assessed	Radiographs (Intraoperative images) of the ankle (with the patient forcibly plantarflexing or dorsiflexing).
FAOS	Composite PROM	Includes two questions relating to joint movement (can you straighten your foot/ankle fully, and can you bend your foot/ankle fully). Patients rate the extent to which they can straighten and bend their foot/ankle on a 5-point scale. The options are: always (0), often (1), sometimes (2), rarely (3) and never (4).
AOFAS	Mixed, composite	Contains two questions relating to joint ROM: sagittal motion (flexion plus extension), and hindfoot motion (inversion plus eversion). For each question, clinicians select from three options. For sagittal motion: normal or mild restriction (30 degrees or more), moderate restriction (15 degrees to 29 degrees), severe restriction (less than 15 degrees). For hindfoot motion: normal or mild restriction (75%-100% of normal), moderate restriction (25%-74% of normal), or marked restriction (less than 25% of normal).
Beuchel-Pappas	Composite, mixed	As AOFAS.
JSSF	Composite, mixed	As AOFAS.

Core domain: Treatment satisfaction

Candidate OMI	Type of OMI	Description of OMI
Likert scale	PROM	Generic scale measuring satisfaction using ordinal data (e.g. very dissatisfied, dissatisfied, neutral, satisfied, very satisfied).
VAS	PROM	Generic 0-10 or 0-100, scale, with satisfaction rated as a numerical value.
NRS	PROM	As VAS.

Circumstance-dependent core domain: structural pathology (inflammation)

Candidate OMI	Type of OMI	Description of OMI
Laboratory markers	Clinician-assessed	Blood tests used to measure inflammation, e.g. CRP, ESR.
MRI	Clinician-assessed	Imaging.
Ultrasound	Clinician-assessed	Imaging.

CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; MRI, magnetic resonance imaging

Circumstance-dependent core domain: structural pathology (joint damage/deformity)

Candidate OMI	Type of OMI	Description of OMI
Radiograph	Clinician-assessed	Imaging.
CT	Clinician-assessed	Imaging.

CT, computed tomography

Circumstance-dependent core domain (healthcare expenses)

Candidate OMI	Type of OMI	Description of OMI
EQ-5D-5L	PROM	Measures health-related quality of life to estimate Quality-Adjusted Life Years (QALYs)
Return to work time	PROM	Time taken to return to work.
Client Service Receipt Inventory questionnaire	PROM	Measures use of healthcare services (e.g. hospitalisation and visits to medical professionals).
PRODISQ	Clinician-assessed	Measures the relationship between health and productivity (e.g. work absence, reduced productivity at work, and productivity costs at an organisational level).

EQ-5D-5L, EuroQol 5-Dimension 5-Level Questionnaire