Developing a Core Outcome Set for Hand Fractures and Joint Injuries in Adults

Sandeep Rajiv Deshmukh

MB/BChir, MA (Cantab), FRCS (Tr & Orth)

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

July 2022

Abstract

Background

Hand fractures and joint injuries are common, with significant impact for patients themselves but also on a wider societal and economic level through healthcare costs and productivity loss. Despite the clear significance of these injuries, there is not a consensus on optimal treatment. The variety of treatment options together with a lack of consistency in outcome reporting and research method standards makes interpretation of the available evidence challenging.

One solution is a core outcome set, which aims to set the minimum outcome measurements in any clinical study. This would improve consistency and comparability between studies, facilitating meta-analysis.

Aim

The overarching aim was to work towards development of a core outcome set for hand fractures and joint injuries in adults which will guide outcome assessment in future studies. The primary purpose of this project was to establish 'what' should be measured when considering the outcome of hand fracture and joint injury management.

Methods

This was a mixed methods study to develop a core outcome set for clinical trials of hand fractures and joint injuries, with initial scoping work to clarify the set of injuries intended to

be covered by the core outcome set, analysis of data from a UK Major Trauma Centre to determine a sense of the scale of the incidence of hand fractures, and then phases of outcome domain generation and consensus prioritisation to reach the final core outcome set. The specific steps were:

- Defining the scope of injuries to be addressed by the core outcome set through one-to-one discussions with expert stakeholder consultation with hand surgeons and therapists
- Collecting data from reports of all hand and wrist radiographs from Queen's Medical Centre (Nottingham) over a one-year period to assess the incidence and anatomical distribution of fractures.
- A systematic review of randomised/quasi-randomised controlled studies and large (≥100 participant) prospective observational studies on treatment of hand fractures and joint injuries to identify outcomes selected in recently published studies. An assessment of outcome reporting bias was also conducted
- Extensive exploratory qualitative research with the patient stakeholder group, to identify their perspective on the injury, treatment and outcomes important to them and to generate outcome domains as well as descriptors using an inductive, thematic approach
- Initial consensus prioritisation of a longlist of outcome domains developed through synthesis of the systematic review and qualitative work through an international three-round Delphi survey
- A final consensus meeting using an adapted nominal group technique format, involving all key stakeholders, to reach consensus on a final core outcome set.

Main findings

In the systematic review of 160 studies vast heterogeneity in outcome selection was found. There were 639 unique outcomes, which were rationalised to 74 outcome domains based on the World Health Organization International Classification of Functioning, Disability, and Health framework. Outcome reporting bias was evident, with only a minority of outcomes appropriately reported across these studies.

To explore which outcomes were relevant to patients with hand fractures and joint injuries, a qualitative study was conducted involving interviews (25 patients) and focus groups (five groups involving a total of 21 patients). A total of 35 outcome domains grouped within six broad themes were generated, along with descriptors for the domains.

The two streams of outcome domains were synthesised to form a longlist of 37 domains that entered a consensus process beginning with an online international Delphi survey. From the original 152 participants who began the survey, 144 (>94%) completed all three rounds (54 patients, 55 hand surgeons, 35 hand therapists). Based on pre-defined consensus criteria, 20 domains reached consensus as ‘very important’ and the remainder reached no consensus.

All outcome domains were discussed at a final consensus meeting with 27 participants (12 patients, seven surgeons, six hand therapists, a health economist and a trial manager). The domains reaching no consensus were discussed and voted upon, with none reaching threshold to salvage and consider for the core outcome set. The 20 ‘consensus in’ domains were discussed and underwent iterative prioritisation steps. A final vote selected seven outcome domains for inclusion in the core outcome set: fine hand use, pain / discomfort at rest, pain / discomfort with activity, self-hygiene / personal care, return to usual work / job, range of movement, and patient satisfaction with outcome / result.

Conclusion

This study has shown the magnitude of the inconsistency in outcome selection for clinical research on hand fractures and joint injuries in adults. A core outcome set to help address this issue was developed based on exploration of the existing literature and the patient perspective. Through a subsequent systematic consensus process, the longlist of outcome domains was refined to a final set of seven core outcome domains. These touch upon several bases including functional tasks (covering basic aspects and a working life role), patient comfort, abstract function (range of motion) and patient satisfaction. They are recommended as the baseline domains to be measured in future clinical research on these injuries, with the optimum way to measure the domains being the subject of future work.

Publications, Presentations, Prizes and Funding

This section contains outputs (journal articles and conference presentations) completed or in progress as a result of the work in this thesis, and prizes/funding awarded during the registered period of study.

Publications

Published

- Deshmukh, S.R., Mousoulis, C., Marson, B.A., Grindlay, D., Karantana, A. and Core Outcome Set for Hand Fractures and Joint Injuries in Adults Group, 2021. Developing a core outcome set for hand fractures and joint injuries in adults: a systematic review. *Journal of Hand Surgery (European volume)*, 46(5), pp.488-495.

Accepted – awaiting publication

- Deshmukh, S.R., Kirkham, J.J., Karantana, A, on behalf of the SO-HANDI study group. Developing a core outcome set for hand fractures and joint injuries in adults: an international Delphi survey and consensus meeting.

In preparation

- Deshmukh, S.R., Davis, T., Kerr, N., Marson, B.A., Leighton, P., Montgomery, A.A., Karantana, A. Incidence and anatomical distribution of adult hand and wrist fractures presenting to a UK Major Trauma Centre.
- Deshmukh, S.R., Leighton, P., Montgomery, A.A., Karantana, A. Experiences of injury, treatment and recovery of patients with hand fractures and joint injuries.

Presentations

International

- Deshmukh, S.R., Mousoulis, C., Marson, B.A., Grindlay, D., Leighton, P., Montgomery, A.A., Karantana, A. on behalf of the Core Outcome Set for Hand Fractures and Joint Injuries Study Group, 2021. Developing a core outcome set for hand fractures and joint injuries in adults: a systematic review of treatment outcomes.

  Oral presentation at FESSH-EFSHT 2020 Congress
- Deshmukh, S.R., Leighton, P., Montgomery, A.A., Kirkham, J.J., Trickett, R., Jerosch-Herold, C., Rodrigues, J.N., Griffin, X., Costa, M., Karantana, A. Developing a core outcome set for hand fractures and joint injuries in adults.

  Oral presentation at IFSSH, IFSHT & FESSH 2022 Combined Congress

National

- Deshmukh, S.R., Leighton, P., Montgomery, A.A., Trickett, R., Jerosch-Herold, C., Rodrigues, J.N., Griffin, X., Costa, M., Karantana, A. Developing a core outcome set for hand fractures and joint injuries in adults: a Delphi consensus study involving key stakeholders.

  Oral presentation at BSSH Autumn 2021 Scientific Meeting
- Deshmukh, S.R., Leighton, P., Montgomery, A.A., Trickett, R., Jerosch-Herold, C., Rodrigues, J.N., Karantana, A. Developing a core outcome set for hand fractures and joint injuries in adults: a qualitative study on the patient perspective.

  Oral presentation at BSSH Spring 2021 Scientific Meeting

Regional

- Deshmukh, S.R., Mousoulis, C., Marson, B.A., Grindlay, D., Leighton, P., Montgomery, A.A., Karantana, A. on behalf of the Core Outcome Set for Hand Fractures and Joint Injuries Study Group, 2021. Developing a core outcome set for hand fractures and joint injuries in adults: a systematic review of treatment outcomes.
  Oral presentation at East Midlands Orthopaedic Research Meeting 2019

Prizes

November 2019	East Midlands Orthopaedic Research Meeting Prize – Proxime Accessit
April 2021	Meeting Best Free Paper Prize, British Society for Surgery of the Hand Spring 2021 Scientific (£500)
January 2022	Arbeitsgemeinschaft für Osteosynthesefragen United Kingdom & Ireland (AOUK&I) Research Prize 2021 (£1,000)

Funding

A grant of £8,801 was awarded by AOUK&I to support the qualitative work, Delphi consensus study and consensus meeting. The PhD studentship was funded by the British Society for Surgery of the Hand and the Centre for Evidence Based Hand Surgery (University of Nottingham).

Acknowledgments

I wish to thank Miss Alexia Karantana, Dr Paul Leighton and Professor Alan Montgomery for giving me the opportunity to undertake this PhD and for their enduring guidance and support throughout.

My thanks also to the various individuals who have supported my research work, including fellow researchers and clinical academics in the academic office and beyond, staff of the academic orthopaedic division, clinicians in the fracture clinic who helped screen potential patient participants, and the research participants themselves for the effort and time they contributed which has enabled this research.

I owe a huge thanks to my parents, Rohini and Rajiv, and my sister Sunita for their tremendous support particularly during my education and early career. I am grateful to them as well as my wider family during these years of PhD research.

Special thanks to my wife Ranjeet, for her support, reassurances and her huge efforts which afforded me the time needed to focus on this work. Finally, I want to thank my daughter Avani – born early on in my time out-of-programme, she frequently found a way to ensure I took some much-needed breaks. Through her eyes I have had the chance to see the world anew, giving me a fresh perspective on what truly matters in life.

Table of Contents

Chapter 5 Delphi consensus study 178

Index of Tables

Index of Figures

Index of Appendices

Declaration and Summary of collaborators of the Standardised Outcomes in HAND fractures and joint Injuries (SO-HANDI) project

I certify that the work presented in this thesis is the result of original research primarily conducted with myself in the lead role. However, I recognise that the research described is collaborative: my supervisors (Miss Alexia Karantana, Dr Paul Leighton and Professor Alan Montgomery) were involved in all steps of the project and there were a number of people who joined a formal collaborative group I established during the study, with various roles and contributions to the ongoing work. There were also some individuals which, although not part of the formal collaborative, contributed to aspects of the research. I wish to thank all those who contributed to the work presented in this thesis (Table 1), and in order to appropriately acknowledge their specific roles I have indicated which elements of the research they were involved in (Table 2) along with their degree of involvement.

For the remainder of this thesis I generally use the first-person singular voice in active sentences for consistency, except in cases where part of the work involved substantial input from one or more other people. For example, in the systematic review where a second reviewer undertook data extraction or some element of data analysis I have used the first-person plural voice.

Table 1 List of SO-HANDI collaborators

Name	Initials	Affiliation
British Association of Hand Therapists Clinical Evidence Committee	BAHT	British Association of Hand Therapists
Matthew L Costa	MLC	University of Oxford
Tim Davis	TD	Queen's Medical Centre, Nottingham
Timothy RC Davis	TRCD	Queen's Medical Centre, Nottingham
Sandeep R Deshmukh	SRD	University of Nottingham
Xavier L Griffin	XLG	Queen Mary University of London
Douglas Grindlay	DG	University of Nottingham
Adeel Ikram	AI	University of Nottingham
Christina Jerosch-Herold	CJ-H	University of East Anglia
Alexia Karantana	AK	University of Nottingham
Nick Kerr	NK	Queen's Medical Centre, Nottingham
Jamie Kirkham	JK	University of Manchester
Paul Leighton	PL	University of Nottingham
Ben A Marson	BAM	University of Nottingham
Alan A Montgomery	AAM	University of Nottingham
Christos Mousoulis	CM	University of Nottingham
Jeremy N Rodrigues	JNR	University of Oxford
Ryan Trickett	RT	University Hospital of Wales

Table 2 Summary of contributions of SO-HANDI collaborators

Task	Lead contribution	Major contribution	Supporting role
Obtaining funding	SRD, AK, PL, AAM		
Conceptualisation	SRD, AK, PL, AAM		
Writing-up chapters	SRD		AK, PL, AAM
Epidemiology work			
Designing protocol	SRD	AK	
Screening of PACS reports	SRD, NK, TD		AK
Data extraction	SRD, NK, TD		AK
Validation steps	SRD	AAM	AK
Data analysis and interpretation	SRD	AK, PL, AAM	
Systematic review			
Designing protocol	SRD	AK, PL, AAM	CJ-H, RT, JRN
Designing search strategy	SRD	DG	AK
Screening titles and abstracts	SRD, DG		AK
Screening full texts	SRD, DG		AK
Designing data extraction form	SRD		
Data extraction	SRD, BAM, CM		AK
Outcome reporting bias assessment	SRD, BAM, CM		AK
Data analysis and interpretation	SRD	BAM, CM,AK, PL, AAM	
Writing manuscript	SRD	AK, PL, AAM	BAM, CM, DG, CJ-H, RT, JRN
Qualitative interviews / focus groups			
Designing protocol	SRD	AK, PL, AAM	CJ-H, RT, JRN, XLG, MLC
Obtaining ethics approval	SRD		AK, PL, AAM
Recruiting participants	SRD		
Developing topic guide	SRD	PL	AK, AAM
Conducting interviews / focus groups	SRD	PL	AK, AAM
Transcription (external service engaged for some transcription)	SRD		
Data analysis and interpretation	SRD	PL	AK, AAM

Table 2 Summary of contributions of SO-HANDI collaborators (cont'd)

Task	Lead contribution	Major contribution	Supporting role
Delphi survey			
Designing protocol	SRD	AK, PL, AAM	CJ-H, RT, JRN, XLG, MLC
Synthesis of outcome domain longlist	SRD	AK, PL, AAM	
Stakeholder engagement to refine survey items & descriptor wording	SRD	AK, PL, AAM, BAHT	CJ-H, RT, JRN, XLG, MLC
Creating online questionnaire	SRD		
Obtaining ethics approval	SRD		AK, PL, AAM
Recruiting participants	SRD		AK, TRCD
Managing survey rounds: data collection, sending reminders	SRD		
Processing of participant suggestions for additional domains	SRD	AK, PL, AAM	
Data analysis and interpretation (between rounds and final)	SRD	AK, PL, AAM	
Consensus meeting			
Designing protocol	SRD	AK, PL, AAM, RT, JK	CJ-H, JRN, XLG, MLC
Obtaining ethics approval	SRD		AK, PL, AAM
Recruiting participants	SRD	AK, CJ-H, TRCD	PL, AAM
Recruiting Chairperson	SRD		AK, PL, AAM
Preparation of pre-meeting participant information pack	SRD		AK, PL, AAM, JK
Chairing meeting	JK	SRD	
Contemporaneous meeting data analysis (live poll creation, analysis)	SRD		AK, PL, AI, BAM, RT, CJ-H, JK
Facilitating small group sessions	AK, PL, AI, BAM, RT, CJ-H		SRD, JK
Transcription (external service engaged for majority)	SRD		
Final data analysis and interpretation*	SRD	AK, PL, AAM, JK	

* Subsequent involvement of collaborative group and several meeting participants in terms of feedback on manuscript submitted for publication

List of abbreviations

AOUK&I	Arbeitsgemeinschaft für Osteosynthesefragen United Kingdom & Ireland
BSSH	British Society for Surgery of the Hand
CEBHS	Centre for Evidence Based Hand Surgery
CENTRAL	Cochrane Central Register of Controlled Trial
CES-D	Center for Epidemiologic Studies Depression Scale
COMET	Core Outcome Measures in Effectiveness Trials
COS	Core outcome set
COSMIN	COnsensus-based Standards for the selection of health Measurement INstruments
COS-STAD	Core Outcome Set Standards for Development
DASH	Disabilities of the Arm, Shoulder and Hand questionnaire
DRF	Distal radius fracture
EuroQOL	European Quality of Life Scale
GSE	General Self-Efficacy Scale
IADL	Lawton Instrumental Activities of Daily Living Scale
ICF	International Classification of Functioning, Disability, and Health
ICHOM	International Consortium for Health Outcomes Measurement
IRAS	Integrated Research Application System
JLA	James Lind Alliance
MHQ	Michigan Hand outcomes Questionnaire
NICE	National Institute for Health and Clinical Excellence
NPV	Negative predictive value

OMERACT	Outcome Measures in Rheumatology (formerly Outcome Measures in Rheumatoid Arthritis Clinical Trials)
PACS	Picture archiving and communication systems
PHQ	Patient Health Questionnaire
PPI	Patient and public involvement
PPPC	Patient Perception of Patient Centeredness
PPV	Positive predictive value
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROM	Patient-reported outcome measure
PRWE	Patient-Rated Wrist Evaluation
PRWHE	Patient-Rated Wrist/Hand Evaluation
PSEQ	Pain Self-Efficacy Questionnaire
PSP	Priority Setting Partnership
QuickDASH	Abbreviated Disabilities of the Arm, Shoulder and Hand questionnaire
SF-12	12-Item Short Form Survey
SF-36	36-Item Short Form Survey
SO-HANDI	Standardised Outcomes in HAND fractures and joint Injuries
TAAQoL	TNO-AZN Adult Quality of Life questionnaire
UK	United Kingdom
USA	United States of America
WHO	World Health Organization

Chapter 1 Introduction

1.1 Overview

This chapter serves to provide an overview of:

- Hand fractures and joint injuries in terms of impact and prioritisation
- Outcomes in health and their measurement more generally
- Core outcome sets and the development process in general
- Previous or ongoing work related to standardised outcome selection for hand and wrist conditions.

This will provide context for the ensuing chapters which describe the work conducted in the Standardised Outcomes in HAND fractures and joint Injuries (SO-HANDI) project – to develop a core outcome set for hand fractures and joint injuries in adults.

1.2 Hand fractures and joint injuries

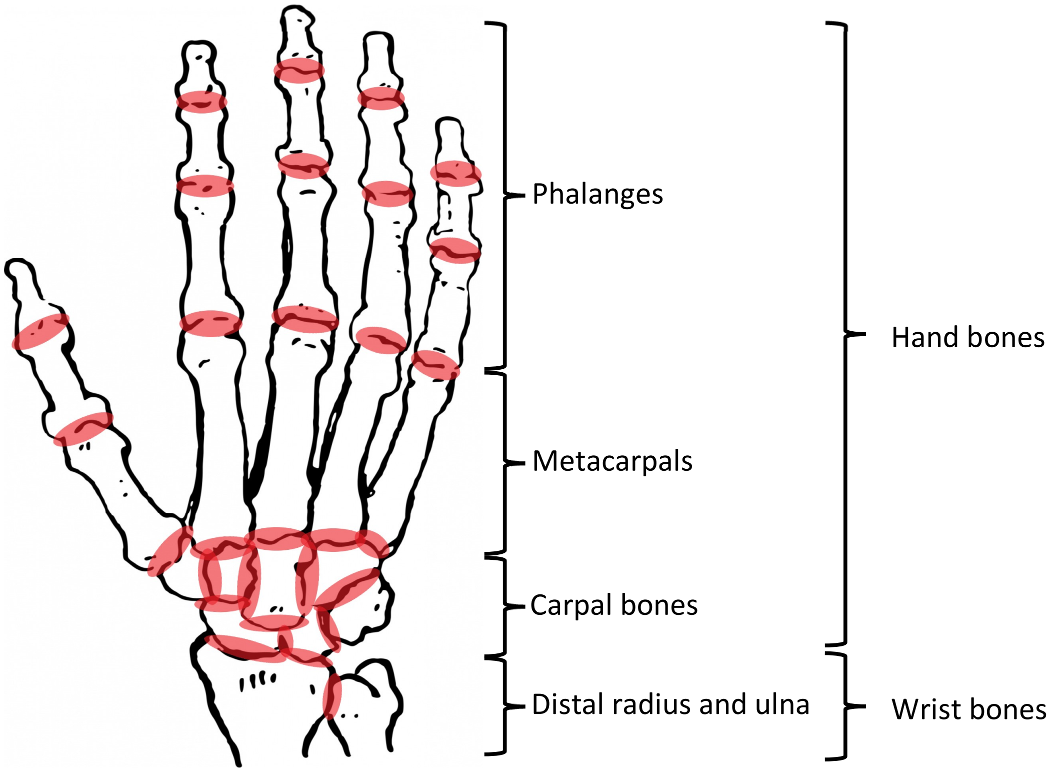
Hand fractures and joint injuries have significant impact on patients, healthcare resources and the economy due to restrictions on use of the hand for activities of daily living, work and leisure. Even six weeks after phalangeal or metacarpal fractures, some patients can struggle to perform routine tasks such as driving, writing or typing, which impacts on return to even office-based work (Smith et al. 1985). Patients who sustain wrist fractures report that alongside pain they also felt a significant psychological burden involving anxiety and shock, as well as financial distress from the prolonged period off work (Watson et al. 2018). Approximately one in five patients with distal radius fractures required admission to hospital in one study, indicating the potential burden of such injuries on specialist resources (O'Neill et al. 2001). Hand fractures and joint injuries are relatively common, so if a significant proportion need hospital admission this will in turn impact greatly on healthcare costs.

Even without admission the healthcare costs can be substantial. A Dutch-based study analysed their national Injury Surveillance System and estimated that metacarpal and phalangeal fractures accounted for approximately 6.3% of total costs of injuries presenting to Emergency Departments in the Netherlands (de Putter et al. 2012). This equated to US$278 million in 2007 and considered a combination of healthcare and productivity costs.

Despite the clear significance of such injuries at both a personal and a macroeconomic level, there typically is not a consensus on optimal treatment. Several reviews of the management of hand fractures and joint injuries highlight “inadequate outcome assessment” and “large variation in reported outcomes” (Handoll and Vaghela 2004; Poolman et al. 2006; Verver et al. 2017). Use of multiple, non-comparable outcomes across what would otherwise be considered ‘good quality studies’ hinders meta-analysis. This then impacts on the conclusions that can be drawn from the body of evidence for management of hand fractures and joint injuries. A core outcome set (COS) is an opportunity to improve the future evidence-base in the field of hand fractures and joint injuries by leading to consistent selection of outcomes, along with the nature of the outcomes themselves being more patient-centred.

For the purposes of this thesis, the term ‘hand fractures and joint injuries’ includes all fractures or breaks and joint injuries of the hand and wrist, as depicted in Figure 1.1. Setting the scope involved expert stakeholder consultation and is described in detail in Chapter 2.

Figure 1.1 Schematic diagram depicting bones of the hand and wrist. (Joints are shaded in red)



1.3 Prioritisation

The importance of further research into hand fractures and joint injuries is evidenced by the Health Technology Assessment Programme funding numerous trials in this area over the last 10 years alone, including the Distal Radius Acute Fracture Fixation Trial (1 & 2), the Scaphoid Waist Internal Fixation for Fractures Trial and POINT (a multi-centre randomised trial for management of proximal phalanx fractures) (National Institute for Health Research Health Technology Assessment 2022).

Through joint stakeholder work involving patients, carers and clinicians a recent James Lind Alliance (JLA) Priority Setting Partnership (PSP) on common conditions affecting the wrist and hand aimed to highlight research topics of most significance to all parties and with enduring ‘uncertainty’ (James Lind Alliance 2017). Although not focused solely on hand fractures and joint injuries, traumatic injuries to the hand were included within the scope of this PSP and in five of the ‘Top 10 Uncertainties’ (Figure 1.2).

The underlying goals amongst these five consisted of aiming to “improve results”, “enable the most efficient return to full function” or consider “functional outcome” or “benefit”. One ‘Uncertainty’ was entirely based on the premise of a patient-centred approach to outcomes being of great utility and focused on the outcome measurement aspect: “What methods are most accurate, user friendly and demonstrate the best clinical utility in measuring patient reported outcomes in common hand conditions?” (James Lind Alliance 2017)

Through the ‘Uncertainties’ described in the JLA PSP report, there is clearly recognition of the importance of defining outcomes for use in clinical studies in order to allow progress to be made in clinical research on hand fractures and joint injuries. Using a COS in future clinical study design would optimise progress in all the areas highlighted by the JLA PSP.

Figure 1.2 James Lind Alliance Priority Setting Partnership Top 10 Uncertainties of common conditions affecting the hand and wrist (not ordered)

(Adapted from James Lind Alliance, 2017)

- Can scar/fibrosis formation be manipulated to improve results following hand surgery/trauma?
- In patients with Dupuytren's disease, what invasive techniques give the best results in terms of function, recurrence and cost?
- In the treatment of common hand conditions, such as peripheral nerve compression syndromes, ganglia or arthritis of the fingers/thumb/wrist, do surgical interventions have a demonstrable benefit in patient reported outcome when compared with non-surgical methods or placebo surgery?
- Regarding patient and cost benefits, which interventions give the best results in the treatment of painful joints in the hand/wrist?
- What are the most effective non-surgical methods for treating early arthritis in the hand and fingers?
- What interventions/treatments will have the most positive effect following nerve injury?
- What methods are most accurate, user friendly and demonstrate the best clinical utility in measuring patient reported outcomes in common hand conditions?
- Which hand therapy techniques enable the most efficient return to full function following surgery or injury?
- Which hand/finger/thumb injuries would benefit from surgical intervention over hand therapy or no formal treatment, considering both functional outcome and societal cost?
- Which patients with acute ligament injuries to the wrist or chronic wrist/distal radio-ulnar joint instability benefit from surgical treatment rather than from non-surgical method?

1.4 Health outcomes

The outcomes or results of healthcare interventions have value in both research and clinical practice settings. The importance of observing the consequences of medical intervention has long been known, and the concept of evidence-based medicine has helped to focus clinicians' attention on the principle of combining the best available published evidence with their own clinical experience and expertise (Sackett et al. 1996; Djulbegovic and Guyatt 2017). Measuring health outcomes allows us to identify the impact that healthcare interventions have. This is necessary in order to make comparisons which can then help to shape best practice for specific health conditions.

Comparison between interventions can also seek to determine the best use of limited healthcare resources, in a form of cost-benefit analysis (Robinson 1993). Such analysis again requires an assessment of the benefits gained through healthcare intervention.

When measuring health outcomes, two basic aspects of the concept of an outcome are (Stucki et al. 2007):

- The domain being assessed – 'what' to measure
- The method being used to make the measurement – 'how' the specified domain is measured

In their analysis of ClinicalTrials.gov data subsequent to the trial registry being expanded to include the reporting of summary results, Zarin et al (2011) also drew a distinction between the outcome domain ('what' one is trying to measure), and that of the measurement instrument ('how' the domain is being measured).

The reasoning behind making this distinction is that 'what' to measure (the outcome domains thought to be relevant and important according to key stakeholders such as patients and healthcare professionals) is likely to be stable over time unless the underlying health condition changes in some material way (Stucki et al. 2007). 'How' to measure various outcome domains is something that has demonstrably changed over time and is quite likely to do so in future. Therefore, a strategy which first determines 'what' to measure and then on 'how' to measure it allows the former to be the foundation which the latter can build upon in an evolving manner. A brief history of the concept of evolving outcome measures is described in the next section.

1.4.1 Measuring health outcomes

This thesis will focus on the 'what' element of outcomes. However, a brief introduction to the 'how to measure' aspect is warranted. There are many forms of measure (Garratt et al. 2002):

- Dimension-specific measures focus on a specific aspect/domain of health
- Generic measures are not population-specific and typically measure several health domains
- Disease or population-specific measures consider aspects of health relevant to particular health problems, again covering multiple domains
- Individualised measure – where each patient determines the priority of different aspects of their life and then a measure is made of how their health condition has impacted them
- Utility measures for economic evaluation that incorporate 'health states' of patients

There is an argument that for a long time a relatively simple measure, that of mortality rates, was the only key measure of the health of populations. However, for many parts of the world mortality rates have not been considered a sensitive enough measure of the change in individual or population health states for several decades (Bergner 1985). Although mortality remains an important measure, for many populations death rates are sufficiently low that they are afforded the option of significantly expanding beyond that basic premise of health to also consider ‘morbidity’. Measuring morbidity is a vastly more complex issue than mortality, not least because the former has a much more amorphous nature. Any symptom or detriment to one’s natural health state can be considered morbidity, and it is these things which health interventions typically aim to resolve or alleviate.

Traditional measures of health have involved outcomes such as physiology (through laboratory test) and clinician observation (at times with the use of formalised scales/scoring systems to improve reliability) but there has been a growing recognition of the value of the perspective of patients in both clinical practice and research (McDowell 2006; Weldring and Smith 2013). This has progressed from asking about patient satisfaction with healthcare to patient-reported outcome measures (PROMs) which are usually questionnaires that attempt to elicit patients’ views on a variety of domains such as function, pain, symptom control, quality of life, etc (Black 2013).

With the use of health measurements beyond simpler measures of inarguable outcomes (such as death rates) towards clinician observation scales and PROMs, there is an awareness that the properties of more complex instruments need to be assessed for reliability, validity and responsiveness (Mokkink et al. 2009). The Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN) Initiative is an international team of researchers who have developed a checklist to guide such evaluations (Mokkink et al. 2010). This is a substantial body of work in its own right, and would need to be conducted for measurement instruments relating to all the outcome domains selected in the ‘what to measure’ phase of core outcome set development. Indeed, if one or more domains are not sufficiently measured by existing instruments then the ‘how to measure’ phase could involve development of a novel instrument. This explains why ‘how’ to measure the outcome domains has been deemed beyond the scope of this thesis, as it is a critical element of work which must not be rushed and it is not pragmatic to expect to complete this alongside the ‘what to measure’ phase within a doctoral thesis timeframe.

1.4.2 Classification systems for outcome domains in health

There is a myriad of health conditions, with many additional factors that may influence the choice of outcomes to be measured. Tackling the complexity of outcome domains in health could be approached by trying to consider the different aspects of health. The following is a brief summary of some of the more commonly referenced classification systems, presented chronologically.

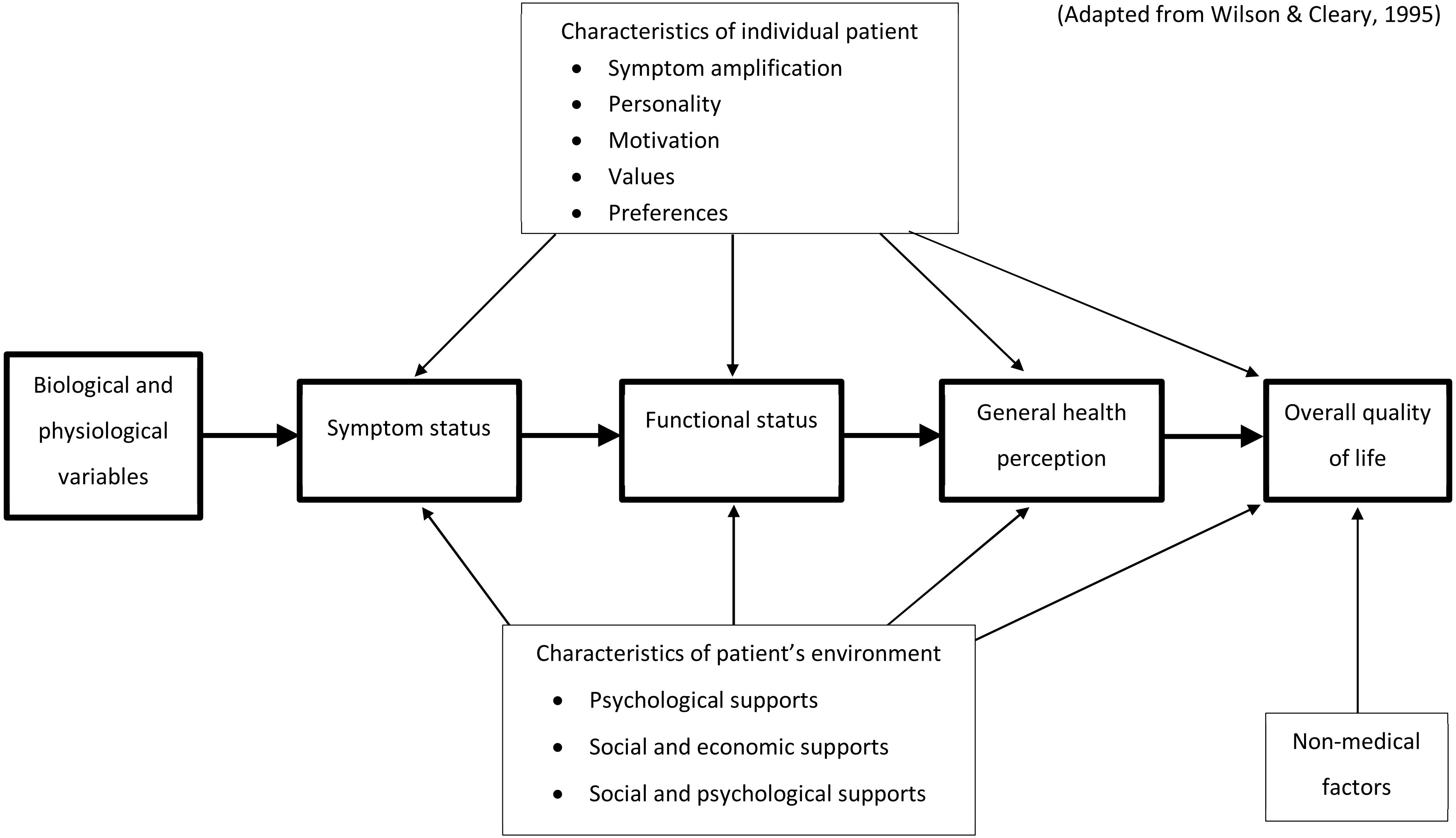
1.4.2.1 *Wilson and Cleary Model of Health-Related Quality of Life* (Wilson and Cleary 1995)

In this model, Wilson and Cleary describe a taxonomy of patient outcomes categorised according to underlying concepts they represent and also the relationships between different concepts (Wilson and Cleary 1995). They consider measures of health as, “existing on a continuum of increasing biological, social, and psychological complexity”. This continuum has five levels, with inputs formed by characteristics of an individual patient and that of the environment or supports that are available to that patient as summarised in Figure 1.3. Wilson and Cleary (1995) considered these five levels as being on a continuum of increasing complexity:

- Biological and physiological variables – focusing on the function of cells, organs and organ systems
- Symptom status – focusing on the patient as a whole, it is the “patient’s perception of an abnormal physical, emotional or cognitive state”
- Functional status – the ability of the patient to perform particular defined tasks in various domains (e.g. physical, social and psychological functions)
- General health perception – an overall evaluation which integrates the previous levels and is subjective in nature
- Overall quality of life – subjective well-being related to the patient’s happiness or satisfaction with their life as a whole

The model was revised by Ferrans et al in 2005, but the essence of the five-level continuum was retained (Ferrans et al. 2005).

Figure 1.3 Original Wilson and Cleary health-related quality of life conceptual model for measures of patient outcome



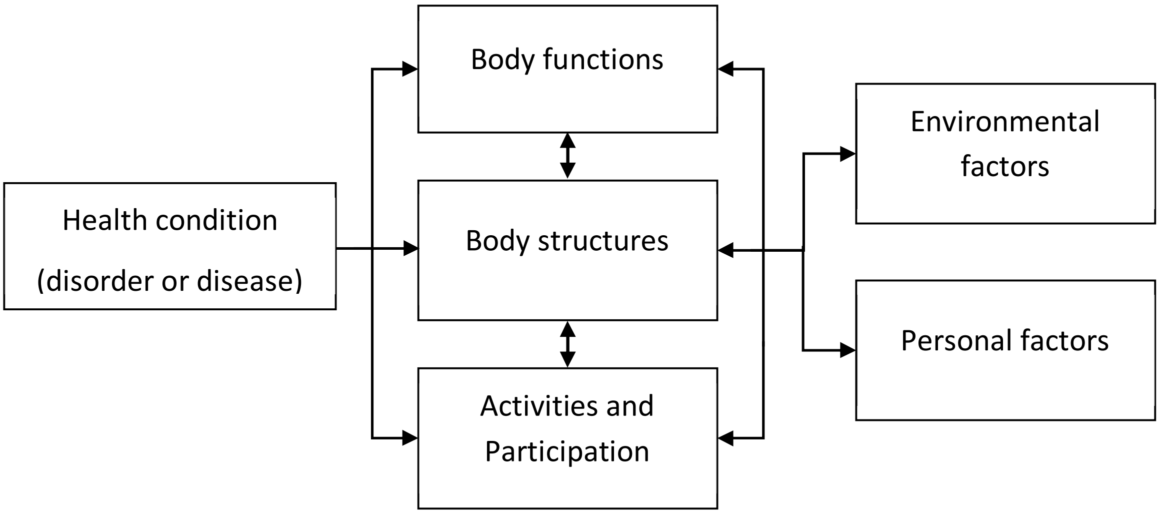
(Adapted from Wilson & Cleary, 1995)

1.4.2.2 World Health Organization International Classification of Functioning, Disability and Health (World Health Organization 2001)

The World Health Organization (WHO) International Classification of Functioning, Disability and Health (ICF) is a generalised framework for health and health-related domains within the four 'components' of 'body functions and structures', 'activities and participation', 'environmental factors' and 'personal factors' (World Health Organization 2001). Although most of these have a highly granularised classification involving a hierarchy of domains within the WHO ICF, 'personal factors' has not yet been classified (World Health Organization 2002).

The hierarchy is structured slightly differently to the framework, as outcome domains of the body functions and body structures components are considered separately in the classification while the 'activities and participation' component is kept intact. The summary of the classification shown in Figure 1.4 reflects this.

Figure 1.4 International Classification of Functioning, Disability and Health Classification



1.4.2.3 Outcome Measures in Rheumatology (OMERACT) Filter 2.0 (Boers et al., 2014)

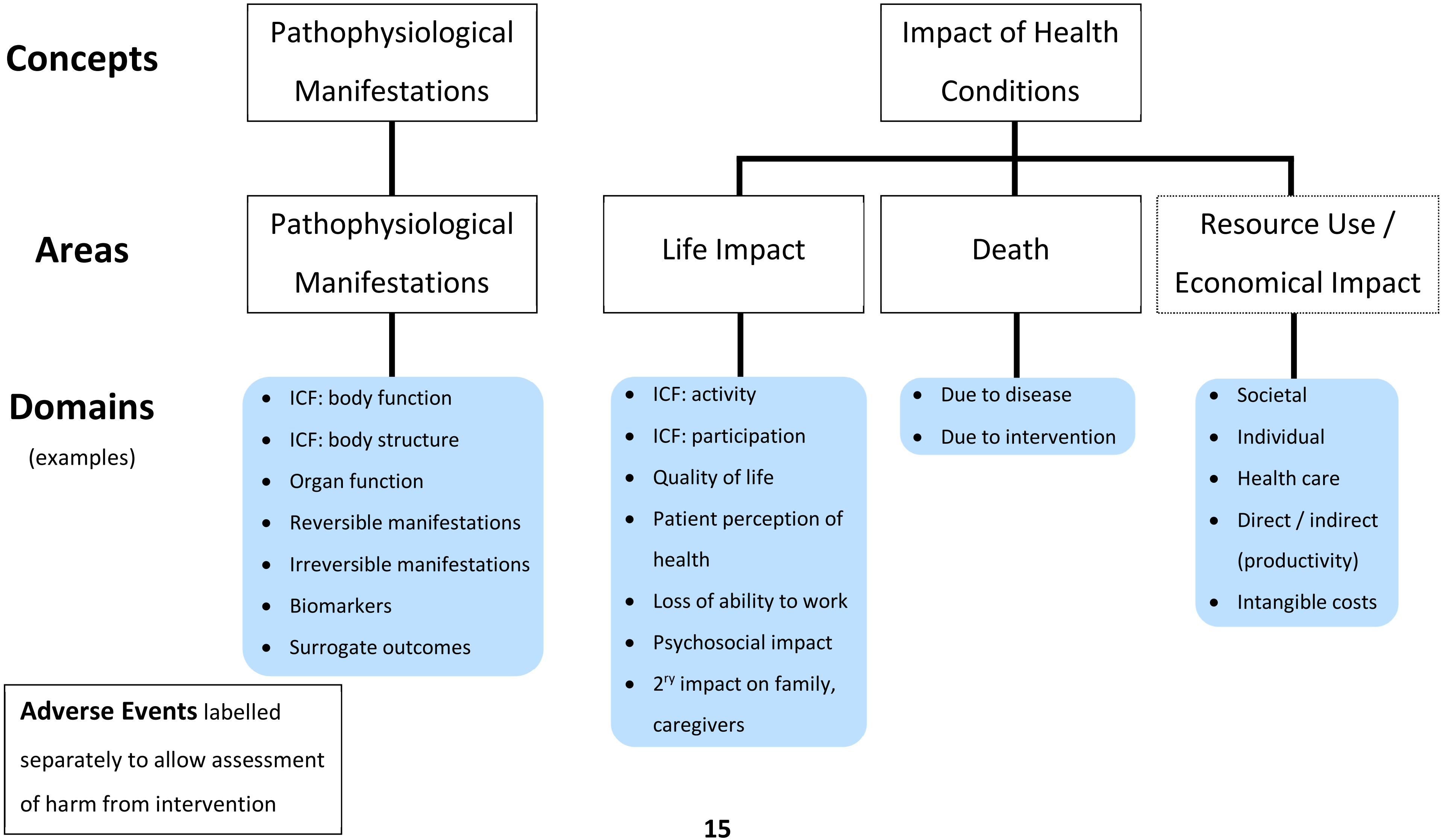
The Outcome Measures in Rheumatology (OMERACT) initiative is an international collaboration which had its first conference in 1992 (Clarke 2007). OMERACT had an overall goal of achieving some form of common approach to the outcomes being assessed or measured in clinical trials in rheumatology but the underlying principles they have developed can be applied much more broadly. The philosophy of OMERACT was summarised in the original OMERACT Filter, which considered facets of 'Truth' (accuracy of measurement), 'Discrimination' (encapsulating reliability and sensitivity to change) and 'Feasibility' (essentially the pragmatism of measurement) (Boers et al., 1998). This Filter focused more on the concepts of outcome validity in the sense of how outcomes are measured rather than what domains are covered.

OMERACT subsequently reviewed several outcome frameworks to determine suitability for use in standardising outcome measurement across clinical trials (Idzerda et al. 2014). This review, where they explored frameworks including the Wilson and Cleary model and WHO ICF framework, informed the development of their own framework which is contained within OMERACT Filter 2.0. Here, they outline four 'areas' that they believed encompassed all possible health outcomes for use in clinical trials (Boers et al. 2014). This comprised of three core areas including 'Death', 'Life Impact' and 'Pathophysiological Manifestations', and a strongly recommended area of 'Resource Use'. OMERACT has related their own framework to the WHO ICF components, as indicated in Figure 1.5.

Filter 2.0 has recently been updated, with the concept of ‘Adverse Events’ being more integrated into the model with the domains more broadly categorised as ‘intended benefits’ or ‘harms’ (Boers et al. 2019).

Figure 1.5 OMERACT Filter 2.0 Conceptual Framework of Outcome Measurement in Health Intervention Studies

(Adapted from Boers et al., 2014)



1.4.2.4 Williamson / Clarke revised taxonomy for outcomes in medical research

(Dodd et al. 2018)

Having reviewed numerous “health research vocabularies”, the authors of this taxonomy believed that there was insufficient comprehensiveness and granularity. They developed their revised taxonomy as an iterative process, based on core areas which were similar to other classifications. This includes mortality/survival, physiological/clinical, functioning, resource use and adverse events/effects. However they then elaborated upon this by specifying a further level to the taxonomy, which lead to 38 outcome domains (Dodd et al. 2018). The authors encourage further subdivision of these outcome domains by researchers where necessary, and several of the outcome domains within the Williamson / Clarke revised taxonomy do appear to encompass quite broad concepts, e.g. ‘musculoskeletal and connective tissue outcomes’ or ‘general outcomes’ within the physiological/clinical core area.

1.5 Core outcome sets and health outcome selection in research

As the brief summary of classification systems of health outcomes demonstrates, there is a vast number of different outcome domains that can be used to measure health. This applies whether measuring health outcomes for clinical practice or for research. Despite the drive towards evidence-based medicine, it has become increasingly recognised that a lack of harmony in research efforts will hinder progress for a number of reasons.

1.5.1 Issues related to outcome selection and reporting in clinical research

1.5.1.1 Heterogeneous outcome selection

There are many factors involved in selection of outcomes for any given study, and this has tended to lead to variation in the outcomes eventually selected across different studies (Williamson et al. 2012). This can take various forms, such as outcomes differing in terms of underlying meaning or in the way in which they are measured. Such inconsistency can lead to difficulty in drawing comparisons between studies even if other key factors such as the participant population and interventions are comparable (Clarke 2007).

1.5.1.2 Outcome reporting bias

Outcome reporting bias is when only a subset of originally stated outcomes of research is actually reported, selected based on knowledge of results (Kirkham et al. 2010). There is empirical evidence to suggest that statistically significant outcomes are more likely to be fully reported (Dwan et al. 2008). This has consequences not only for the research work concerned but also any systematic review or meta-analysis that attempts to examine the non-reported outcome. For example, an outcome many not be reported because analysis showed statistical non-significance. However, if this is a repeated pattern then the clinical trials input into a meta-analysis will tend to be those where a statistically significant finding occurred and the final conclusion has reduced validity. In their study on this issue, Kirkham et al (2010) examined a cohort of Cochrane reviews and found that of 42 meta-analyses where a statistically significant result was reported, "eight (19%) were non-significant after adjustment for outcome reporting bias and 11 (26%) had overestimated the treatment effect by ≥20%".

1.5.1.3 Research waste

Health research involves patient participants and funding, but there is an issue of avoidable waste in research though poor design, conduct or reporting (Chalmers and Glasziou 2009). There are thus many different facets to the concept of research waste, but in relation to health outcomes this revolves around both heterogeneous outcome selection and outcome reporting bias. Both these issues can lead to clinical trial results not contributing to meta-analyses (Yordanov et al. 2018). There is a further dimension to research waste though – even if outcomes are chosen homogeneously and fully reported, if they are not clinically relevant and do not include the priorities of patients then the conclusions drawn may be of suboptimal value (Ioannidis et al. 2014).

1.5.2 Core outcome sets

A core outcome set (COS) is a minimum set of outcomes that should be assessed/measured in all clinical trials for a specified condition or group of conditions (Clarke 2007). Use of the COS does not preclude researchers from setting additional outcomes for their studies, nor indeed are any of the outcomes that comprise a COS required to be the primary outcome of a clinical study. It would simply mean that any clinical study in that field would at least collect outcome data on the agreed core outcomes.

The concept of a COS has developed over several years as a response to the various outcomes-related issues summarised in the previous section. Standardising outcome selection would facilitate meta-analysis and potentially reduce outcome reporting bias because a deficiency of reporting of the COS would be conspicuous by its absence (Kirkham et al. 2013). By enabling meta-analysis and pooling of results from more studies along with a positive impact on outcome reporting bias, research costs/wastage overall could theoretically be reduced (Ioannidis et al. 2014; Yordanov et al. 2018).

In recognition of their importance, COSs are now endorsed by major funders of clinical research (e.g. National Institute for Health Research, Versus Arthritis, Association of Medical Research Charities, Health Research Board, Horizon 2020) (COMET Initiative 2022c), trialists (Standard Protocol Items: Recommendations for Interventional Trials guidelines 2022), Cochrane (Cochrane Community Blog 2022), NICE (National Institute for Health and Care Excellence (Great Britain) 2015), some journal editors (e.g. Core Outcomes in Women's and Newborn Health Initiative 2019) and patients and the public (COMET People and Patient Participation, Involvement and Engagement working group (COMET Initiative 2022b).

1.5.2.1 Early history of core outcome set development

One of the earliest organised attempts to tackle these issues was by the OMERACT Initiative (Clarke 2007). With its first conference in 1992, OMERACT's goals included reaching a consensus on the minimum number of outcome measures to be used in all rheumatoid arthritis trials (Tugwell et al. 2007). Although the first meeting did not have input from patients, there have been conferences to refine the initially agreed core outcome measures every 2 years and the perspective of patients has since been sought (Kirwan et al. 2003).

The 'what to measure' and 'how to measure' aspects of health outcomes applies for a COS as well, as it is simply a standardised set of outcomes for a given health condition. Through its iterative consensus approach towards a COS for rheumatoid arthritis, OMERACT has worked to define both these aspects (Stucki et al. 2007).

The COSMIN Initiative was founded in 2005 and focused on improving the selection of outcome measurement instruments of health outcomes (Li et al. 2015; COSMIN Initiative 2022). COSMIN have developed definitions of the properties of measurement instruments and tools to assess these properties in existing and newly developed instruments. This aims to provide key information that should improve the ability to select the most appropriate outcome instrument for a given outcome domain when designing studies.

The Harmonising Outcome Measures for Eczema (HOME) Initiative was founded in 2008, bringing together key stakeholders in the eczema research community including patients (and parents/carers), healthcare professionals, the pharmaceutical industry, health regulatory authorities and journal editors (HOME Initiative 2022). In developing a COS for atopic eczema, HOME also began by initially identifying which outcome domains should be considered ‘core’ (Schmitt et al. 2010). They then tackled the issue of how to measure the identified domains in a stepwise approach (Schmitt et al. 2014; Spuls et al. 2017).

1.5.2.2 The Core Outcome Measures in Effectiveness Trials Initiative

The Core Outcome Measures in Effectiveness Trials (COMET) Initiative had its inaugural meeting in 2010, bringing together people with an interest in developing, applying and promoting COSs (Williamson et al. 2012). Their first meeting included triallists, systematic reviewers, patients, clinicians, journal editors, funders, policy-makers and regulators (Williamson et al. 2017). The specific aims of COMET are (Tunis et al. 2016):

1. Raising awareness of current problems with outcomes in clinical trials
2. Encouraging COS development and uptake
3. Promoting patient and public involvement in COS development
4. Providing resources to facilitate these aims
5. Avoiding unnecessary duplication of effort
6. Encouraging evidence-based COS development.

COMET has created a repository of COS development-related studies. This includes those directly involving COS development and those which could inform subsequent COS development (COMET Initiative 2022a). Database and website usage figures suggest that awareness and interest in COMET and perhaps more generally COSs is growing (Gargon et al. 2017).

1.5.3 Standards of core outcome set development

There is currently no established ‘gold standard’ for COS development, with many COSs developed through various methods (Gorst et al. 2016). However, consensus work has been done to agree on some standards by which COSs should be developed – the Core Outcome Set-Standards for Development (COS-STAD) (Kirkham et al. 2017). These recommendations highlight the importance of involving key stakeholders, including patients and/or their representatives. Historically, researchers have chosen the outcomes used in clinical studies and were likely to be influenced by the outcomes selected in previous studies. This could lead to study results which are not meaningful to patients or other stakeholders. By taking a patient-centred approach in developing a COS, the tendency for outcomes to be chosen based simply on convention can be ameliorated.

COMET has laid out some of the key issues that COS developers should consider, including aspects such as setting the scope, identifying existing knowledge, involving stakeholders and consensus methods to be considered (Williamson et al. 2012). Both HOME and OMERACT have issued guidance based on their respective experiences, with HOME producing a ‘roadmap’ of the steps to take (Schmitt et al. 2014) while OMERACT has developed a handbook (OMERACT 2022). COMET has since produced a handbook that provides guidance and recommendations for COS developers (Williamson et al. 2017). There is a great deal of similarity in the broad steps described in this and the OMERACT and HOME documents. According to the COMET handbook the key steps in COS development are (Williamson et al. 2017):

- Step 1: Define COS scope
- Step 2: Check whether a new COS is required

- Step 3: Develop a protocol for determining what to measure
- Step 4: Determine what to measure
- Step 5: Determine how to measure the COS

COMET has collaborated with COSMIN to produce a practical guideline for Step 5 (Prinsen et al. 2016). However, this thesis focuses on ‘what’ to measure and therefore the work done reflects the first four of the steps outlined above.

1.6 Relevant previous and ongoing work

Having determined the intended scope of a COS, COMET recommends that Step 2 is to check whether a new COS is required (Williamson et al. 2017). This has two elements: one is to determine whether a relevant COS is already available, to avoid duplication of effort; the second is to gauge whether a COS is needed (i.e. does an issue of inconsistent outcome selection, measurement or reporting exist). An indication of the need for a COS for hand fractures and joint injuries was briefly touched upon in statements from some of the systematic reviews in this field in Section 1.2. However, the systematic review of treatment outcomes detailed in Chapter 3 aids in quantifying the issue.

The COMET database of existing and ongoing COS development projects is perhaps the most comprehensive COS database available and is updated annually (COMET Initiative 2022a). As part of the preliminary work for this project, the COMET database was checked to see if a relevant COS already existed. I also conducted scoping searches of the literature. The following is a summary of the findings.

1.6.1 **Core set of domains for distal radius fractures** (Goldhahn et al. 2014)

In this study, Goldhahn et al (2014) conducted a literature review and then consensus workshops to arrive at a shortlist of core domains relating to distal radius fractures. The literature review summarised outcomes and outcome instruments used in published research and the information was conveyed to workshop participants. Two sets of group discussion were held with international participants, one in Munich in 2009 and the second in Zurich in 2011. There were ten participants in the first round, with each participant contributing their five most important outcomes for distal radius fracture patients. A survey based on the generated items was completed by 13 respondents, and informed the second meeting. This meeting involved 16 participants, five of whom were participants in the first meeting. A consensus was then reached on nine outcome domains.

Patients were included in the discussions, which involved stakeholder groups of clinicians, physiotherapists, methodologists, epidemiologists, researchers and industry representatives. Unfortunately there is no detailed description of the methodology which clarifies how many patient representatives were present – given there were only 21 unique participants across the two meetings it is not clear that the patient perspective would have been sufficiently elicited. Certainly no in-depth qualitative work to explore the patient perspective appears to have informed the consensus meetings. It is quite possible that this COS is skewed towards the professional perspective rather than that of patients. It would certainly lack some of the more detailed understanding of reasons that patients raise certain outcome domains when interviewed or during focus group discussion. The SO-HANDI COS involves extensive qualitative work with patients to contribute towards a longlist of outcome domains that then undergoes a consensus process to refine to a shortlist and then the eventual COS. Finally, the core set of domains for distal radius fractures has, by its nature, a much narrower scope than the one established for the SO-HANDI COS. This is not necessarily a criticism of the Goldhahn et al (2014) COS, as a COS can be extremely focused in terms of the scope of the health condition covered, or quite broad. However, it is a distinguishing feature between their COS and ours, and one could argue that there is greater utility in a COS that has broader application. Of course, the broader COS needs to be valid for the wider scope of injuries. Appropriate qualitative work could help establish whether outcome domains raised by patients with the broader range of injuries show sufficient overlap to have a unified COS.

1.6.2 World Health Organization Comprehensive and Brief ICF Core Sets for Hand Conditions (Rudolf et al. 2012)

In Section 1.4.2.2 the WHO International Classification of Functioning, Disability and Health was briefly described. This framework has been used to develop ways in which the effectiveness of healthcare interventions can be monitored, including the ICF Core Sets for Hand Conditions (Rudolf et al. 2012). These have been developed sequentially, with first the Comprehensive ICF Core Set (117 outcome categories) and then further work to narrow down to the Brief ICF Core Set (23 outcome categories).

There was first a preparatory phase to generate a pool of potential ICF categories to include in the final Core Sets. This phase consisted of four parts (Rudolf et al. 2012):

1. Focus groups involving 59 patients with hand conditions to elicit the patients' views on important aspects of functioning, environment and personal factors. The statements of patients were translated to the WHO ICF domains.
2. International online survey of 162 healthcare professionals working in the field of hand conditions, again linking responses to the WHO ICF domains
3. Systematic literature review on outcomes of studies on patients with hand conditions which were published in 2003-2008
4. Cross-sectional study of hospitals and rehabilitation centres in Germany, where clinicians rated the functioning and health of 210 patients with hand conditions using the classification system of the ICF

Based on information from the preparatory phase, 23 professional stakeholders comprising of an international panel of physicians, occupations therapists, physiotherapists, nurses, social workers and psychologists determined the ICF Core Sets for Hand Conditions through a three-day consensus conference. They selected ICF categories for the Comprehensive ICF Core Set first, and then some categories from this to form the Brief ICF Core Set.

The purpose of a COS is to represent the minimum outcomes to be measured and reported in clinical studies as a matter of consensus priority rather than being comprehensive, with researchers retaining the freedom to have additional outcomes of specific relevance to a particular study (Clarke and Williamson 2016). Therefore, the Brief ICF Core Set is the one with closest comparison to the SO-HANDI COS, in that it aims to be a minimal standard to describe functioning for any patient with any hand condition.

The Brief ICF Core Set covers an extremely broad scope of health conditions including, “disease or hand injuries originating external to the hand but affecting the hand (such as rheumatoid arthritis, stroke, Parkinson’s disease, or brachial plexus injuries)” (Kus et al. 2011). This correspondingly seems to have resulted in a longer list of core outcomes than would be anticipated from a typical COS. The 23 outcome domains of the Brief ICF Core Set would make its use as a COS impractical. Certainly the broad scope results in some domains that would not seem to apply to a COS dedicated to hand fractures and joint injuries (e.g. ‘spinal cord and related structures’) (Rudolf et al. 2012).

Importantly, although the development process follows the methodological steps specified for WHO ICF Core Set development (Selb et al. 2015) it did not incorporate direct patient involvement when selecting the outcome domains to be included in the ICF Core Sets (Rudolf et al. 2012). Patient participation was limited to the preparatory phase which only informed development of the Comprehensive ICF Core Set. The Brief ICF Core Set was similarly decided by non-patient participants. The Brief ICF Core Set in particular may lack some outcome domains that patients might deem very important, as it is a set of outcome domains filtered by non-patients from a longlist (the Comprehensive ICF Core Set) that was already filtered by non-patients. Indeed this is borne out by the ICF Core Set developers attempting to validate the Brief ICF Core Set with patients and in conclusion recommending that seven additional ICF outcome domains are added to complement the Brief ICF Core Set (Kus et al. 2012). COMET advocates patient participation in development precisely because otherwise important outcome domains may be neglected (Williamson et al. 2017).

Even the Brief ICF Core Set then has 30 outcome domains recommended for assessment in patients with hand conditions, which is at odds with the principle of a core outcome set.

1.6.3 International Consortium for Health Outcomes Measurement standard set for hand and wrist conditions (Wouters et al. 2021)

The International Consortium for Health Outcomes Measurement (ICHOM) group originates from the United States of America (USA). ICHOM has developed a standard set for hand and wrist conditions. The ICHOM 'standard set' is different to a COS and neither precludes the other. Although somewhat narrower than that of the ICF Core Sets, in covering all hand and wrist conditions the ICHOM standard set scope is very broad. It is therefore presumably less specific and sensitive for a specific health area such as hand fractures and joint injuries. Probably in recognition of the cumbersome nature of such a breadth of scope, the ICHOM working group opted to divide the standard set into five 'tracks': thumb, finger, wrist, nerve and severe hand and/or forearm trauma. ICHOM explored and reached a consensus on an outcome set per track.

There are several differences beyond the scope of health conditions between the SO-HANDI COS and the ICHOM standard set. The former has selected key outcomes to be used in clinical research. This differs from the ICHOM standard set which is targeted towards use by healthcare providers in more clinical settings. Perhaps most importantly though, the SO-HANDI COS is underpinned by extensive qualitative work with patients as per COMET methodological recommendations (Williamson et al. 2017); whereas ICHOM involved patients only through a survey after outcome domains were selected by the working group. That selection was from a longlist of domains developed by the working group through systematic review of the recent literature on management of hand and wrist conditions. The working group consisted of 22 hand surgeons, hand therapists and researchers, and therefore no patients had any role in development nor selection of the outcome domains. The patient survey was to check which domains patients felt were a priority, but did not appear to then incorporate these views by making any changes to the working group's selected domains.

The ICHOM working group proceeded with selection and even generation of outcome measurement instruments to match the selected domains, but this work is beyond the scope of the SO-HANDI COS and therefore will not be considered in detail in this thesis.

1.6.4 Miscellaneous works

By checking the COMET database and discussing with the COMET project co-ordinator I was able to find other registered projects with ongoing or completed development of standardised outcome sets for hand conditions including:

- Thumb carpometacarpal osteoarthritis (Copeland et al. 2020; Hoang-Kim et al. 2022)
- Hand osteoarthritis more generally (Kloppenburg et al. 2014)
- Scars of the hand (Kennedy et al. 2022)
- Dupuytren’s disease (Ball et al. 2013)

These are for conditions of the hand outside the scope of this COS for hand fractures and joint injuries.

Having ascertained that a COS with the planned scope would be a novel contribution, I registered the SO-HANDI project with the COMET Initiative (Deshmukh et al. 2022) in October 2018. This included an outline of the project as well as details of project collaborators from several parts of the United Kingdom (UK) who formed the broader SO-HANDI study group.

1.7 Summary

Hand fractures and joint injuries are common, with impact not only to affected patients but also on wider society through healthcare resource use and lost productivity. Despite the clear significance of such injuries at both a personal and a macroeconomic level, there is typically not a consensus on optimal treatment. The variety of treatment modalities, and lack of consistency in outcome reporting and research methodology standards, makes it challenging to interpret the available evidence. Reviews of the management of hand fractures and joint injuries have commented on issues such as “inadequate outcome assessment” and “large variation in reported outcomes” (Handoll and Vaghela 2004; Poolman et al. 2006; Verver et al. 2017).

Over the last few decades there has been a growing recognition of the challenges posed by lack of outcome standardisation in healthcare research more generally (Clarke 2007). There are now numerous COS development groups working towards standardising outcome measurement in different health conditions as shown by the growing COMET database of COS development projects (Gargon et al. 2019b).

1.8 Thesis aim

To establish what outcome domains should be measured in a core outcome set on hand fractures and joint injuries in adults.

Objectives

- Chapter 2
  - Define the scope of the COS with particular focus on the scope of injuries to be covered, with exploration through expert stakeholder (hand surgeon and therapist) discussions
  - Determine the incidence of fractures relevant to the COS scope through analysis of reports of all hand and wrist radiographs done over a one-year period at a UK Major Trauma Centre, including a validation step and drawing comparisons with existing literature
- Chapter 3
  - Identify the outcomes prioritised from the healthcare professional/researcher perspectives by systematic review of the recently published clinical research in this field. Verbatim outcomes are first categorised into unique outcomes, which are then categorised into outcome domains using the World Health Organization International Classification of Functioning, Disability and Health
  - Assess outcome reporting bias in the recently published clinical research in this field
- Chapter 4
  - Explore the perspective of patients who have experienced hand fractures and joint injuries in terms of the injury itself, its treatment and outcomes which are

  important to them, through one-to-one semi-structured interview and focus group methods and thematic analysis to develop outcome domains of relevance to patients
  - Consider variation between patients who have sustained distal radius fractures and those who sustained other forms of hand fractures and joint injuries

- Chapter 5
  - Conduct a three-round international Delphi consensus study of key stakeholders (patients, hand surgeons and hand therapists) to process the longlist of outcome domains synthesised from the findings of the work described in Chapters Chapter 3 and Chapter 4. This achieved initial consensus prioritisation of the longlist of outcome domains and informed a final consensus meeting
  - Consider and explore general aspects of Delphi methodology

- Chapter 6
  - Synthesise the different strands of data and reach final consensus on a COS through a consensus meeting involving all key stakeholders
  - Develop further insight about the rationale underlying participants' prioritisation of outcome domains by analysis of transcripts of meeting discussions

Chapter 2 Scope of core outcome set and incidence of hand fractures and joint injuries

In this chapter I briefly summarise some of the exploratory work done at the outset to define the scope of the SO-HANDI COS, as well as a study at a large tertiary care UK major trauma centre which adds to the wider epidemiological literature on hand and wrist fractures. The latter helps to demonstrate the extent of these fractures and also provided guidance towards purposive sampling in the subsequent qualitative study (Chapter 4).

2.1 Defining the scope of the COS on hand fractures and joint injuries

Step 1 of the COMET guidance on COS development involves defining the scope of the COS (Williamson et al. 2017).

2.1.1 Defining the population, interventions and settings of use for the COS

Defining the scope of a COS is an important early step in the process of COS development as highlighted in the Core Outcome Set Standards for Development (COS-STAD) recommendations (Kirkham et al. 2017). Within the domain of 'scope' the authors of COS-STAD state the following should be clarified with regards to a COS:

1. Research or practice setting(s) in which it is to be applied

2. Health condition(s) covered
3. Population(s) covered
4. Intervention(s) covered (Kirkham et al. 2017 p.4)

I developed the scope of this COS in-line with these recommendations.

This thesis is focused on the determination of what outcomes should be measured/assessed – these are applicable to both research and practice settings. In COS development, questions persist about a potential need for a difference in 'how' to measure the outcomes depending on the specific setting. Therefore, while the complete COS taking into account both the 'what' and 'how' aspects may be more selective in applicability, the current project focused on determining core outcome domains and arguably does not share this restriction.

2.1.2 Exploratory work on defining the health conditions for the COS

Defining the scope of the COS in terms of the health condition(s) was challenging. I intended for this COS to cover hand fractures and joint injuries; this did not have an established definition at the commencement of the thesis. To explore this, I conducted informal one-to-one discussions with a range of clinical experts in the management of these injuries. This included face-to-face or telephone discussions with hand therapists and hand surgeons from across the UK, with a mix of professional backgrounds to capture a broad set of perspectives (Table 2-1).

The discussions began with delivery of background information on core outcomes sets and the intent to develop one for hand fractures and joint injuries. Discussion focused on two areas: the anatomical scope of the hand and the range of injuries to be covered by the COS (focused on the term 'joint injuries').

Table 2-1 Professional background of experts involved in scoping discussions

Characteristic		Number (n = 12)
Clinical role	Hand surgeon	6
	Hand therapist	6
Surgeon training background	Orthopaedic surgery	5
	Plastic surgery	1
Therapist training background	Occupational therapy	2
	Physiotherapy	4

2.1.2.1 Anatomical scope of the hand

There were a range of views amongst the experts in terms of the anatomical limits of the hand. A few drew a clear anatomical distinction between the hand and wrist, such as the metacarpal/carpal boundary, the boundary between the two rows of carpal bones, or the radiocarpal joint. However, when challenged as to the reasoning behind their chosen boundary of the hand or when presented with alternative boundaries, the experts did not have a clear rationale for their choice.

The aim of this line of enquiry was to set the scope, in particular whether the radiocarpal joint (and subsequently the distal radius and/or ulna) should be included. The prevailing view was that excluding fractures in the distal radius and ulna, due to these being so closely linked in function to the carpus as part of the wrist joint.

The final anatomical scope included all the bones and joints as shown previously in Figure 1.1.

2.1.2.2 Injury scope for 'joint injuries'

There were differing opinions on the types of injuries deemed 'joint injuries' that should be included. General consensus was that certain injury types and patterns – primary nerve injuries, traumatic amputations and the multi-injured 'mangled hand' – should be excluded from the scope as their treatment and recovery do not align with less severe hand injuries. Some felt there was not much difference between tendon injuries and ligamentous injuries at a joint. Further discussion centred on isolated tendon injuries away from a joint (e.g. pure lacerations of flexor/extensor tendons); it was concluded that primary tendon injuries of an open wound nature, not directly localised to a joint, should not be included within the scope of this COS.

2.1.3 Finalised scope of COS

For the purposes of this COS, hand fractures and joint injuries therefore include the following:

- Phalangeal fracture(s)
- Metacarpal fracture(s)
- Carpal fracture(s) (scaphoid, lunate, triquetral, pisiform, trapezium, trapezoid, capitate, hamate)
- Distal radius and/or distal ulna fracture
- Any injury with physical damage localised to a joint between the bones listed above, including dislocation, subluxation, volar plate injury, avulsion injury, ligamentous tears/sprains/ruptures and closed tendon ruptures/tears

The scope includes the treatment of acute injuries and chronic problems such as malunion and non-union.

In terms of the population, this COS is for adults. It is intended to cover all treatment interventions pertaining to the management of hand fractures and joint injuries. It does not cover prophylaxis or prevention.

Primary nerve injuries, burns and open tendinous injuries are excluded from the scope of this COS. Complex hand injuries (e.g. mangled hand, amputations requiring replantation) are outside the scope of the COS. However, a COS developed from this work could form part of the range of outcomes selected for a clinical study on these types of complex hand injuries.

2.2 Incidence of hand fractures and joint injuries

2.2.1 Introduction

Hand fractures and joint injuries are common (O’Neill et al. 2001; van Onselen et al. 2003). They are important to patients and due to various economic and health-care resource consequences they impact on wider society (O’Neill et al. 2001; de Putter et al. 2012; Watson et al. 2018).

Epidemiological studies of these injuries lack comprehensiveness in their coverage. They generally report the data for either fractures of distal radius or of carpus/metacarpus/phalanx in their given study population. They often involve small or non-representative populations with arbitrary age restrictions (O’Neill et al. 2001; Thompson et al. 2004; Court-Brown et al. 2014) or a combination of paediatric and adult patients (van Onselen et al. 2003; Laugharne et al. 2013). Several lacked detailed breakdown of the injuries such as specific anatomical location (Feehan and Sheps 2006; Court-Brown et al. 2014; Karl et al. 2015) or calculation of incidences (Hove 1993; van Onselen et al. 2003; Laugharne et al. 2013). All these factors reduce accuracy and meaningfulness of the reported incidence of specific hand fractures. Therefore, a more detailed study on fractures pertaining to this COS in an adult UK population would be of value and could help inform future decisions on allocation of healthcare resources and research design.

2.2.2 Aim

To describe the incidence and anatomical distribution of all fractures of the finger rays, carpus, and distal radius/ulna in adults (aged ≥18) presenting to a UK major trauma centre, through analysis of all radiograph reports over a one-year period which stated the presence of a fracture of these bones.

2.2.3 Methods

I identified all patients with fractures presenting to a large tertiary care UK major trauma centre in an urban area (Nottingham University Hospitals NHS Trust) over 12 months, from 1st August 2016 and 31st July 2017. All reports from the picture archiving and communication systems (PACS) for all radiographs taken of the hand or wrist during this period were identified and reviewed This included patients referred for radiographs from primary care, the emergency department, local minor injury units and fracture clinic.

Fractures were focused upon, as data collection to determine the incidence of ‘joint injuries’ is extremely challenging for a number of reasons. ‘Joint injuries’ is considered an umbrella term for a number of pathologies including dislocations and subluxations (with either soft tissue injury or associated fractures), avulsion fractures and intra-articular fractures (Freiberg, 2007). However, it can also include purely tendinous (e.g. mallet finger, acute Boutonniere’s) and ligamentous (e.g. collateral ligament rupture) injuries. As a readily accessible data source, PACS reports would be expected to reliably capture fractures including joint injuries involving a fracture, but not other forms of joint injury (some of which may not even present to hospital, such as a patient with a finger joint subluxation that spontaneously reduced).

The recorded injuries were categorised into four groups to aid analysis:

- Isolated fracture involving distal radius and/or ulna
- Isolated fracture involving phalanx/metacarpus
- Isolated fracture involving carpus
- Multiple fractures (involving any combination of above bones)

Radiographs taken at the first point of patient contact are formally reported by either a suitably trained radiographer or radiologist, attached to corresponding radiographs and stored in the central PACS system. These reports were screened to identify those where injury according to any of the groups listed above were reported. On occasions where no report was available, the radiographs were reviewed directly by a reviewer and the radiograph event (an occasion when plain x-ray investigation was undertaken) was categorised as ‘fracture’ or ‘no fracture’ accordingly.

Data extracted from reports with eligible fractures included: specific bone(s) fractured; zone of injury within the fractured bone(s) (i.e. distal, shaft, proximal); whether multiple fractures were identified in a given injury event; laterality of injury.

Demographic data including patient age at time of radiograph and sex were recorded. Population estimates for calculation of incidences used the Office for National Statistics 2016 population data (Office for National Statistics - UK Statistics Authority 2016); 95% confidence intervals for age- and sex-specific incidences are Poisson confidence intervals, given the relatively low frequency of injuries compared to population figures for each age band. For demographic data analysis, the unit of analysis was the individual patient: in cases where a patient had two or more separate fracture events over the data collection period, their age and sex data at first presentation was analysed. For all other analysis including age- and sex-specific incidences, the unit used was the fracture event: patients with multiple fractures in a single fracture event counted just once but those with two or more separate fracture events over the study period counted more than once.

Screening of PACS reports was by one of two reviewers initially, with a validation step in which a third reviewer checked a random sample of 100 reports screened as reporting 'no fracture' and 200 reports screened as reporting 'fracture'. In the case of errors being found, all reports were to be screened again but by the other of the initial two reviewers (i.e. there would effectively be independent two reviewer screening of all reports). In such a scenario, any discrepancy between reviewers would be checked by SRD and resolved. The final step in this screening stage would then repeat the validation check of a random sample of 100 reports screened as reporting 'no fracture' and 200 reports screened as reporting 'fracture'.

Further validation involved assessing a random sample of ~5% of the original radiographs to determine accuracy of PACS reporting of fractures.

This work was performed as part of service evaluation with the host NHS Trust, with the aim to assist in identifying the anatomical distribution and incidence of these injuries; no ethics approval was required.

2.2.4 Results

Between 01/08/2016 and 31/07/2017 there were 13,941 radiograph events involving the hand or wrist in 13,491 patients. No report was available for 670/13,941 (4.8%) of radiograph events, so radiographs for these were reviewed directly.

There were a total of 3,925 fracture events (an instance of one or more hand or wrist fractures being reported) across 3,886 patients. Several patients had concurrent fractures (more than one fracture in a given fracture event), while some had two or more separate instances of fracture events over the study period.

2.2.4.1 Demographic characteristics and fracture incidence

There were 2,145 fracture events sustained by men, and 1,780 by women. Median age of the overall patient group was 44 years (IQR 27 to 65). Table 2-2 shows the summary of age and sex distribution of fracture events amongst the four fracture groups based on patients' first fractures within the study period.

Table 2-2 Demographic characteristics of patients according to injury group (first fracture only)

Characteristic	Fracture group				
	Single distal radius +/- ulna (n = 1290)	**Single phalanx or metacarpus (n = 2007)**	**Single carpus (n = 300)**	**Multiple fractures (n = 289)**	**All groups combined (n = 3886)**
Age in years Median (upper, lower quartiles)	62.7 (46.5, 77.9)	33.9 (23.9, 50.7)	38.0 (24.4, 60.4)	42.4 (26.0, 64.6)	44.0 (27.0, 65.0)
Sex					
Male	349 (27%)	1375 (69%)	204 (68%)	194 (67%)	2122 (55%)
Female	941 (73%)	632 (31%)	96 (32%)	95 (33%)	1764 (45%)

There were an estimated 632,100 adults (age ≥18 years) in Nottingham and its immediate surrounding Districts, comprised of approximately 311,700 men and 320,400 women (Office for National Statistics - UK Statistics Authority 2016). This meant an overall crude hand or wrist fracture incidence of 62 per 10,000 adults per year. The crude incidences by injury group for men, women and the total adult population are summarised in Table 2-3.

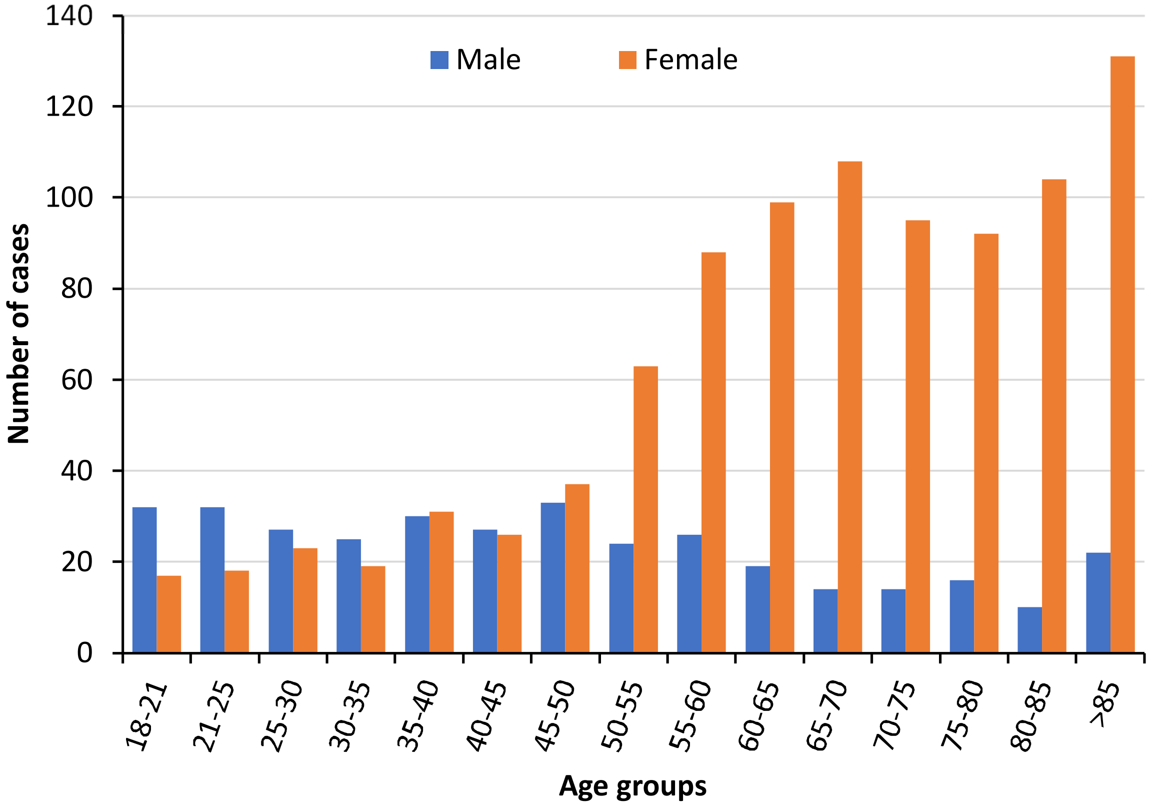
Table 2-3 Crude annual incidences of fractures by injury group and sex

Crude annual incidence	**Fracture group**				
	Single distal radius +/- ulna	**Single phalanx or metacarpus**	**Single carpus**	**Multiple fractures**	**All groups combined**
Rate per 10,000 men per year	11	45	6.6	6.4	69
Rate per 10,000 women per year	30	20	3.0	3.0	56
Rate per 10,000 adults per year	21	32	4.8	4.7	62

Figures 2.1 to 2.4 illustrate a detailed breakdown of age- and sex-specific distribution across the four fracture groups.

Figure 2.1 Age- and sex-specific (a) numbers of cases and (b) incidences with 95% confidence intervals for isolated distal radius +/- ulna fracture

(a)



(b)

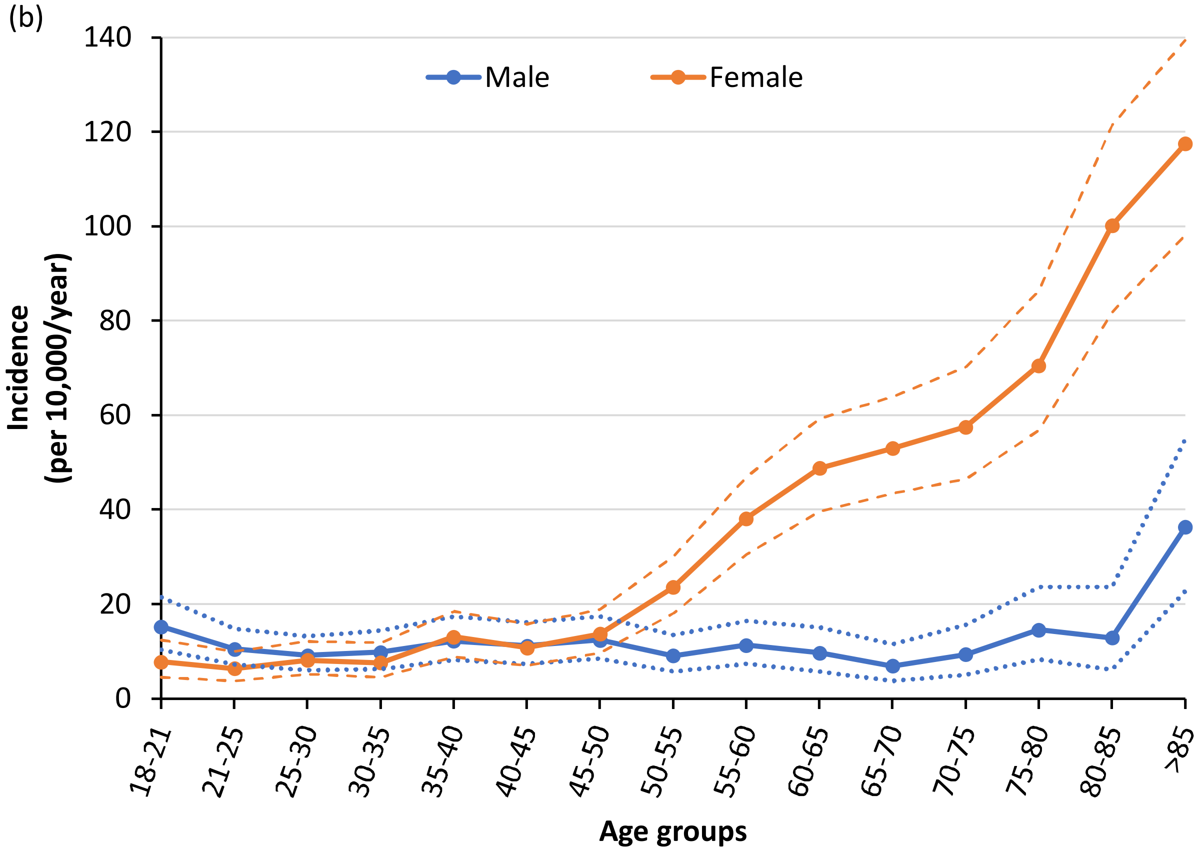
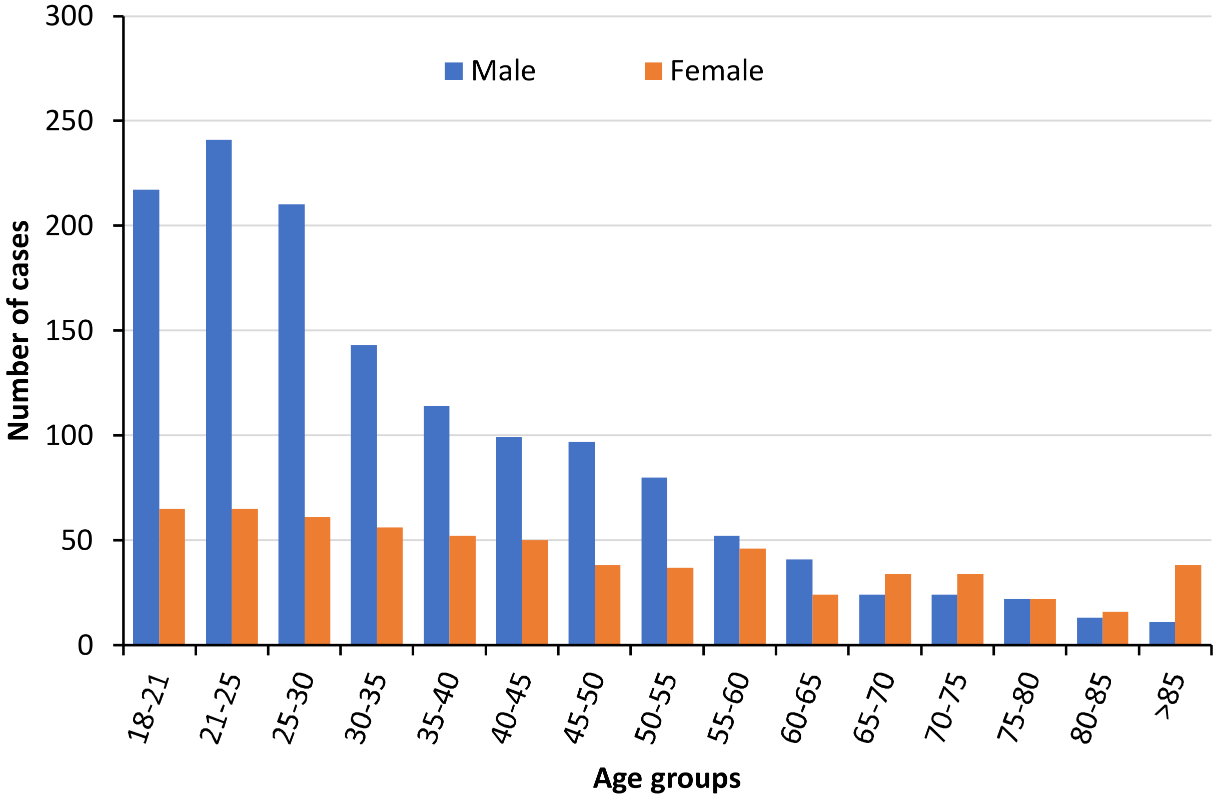


Figure 2.2 Age- and sex-specific (a) numbers of cases and (b) incidences with 95% confidence intervals for isolated phalanx/metacarpus fractures

(a)



(b)

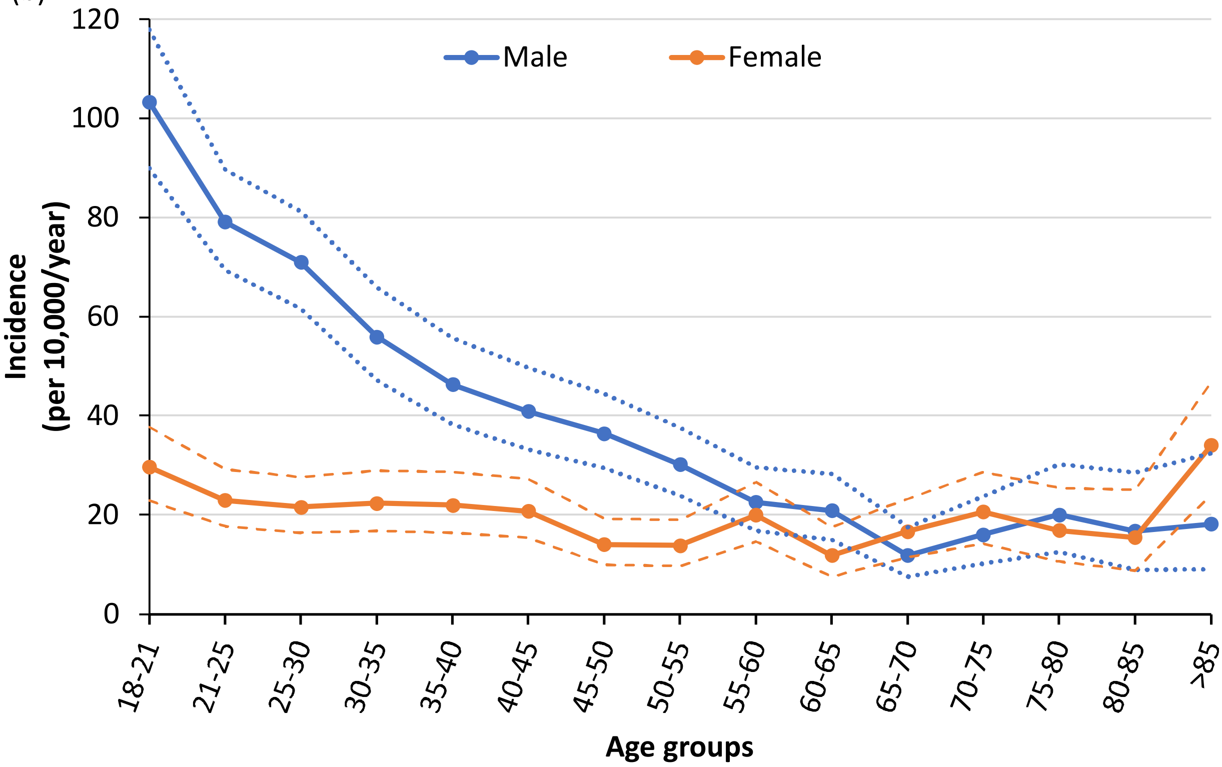


Figure 2.3 Age- and sex-specific (a) numbers of cases and (b) incidences with 95% confidence intervals for isolated carpus fracture

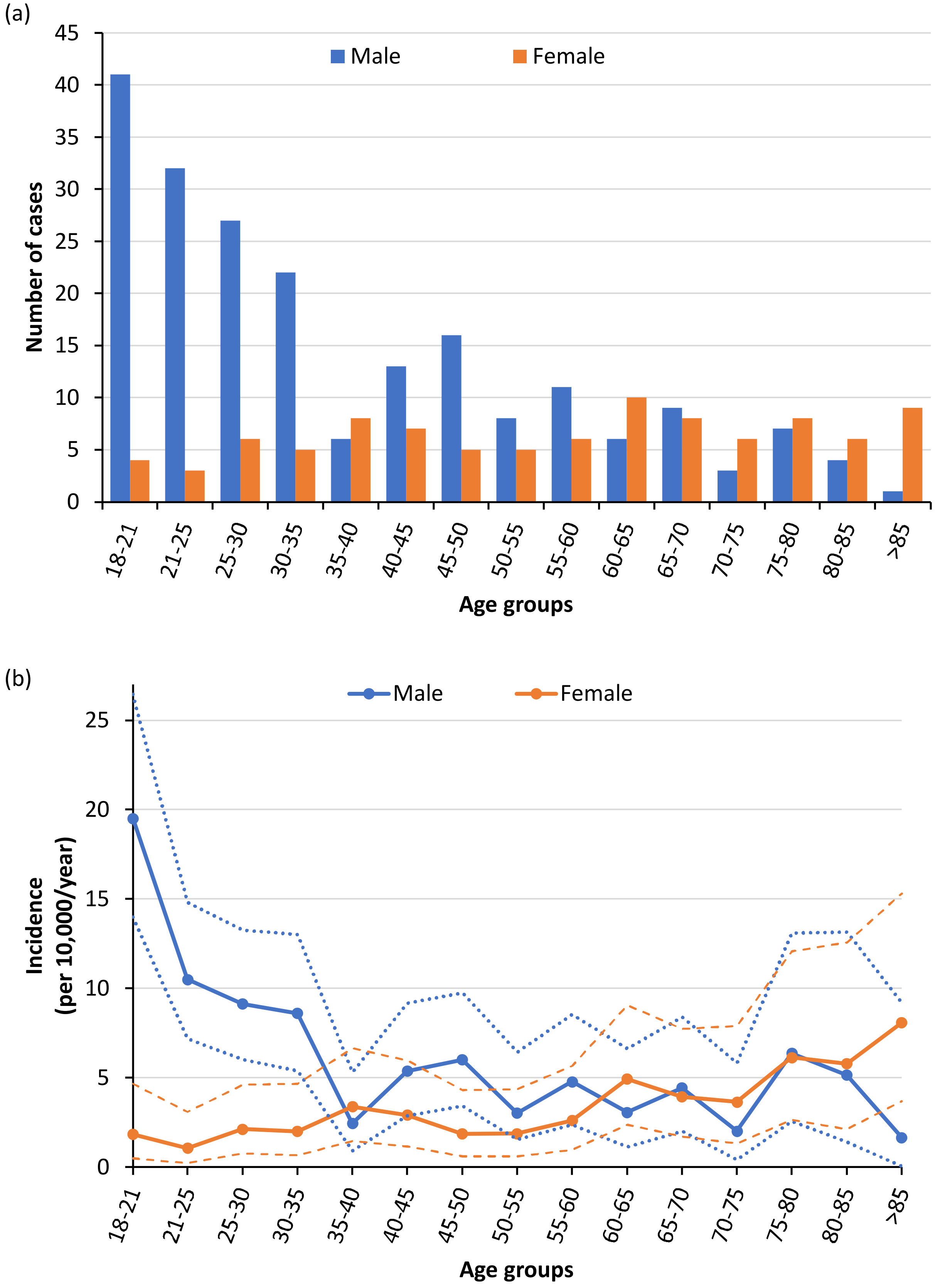
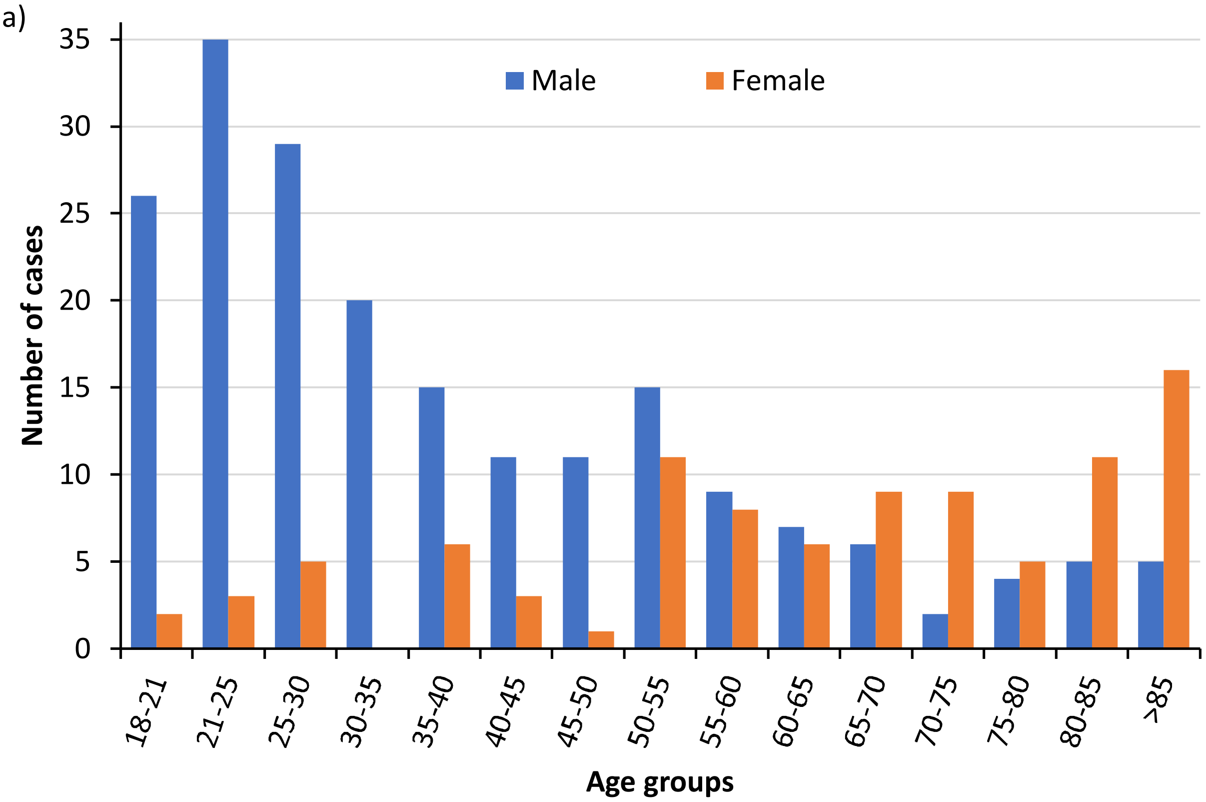
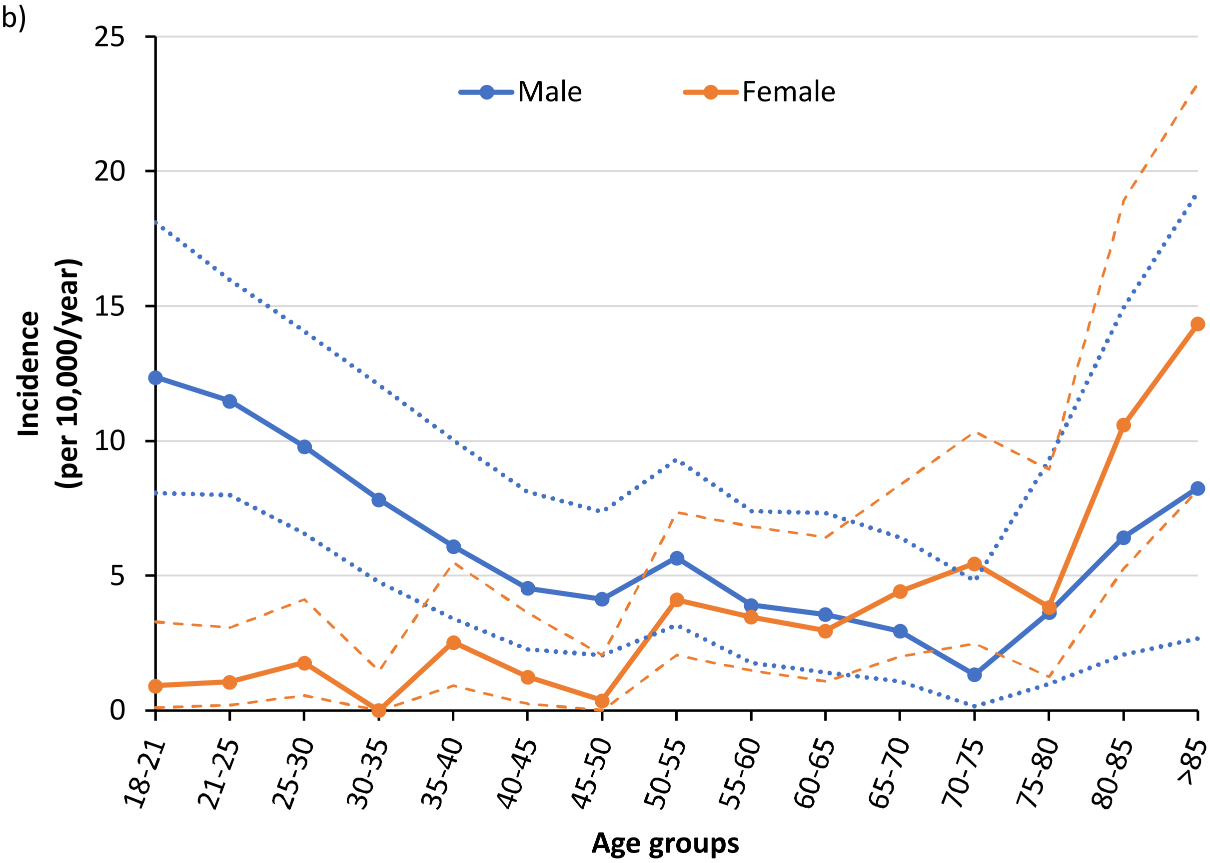


Figure 2.4 Age- and sex-specific (a) numbers of cases and (b) incidences with 95% confidence intervals for multiple fractures

(a)



(b)



2.2.4.2 Single distal radius and/or ulna fracture

There were 1,302 distal radius and/or ulna fractures. Seven patients had two separate fracture events involving distal radius and/or ulna fracture. Just over 60% (786/1,302) were reported as isolated distal radius fractures, while 35% (462/1,302) were a combination of distal radius and distal ulna fractures. Only 4.1% (54/1,302) were isolated distal ulna fractures. For cases where the distal ulna was reported as fractured, the ulnar styloid was involved for 84% (435/516).

2.2.4.3 Single phalangeal/metacarpal fracture

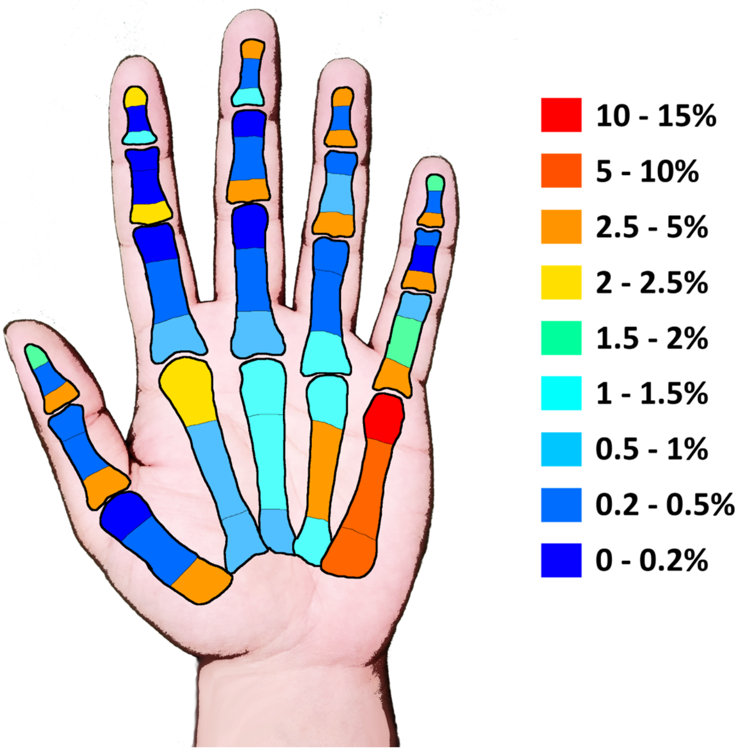
There were 2,026 isolated phalanx or metacarpal fractures, and three thumb sesamoid fractures. Ten patients had two separate instances of isolated phalanx or metacarpal fracture and one had three such instances. Table 2-4 summarises the fractures according to hand ray and specific bones. Over 40% of phalangeal/metacarpal fractures were of the little finger ray, with 25% (505/2,026) to the metacarpal alone.

Table 2-4 Frequency of single fractures to the bones of the hand

	Distal phalanx	**Middle phalanx**	**Proximal phalanx**	**Metacarpus**	**Sesamoid**	**Total number (% of total)**
Thumb	121	N/A	72	57	3	253 (12%)
Index	82	43	29	76	N/A	230 (11%)
Middle	114	80	22	58	N/A	274 (14%)
Ring	134	110	44	103	N/A	391 (19%)
Little	131	90	152	505	N/A	878 (43%)
Total number (% of total)	582 (29%)	323 (16%)	319 (16%)	799 (40%)	3 (0.15%)	**2026 (100%)**

Figure 2.5 shows a heatmap of the distribution of fractures across the hand and by zones of each hand bone.

Figure 2.5 Heatmap of the distribution of phalangeal and metacarpal fractures



The percentages shown were based on the total number of isolated phalangeal and metacarpal fractures where the zone of bone which was injured had been specified. This was the case for 92% (1,861/2,026) of the fractures.

2.2.4.4 Single carpus fracture

There were 302 carpus fractures (Table 2-5). The scaphoid was fractured in 61% (185/302) of cases while the second commonest carpal fracture was of the triquetral at 26% (80/302). The latter were typically dorsal avulsion type injuries.

Table 2-5 Frequency of single fractures to the carpal bones

Specific carpus fractured	**Frequency**
Scaphoid	185
Triquetral	80
Hamate	28
Pisiform	4
Trapezoid	3
Trapezium	2
Total	**302**

2.2.4.5 Multiple fractures

Multiple concurrent fractures within single fracture events were the most challenging to describe. Many different combinations of injury occurred. There were 295 fracture events involving multiple fractures in 293 patients, with two bones fractured in 265/295 (90%) of fracture events. Further combinations of fracture regions (i.e. hand, carpus, distal radius +/- ulna) are provided in Appendix 1.

2.2.4.6 Validation

Accuracy of radiograph report screening

Initial check of the accuracy of report screening was performed as follows:

- Review of 100 randomly selected cases where a screening reviewer had identified a fracture in the report. In all 100 there was indeed a fracture reported.

- Review of 200 randomly selected cases where the screening reviewer had not found a fracture in the report. There were seven cases where a fracture had been reported.
- Therefore proceeded as described earlier, with double screening and a subsequent repeated check afterwards. The final screening accuracy was 100%, with all 100 cases of ‘fracture seen in report’ indeed having a fracture reported, and all 200 cases of ‘no fracture seen in report’ having no fracture reported.

Accuracy of radiograph report content

A separate validation check was performed on the accuracy of the reporting content. We assessed radiographs from a random sample of 700/13,941 (~5%) of ‘radiograph events’ and compared the radiographic finding to the study report (Table 2-6). PACS reports had a sensitivity of 98% (190/194) and a specificity of 98% (498/506). The positive predictive value (PPV) of a PACS report stating presence of a fracture was 96% (190/198) while the negative predictive value (NPV) was 99% (498/502).

Table 2-6 Validation check of PACS reports through review of original radiographs

		Direct radiograph review findings	
		Fracture present	**No fracture present**
PACS report	**Fracture present**	190	8
	No fracture present	4	498

2.2.5 Discussion

2.2.5.1 Key findings

This study summarises the incidence and anatomical distribution of all fractures of the finger rays, carpus, and distal radius and ulna in adults (aged ≥18) in a UK population. This bridges the gap between breadth of coverage and detail in existing epidemiological studies. This is of value in guiding resource allocation for service provision and when planning and prioritising future research.

The relatively low numbers of cases of single carpal and multiple fractures lead to wide 95% confidence intervals for those fracture groups; nevertheless, the incidence of these injuries was higher in young men than in young women (Figures 2.3 and 2.4). This was also the case for phalangeal and metacarpal fractures, though it was not until age 55-60 years before men and women had overlapping incidences (Figure 2.2). Finally, for isolated distal radius and/or ulna fractures men and women had similar incidences until approximately the 50-55 year age band, after which women's continued to increase with age (Figure 2.1).

Isolated distal radius and/or ulna fractures were sustained almost three times as often in women as in men, in contrast to the circa 2:1 male:female ratio for patients sustaining an isolated hand fracture, a carpal fracture or multiple fractures. Most cases of distal forearm fracture involved older women, while all other fracture groups comprised mainly younger men. The difference in median age between the distal forearm fracture group and other fracture groups was 21 years or more.

The commonest fractured bone was the distal radius, and second commonest was the little finger metacarpal.

Validation on the accuracy of reporting of radiographs found high sensitivity and specificity of fracture reporting at 98%. The PPV of 96% and NPV of 99% suggest a slight overcalling of fractures.

2.2.5.2 Findings in relation to existing literature

Direct comparison with other studies is difficult due to differences in population demographics. Typically, studies have focused on only a subset of the fractures covered in this chapter.

The pattern of phalangeal fractures being more common than metacarpal was consistent with previous studies from Canada and Norway (Hove 1993; Feehan and Sheps 2006). The incidence of phalanx or metacarpal fractures was in keeping with the 36 per 10,000 per year of the Canadian study (Feehan and Sheps 2006). A large, multiple-state USA study reported on incidences of phalangeal and metacarpal fractures, though reported these in 15-year age brackets (Karl et al. 2015). Nevertheless, there was a similar trend of decreasing incidence from young to older age groups: 16.1 and 12.5 per 10,000 person-years respectively, in the 18-34 year age group, reducing to 4.2 and 8.6 per 10,000 person-years respectively, in the ≥65 year age group. In terms of UK-based studies, Court-Brown et al (2014) reported data from 2011 covering fractures of both the hand and wrist, but only for adults aged ≥35, without providing detail on the distribution of hand fractures. Laugharne et al (2013) reported details on the anatomical distribution of hand fractures in a UK population, but combined paediatric and adult data and did not report on incidence. They also did not report on carpal fractures. In agreement with the presented findings (Table 2-4), the little finger metacarpal was the commonest fractured bone amongst the phalanx and metacarpal fractures studied by Laugharne et al (2013). However, they found that metacarpals as a combined group were fractured more often than phalanges, whereas the converse was true in our dataset.

The rate of carpal fractures was substantially lower than a Norwegian study, which reported 10.6% of hand fractures (combining phalanx, metacarpus and carpus) studied were of a carpal bone (Hove 1993). UK studies offered largely varying incidences for carpal fractures. The annual incidence of 4.8 per 10,000 adults/year was substantially lower than the 18 per 10,000 adults/year found by van Staa et al (2001) and more in keeping with the 2.7 per 10,000 adults/year reported by Court-Brown et al (2014).

Several epidemiological studies have described the incidence of distal radius fractures. These suggested a varying incidence around the world which has changed over several decades, but heterogeneity in sampling criteria and method of fracture identification (diagnostic codes, radiograph review, etc) make direct comparison challenging (MacIntyre and Dewan 2016). The annual incidence of distal radius/ulna fracture in our study population was somewhat lower than the 28 per 10,000/year of a recent study based on a Swedish population (Jerrhag et al. 2017). In terms of studies specific to the UK, most were based on data from around 20 years ago (O'Neill et al. 2001; van Staa et al. 2001; Thompson et al. 2004). van Staa et al (2001) reported an annual incidence of 22 per 10,000 while Court-Brown et al (2014) reported an overall incidence (regardless of sex) of 30 per 10,000/year. The difference between our results and that of the latter might be explained by their study population only including adults aged ≥35.

On considering distal radius/ulna fractures in the context of men and women, in prior UK-based studies the overall trend in incidences across age groups for men and women were broadly similar to our findings. Thompson et al (2004) showed much higher incidences in women of age 55+ years compared to men, while O'Neill et al (2001) reported a crude incidence in women of 36.6 per 10,000/year and in men of 8.9 per 10,000/year. In the US, although Karl et al (2015) also reported an overall trend of increasing incidence with age, they differed substantially from our findings in that their ratio of distal radius/ulna fractures in men to women was almost 1:1 (with women sustaining 54.6% of these fractures). Unfortunately, they did not report a breakdown of cases combining age and gender stratification, and thus reasons behind this difference to our results remain unclear.

The validation steps were useful in demonstrating the importance of double screening of PACS reports in studies such as this, which eliminated errors that would have occurred by relying on single reviewer screening. In terms of fracture reporting, a meta-analysis on reporting of plain radiographs found that radiographers report these with a pooled sensitivity of 92.6% and specificity of 97.7%, and also concluded that there was no evidence of a difference between radiographer and radiologist reporting accuracy for emergency department radiographs (Brealey et al. 2005); this meta-analysis did not selectively examine interpretation of skeletal radiographs or even more specifically those of the wrist and hand. Buskov et al (2013) concluded that radiographers tended to miss fewer fractures but overcall fractures more often than radiology trainees. In our study, radiographers produced approximately 70% of the PACS reports so a skew towards this pattern of PPV and NPV was consistent with these findings. Slight over-reporting of fractures is probably an advantage from a patient safety perspective compared to the alternative of missed injuries.

2.2.5.3 Strengths and limitations

The strengths of this study include that it was a consecutive series over a year, it was based on data from a busy UK major trauma centre involved in managing a large volume of relevant injuries, and that it underwent validation via independent data extraction. By focusing on PACS reports, we extracted data on fracture characteristics (such as the zone of bones fractured) which were more granular than typical hospital diagnostic coding allows. Such details can be of value in clinical and research resource allocation and planning.

A limitation of the study was that it relied upon accurate interpretation and reporting of radiographs. However, the validation step revealed very high positive and negative predictive values for the findings in PACS reports. This was reassuring and suggests that review of PACS reports is a reasonable method of data collection for these purposes. The numbers described were likely an underestimation, as they could not account for people who did not seek medical attention for their injury, nor did they typically account for patients who had community/primary care management of their injury (unless a hospital radiograph was requested).

Another limitation, common to epidemiological studies, was having to estimate the population in order to calculate incidences. This was done using data for the main city and directly encompassing districts. As a major trauma centre and secondary care institution, the host Trust also receives referrals from a wider area with a population, which is challenging to accurately define due to overlap with other secondary care institutions. A broader limitation is that the population being studied may not accurately represent the wider UK population, and even more so may not be generalisable to populations around the world. On checking population statistics available for the UK at a similar time point I found that the age and gender profile broadly matched that of the overall UK adult population, but the study population did have a higher proportion in the 18-25 year old age range. In addition, the 2015 indices of deprivation report suggest that the deprivation profile of the area containing the study population consisted of a higher proportion living within the most deprived 10% of neighbourhoods compared to the more typical profile elsewhere in the UK (Smith et al. 2015). However, the study involved a comprehensive dataset of sequential injuries in an urban population, and the incidences of these injuries would likely apply in a broad sense in similar populations elsewhere.

Finally, a limitation with regards to the applicability of the findings to the scope of the COS is that this study has focused on fractures, whereas the COS includes joint injuries not involving a fracture as well. Information on the latter could not be reliably obtained by examination of the PACS reports which was the underlying method of data collection for this work.

2.2.5.4 Conclusion

This study clarified the incidence and anatomical distribution of fractures of the hand and wrist in an urban adult population in the United Kingdom. It offers an update to previously reported findings which tended to involve data from other countries (much of which is over a decade old) and were not as comprehensive in their coverage of the full range of hand and wrist fractures. This makes comparisons difficult, given the tendency to study populations that differed due to arbitrary age limits or mixing of adult and paediatric injuries.

The methods described here are novel and can be applied by others – extracting data from electronic reports of radiographs minimises the need for time-consuming direct clinician review of thousands of radiograph images, yet provides detailed data that cannot be obtained by simple diagnostic coding records alone. The findings can assist in decisions of prioritisation and healthcare resource allocation for the range of injuries covered, as these rely on a clear grasp of the scale and scope of healthcare conditions.

2.3 Chapter conclusion

The information derived from the work in this chapter guided the development of the COS. Setting a clear scope of injuries was vital for all subsequent steps of the process. Meanwhile, awareness of the age and gender variation in demographics for distal radius fractures as opposed to other hand fractures and joint injuries was important as this could theoretically have influenced the outcome domains selected in existing and future clinical research. This is further explored in the systematic review (Chapter 3) and thereafter played a key part in guiding the construction of purposive sampling frames for the patient qualitative interviews and focus groups (Chapter 4).

Chapter 3 Systematic review of treatment outcomes

3.1 Introduction

Having established the scope of the COS and the epidemiology of hand fractures and joint injuries, I proceeded with the next step recommended in the COMET Initiative handbook: that of identifying 'existing knowledge' in terms of outcome domains to be measured in clinical studies (Williamson et al. 2017). The COS-STAD recommendations mention literature review as a means of identifying an initial list of outcome domains which are then refined through consensus process to form a COS (Kirkham et al. 2017).

In developing such a longlist of outcome domains for further consideration in the consensus process it is useful to include those outcomes which have been used in recent clinical studies on hand fractures and joint injuries. The common outcomes amongst these could be considered an informal consensus amongst researchers by virtue of being selected more frequently (Clarke and Williamson 2016). By conducting a systematic review of outcome domains used in recent clinical studies, I expected that a fairly comprehensive representation of outcome domains considered important by the researcher stakeholder group could be derived.

This systematic review also served to quantify the issue of heterogeneity in outcome selection in clinical research in this health area, highlighting the underlying need for a COS.

3.2 Aims

The primary aim of this systematic review was to identify outcome domains previously reported in recently published studies on the management of hand fractures and joint injuries in adults.

The secondary aim was to assess reporting bias of outcomes in these studies.

3.3 Methods

The design of this systematic review was guided by the Core Outcome Set-STAndards for Development (COS-STAD) (Kirkham et al. 2017). The protocol for the review was registered prior to commencement of data extraction on the PROSPERO International prospective register of systematic reviews (CRD42019126299) (Deshmukh et al. 2019).

3.3.1 Criteria for considering studies for this review

3.3.1.1 Types of studies

Inclusion criteria:

- Publication in English or with English translation available
- Randomised or quasi-randomised controlled trials or pilot studies
- Prospective observational cohort studies or case series with ≥100 patients enrolled

Exclusion criteria:

- Systematic reviews

- Biomechanical studies
- Cadaveric studies
- Reports where only abstract (rather than full report) was available, e.g. conference abstracts
- Unpublished and ongoing studies
- Studies not assessing treatment of patients, i.e. purely diagnostic or epidemiological studies
- Purely clinimetric studies only evaluating or validating measurement instruments

3.3.1.2 Types of participants

Inclusion criteria and exclusion criteria were as per the scope of the COS detailed in Section 2.1.3. The following were included:

- Studies of mixed populations (e.g. adults and children) if ≥90% of the population meets the review inclusion criteria.
- Studies of mixed diagnoses (e.g. distal radius, mid-shaft radius and ankle fractures) if ≥80% of the population met the review inclusion criteria (e.g. total of ≥100 patients with a condition of interest). Alternatively, mixed diagnosis studies were included if the outcomes were reported in a manner which allowed differentiation between diagnoses and the study involved ≥100 patients with a condition of interest.
- Studies where the publication of the secondary report or of further follow-up of original study participants was within the search dates specified as long as all other criteria were met. Analysis of such additional report or further follow-up studies also included data from the primary study regardless of whether the original study fell

within the search period. Such cases were considered as a single study, since the study population was the same.

3.3.1.3 Types of interventions

Studies on any treatment interventions for hand fractures and joint injuries were included, whether conservative or surgical. Studies on prophylactic or preventative interventions were excluded.

3.3.1.4 Types of outcome measures

Any outcome which eligible studies aimed to assess or measure and/or report in the published record were included, whether pre-specified or not. Studies which only reported early anaesthesia/analgesia-related outcomes (i.e. within first 24 hours of intervention) were excluded.

3.3.2 Search methods

3.3.2.1 Electronic search terms and databases

I used the following key search concepts:

A. Bones, joints, tendons and ligaments of the hand, carpi and distal radius
B. Generic terms for fractures and joint injuries
C. Specific hand fractures and joint injury terms

Free text terms and subject headings for each database for these concepts were identified. Searches combined the electronic search terms of search concept [A] with [B] using the

Boolean ‘AND’. I then added to these by combining with search concept [C] using the Boolean ‘OR’. Details of the databases searched and interfaces utilised are summarised in Table 3-1. The electronic search was designed to remain inclusive, by focusing on anatomy and injury types rather than combining with intervention-specific terms in the query. The search was not filtered by language.

The search strategies were compiled with guidance of an information specialist.

Table 3-1 Databases searched and interface utilised

Database	Interface	Coverage
MEDLINE	Ovid (Epub Ahead of Print, In-Process & Other Non-Indexed)	1946 - 2019
Embase	Ovid	1974 - 2019
Cochrane Central Register of Controlled Trials (CENTRAL)	Wiley	1999 - 2019
PubMed	PubMed	1946 - 2019
CINAHL	CINAHL Plus with Full Text	1937 - 2019
PEDro	NeuRA	1966 - 2019
PsycINFO	Ovid	1806 - 2019

3.3.2.2 Search strategy and time frame

Examples of the search strategies for two databases are included in Appendix 2. For this review I undertook a staged search strategy approach, as per the COMET Initiative handbook (Williamson et al. 2017) with an initial search for a period just over five years (the search was run on 29/03/19, and the search time period was from 01/01/2014 to 29/03/2019). Outcomes extracted from the most recent four-year period (01/01/2015 to 29/03/2019) were compared to those extracted from studies published in the earliest year of the search (01/01/2014 to 31/12/2014).

Following the COMET suggestion, if any new important outcomes were identified in that earliest year, the search would have been extended by a further year proceeding until outcome saturation was reached or the search was extended to a maximum of ten years. This staged approach achieved a balance between resource use and diminishing returns.

3.3.2.3 Searching other resources

I did not expand the search through personal contact with authors, search the reference lists of included studies or search the grey literature.

3.3.3 Data collection and analysis

3.3.3.1 Selection of studies

Study selection is reported in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Moher et al. 2009).

Titles and abstracts of retrieved articles were imported into bibliography management software EndNote version X8 (Thomas Reuters, New York City, NY, USA) and duplicates removed. The remaining titles and abstracts were independently screened against the inclusion criteria by two reviewers to determine potentially eligible articles. In the case of being unable to exclude an article based solely on the title and abstract, then it was categorised as a potentially eligible article. Outcomes from the screening by both reviewers were compared and any disagreements resolved by discussion.

We assessed the full text of all potentially eligible articles for inclusion, with any queries being resolved through discussion with a third reviewer. In the case of an article being part of a follow-up or secondary analysis of a study, the original study report or primary analysis was located and included.

3.3.3.2 Data extraction

Data were extracted using a piloted data collection form.

The following baseline data were extracted by one reviewer: author details, country where study was conducted, single- or multi-centre, year of publication, journal of publication, whether time points for outcomes were since injury, randomisation or intervention, type of study and (if a randomised/quasi-randomised study) whether the study was registered.

All other data, as listed below, were extracted independently by two reviewers and checked for agreement that all outcomes were identified. Disagreement was resolved through discussion, with any unresolved cases being decided through discussion with the senior author. If any necessary data appeared unclear or unavailable, I attempted to contact the study’s corresponding author by email to seek clarification.

The following data were extracted from each study:

- injury/condition being investigated
- intervention(s) under investigation
- study population age
- study population sex
- outcomes to be assessed/measured (as stated in a registry, if the study has been registered, or otherwise at any point in the publication of the study)
- outcome definitions as described by study authors
- outcome time points
- outcome classification (e.g. primary, secondary)
- outcome measures/instruments utilised including mode of assessment – e.g. telephone, online, face-to-face
- reporting of outcomes – as per outcome matrix described by Kirkham et al (2010) but modified as described later in Section 3.3.4

3.3.3.3 Data analysis

We analysed all extracted verbatim outcomes (i.e. the literal outcome as stated by study authors) for similarity in meaning through discussion between two reviewers for all outcomes. For example, “finger flexion” and “flexion of the finger” are technically two different verbatim outcomes. However, if the two are actually measured in the same way the only difference is in the label used by each study. Conversely, some verbatim outcomes might be labelled the same across different studies but have differing methods of measurement. For example, “finger flexion” would constitute two unique outcomes if reported in degrees of joint movement in some studies but as a percentage compared to the contralateral limb in others.

As such, we examined all verbatim outcomes in light of the method of measurement used. Where these simply had differing terminology but the same meaning, they were categorised into one unique outcome. However, where verbatim outcomes had similar terminology but different measurement methods across studies, they were split into two or more unique outcomes as appropriate. In cases of disagreement between the two reviewers, a third reviewer was consulted.

Two reviewers independently categorised the outcomes from each study into second-level WHO ICF outcome domains (World Health Organization 2001). Any disagreements were resolved through discussion, with consultation with the senior author if necessary.

Outcome domains and their frequency of selection were summarised. I also compared outcomes from studies on patients with distal radius fractures with those from studies on the other types of hand and wrist injuries within the scope of the COS.

I anticipated that in many cases one or more outcomes of a study would actually be the score from a patient-reported outcome measurement instrument (e.g. Patient Evaluation Measure, Patient Rated Wrist Evaluation). Therefore, as part of the data analysis for this study, I followed the principles outlined by Macefield et al (2014). This involved analysis of Patient Reported Outcome Measure (PROM) tools with categorisation of scale and single item components into ‘outcome domains’. This was done independently by the same reviewers who categorised the simpler ‘unique outcomes’ into outcome domains. Any disagreement was discussed by both reviewers and resolved through consultation with the senior author as necessary.

3.3.4 Assessment of risk of outcome reporting bias in included studies

Outcome reporting bias assessment was accomplished through a use of a modified version of the outcome matrix described by Kirkham et al (2010). The outcome matrix essentially conveys information on which outcomes are primary or secondary (if clarified in the study) and whether reporting of each outcome was ‘full’, ‘partial’ or ‘not done’. This does not take into account the actual results obtained, but instead works on the premise that any outcome specified for inclusion should then have the corresponding results reported in the final publication.

The primary outcome was defined as the following (in decreasing order of preference):

i. The primary outcome specified in the study
ii. The outcome upon which the study sample size calculation was based (this was also the selected primary outcome if the study inappropriately specified multiple primary outcomes)
iii. The outcome which appeared to most closely correspond with the study aim (where the aim itself was specific enough to allow this determination)

Where none of the above resulted in a clear primary outcome being determinable, all outcomes in the study were considered as secondary outcomes. Primary and secondary outcomes should have time points specified in all cases.

I modified the matrix suggested by Kirkham et al (2010) with an additional ‘unexpected’ category for those outcomes which were not listed in a study’s registration (for randomised/quasi-randomised controlled trials, where available) or otherwise in the study publication prior to the ‘Results’ section, but were reported in the ‘Results’ or ‘Discussion’ (hence would not have been ‘expected’ to be amongst the outcomes reported). For increased granularity in categorisation, I added a ‘minimal’ category which covered those outcomes which were effectively just ‘mentioned in passing’ with no numerical values reported. Table 3-2 shows the full definitions used for the outcome matrix in this review.

We performed independent two-reviewer assessment of outcome reporting status.

We excluded generic ‘adverse event’ or ‘complication’ outcomes from the assessment, except in cases where specific named complications were identified as being standalone study outcomes.

For randomised/quasi-randomised controlled trials, I aimed to use protocols where available but otherwise defaulted to trial registration data. If the latter was not clarified in the publication itself then it was searched for online on the WHO International Clinical Trials Registry Platform (World Health Organization 2022). If still not locatable, then I contacted the study’s corresponding author by email. Where there was a lack of response to the email, I assumed that no trial registration was completed for that study. These studies have not been ignored in the analysis of outcome reporting bias, but instead were categorised separately to those where a trial registration was located.

Table 3-2 Modified outcome matrix reporting status categories for risk of outcome reporting bias

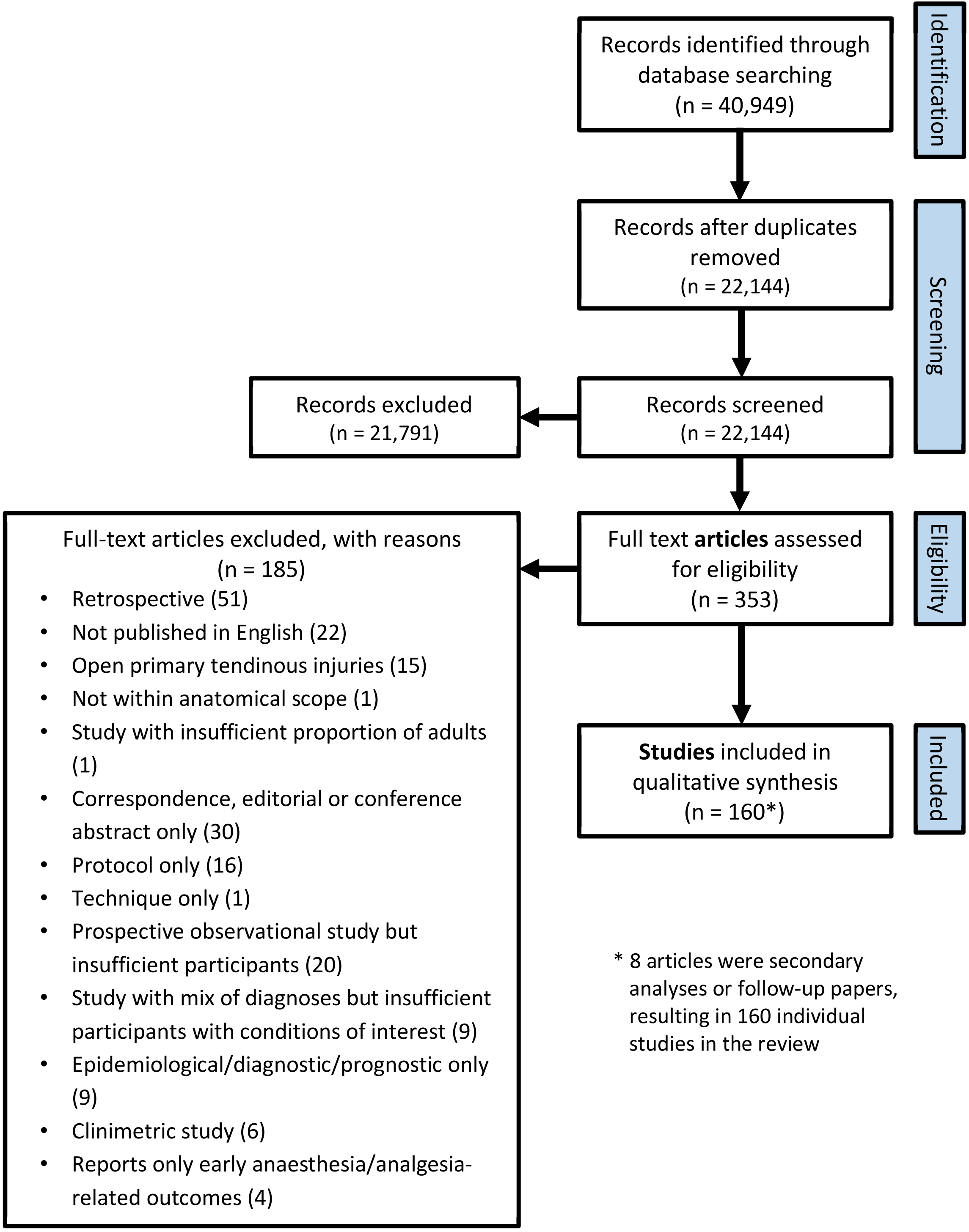
Category	**Definition**
Not done	No clear reporting of an outcome either through description, a table or figure
Minimal	Outcome reported merely by a summary comment (e.g. 'there was no significant difference between the intervention arms') but with no numerical values provided, or if there was such deficiency of information that the reporting was no longer meaningful (e.g. values given but no indication of time point)
Partial	Outcome reported but not at all time points specified elsewhere in the study/registration or lacked sufficient detail to be included in a meta-analysis (e.g. mean value was reported but not variance or p-value for the difference in means between intervention arms)
Complete	Outcome reported at all time points specified elsewhere in the study and with sufficient detail to allow inclusion in a meta-analysis
Unexpected	Outcome reported but was not specified in the study registration or prior to the 'Results' section of the study

3.4 Results

3.4.1 Results of the search

A PRISMA flow diagram for the review is shown in Figure 3.1. A total of 160 studies were included in the review.

Figure 3.1 PRISMA Flow Diagram



3.4.2 Included studies

The 160 included studies collected data from a total of 18,712 participants. The majority were randomised controlled trials. There was a widespread geographic distribution but included studies did predominately originate in the Northern Hemisphere. Details of the study characteristics are summarised in Table 3-3.

Table 3-3 Table of study characteristics

Study Characteristic	n/N (%)
Type of study	
Randomised controlled trial	99/160 (61.9%)
Quasi-randomised controlled trial	11/160 (6.9%)
Prospective cohort study	24/160 (15.0%)
Prospective case series	21/160 (13.1%)
Randomised pilot/feasibility study	5/160 (3.1%)
Geographic distribution of recruitment	
Africa	4/160 (2.5%)
Asia	56/160 (35.0%)
Australasia	7/160 (4.4%)
Europe	65/160 (40.6%)
North America	22/160 (13.8%)
South America	6/160 (3.8%)
Single/Multi-centre	
Single	136/160 (85.0%)
Multi	24/160 (15.0%)
No. of participants (in randomised/quasi-randomised studies)	
≤50	49/110 (45%)
51-100	41/110 (37%)
>100	20/110 (18%)

Table 3-3 Table of study characteristics (cont'd)

Study Characteristic	**n/N (%)**
Age distribution of participants (Specified in 146 studies)	(years)
Range of mean age	26.0 – 77.1
Weighted mean ± SD	53.9 ± 12.1
Gender distribution of participants (Specified in 150 studies)	(% female)
Range	0 – 100
Weighted mean ± SD	61.9 ± 24.4

3.4.3 The outcomes selected and reported in recent literature on hand fractures and joint injuries

There were 1,777 verbatim outcomes across all included studies. Through deduplication and rationalisation, these were categorised to a total of 639 unique outcomes. We then further categorised the unique outcomes using the WHO ICF framework (World Health Organization 2001) into 74 outcome domains. A median of 17 (IQR 7 to 22) outcome domains were identified in the included studies.

The full set of 74 outcome domains as well as the number (and percentage) of studies in which each of them were used are shown in to Table 3-8, separated according to the WHO ICF 'components' as outlined in Section 1.4.2.2. There were several unique outcomes which were either not definable or not covered by the WHO ICF framework (World Health Organization 2001) – the outcome domains recorded for these were given the prefix or 'nd-' and 'nc-' respectively, as suggested by the ICF linking rules guide (Cieza et al. 2005). In each table the outcome domains are presented in descending order of frequency across all (DRF and non-DRF combined) studies.

3.4.3.1 'Body functions' component outcome domains

Table 3-4 Outcome domains within the 'body functions' component

Outcome domain	DRF studies (n=121)	Non-DRF studies (n=39)	Total (n=160)
b280 Sensation of pain	108 (89%)	39 (100%)	147 (92%)
b710 Mobility of joint functions	102 (84%)	35 (90%)	137 (86%)
b730 Muscle power functions	94 (78%)	29 (74%)	123 (77%)
b265 Touch function	78 (64%)	25 (64%)	103 (64%)
b134 Sleep functions	69 (57%)	25 (64%)	94 (59%)
b126 Temperament and personality functions	59 (49%)	14 (36%)	73 (46%)
b640 Sexual functions	46 (38%)	11 (28%)	57 (36%)
b152 Emotional functions	33 (27%)	5 (13%)	38 (24%)
b180 Experience of self and time functions	15 (12%)	5 (13%)	20 (12.5%)
b130 Energy and drive functions	17 (14%)	1 (2.6%)	18 (11%)
b820 Repair functions of the skin	12 (9.9%)	1 (2.6%)	13 (8.1%)
b455 Exercise tolerance functions	10 (8.3%)	1 (2.6%)	11 (6.9%)
b289 Sensation of pain, other specified and unspecified	8 (6.6%)	1 (2.6%)	9 (5.6%)
b270 Sensory functions related to temperature and other stimuli	7 (5.8%)	1 (2.6%)	8 (5%)
b830 Other functions of the skin	7 (5.8%)	1 (2.6%)	8 (5%)
b164 Higher-level cognitive functions	7 (5.8%)	0 (0%)	7 (4.4%)
b760 Control of voluntary movement functions	5 (4.1%)	1 (2.6%)	6 (3.8%)

Table 3-4 Outcome domains within the 'body functions' component (cont'd)

Outcome domain	DRF studies (n=121)	Non-DRF studies (n=39)	Total (n=160)
b140 Attention functions	3 (2.5%)	0 (0%)	3 (1.9%)
b160 Thought functions	3 (2.5%)	0 (0%)	3 (1.9%)
b715 Stability of joint functions	2 (1.7%)	1 (2.6%)	3 (1.9%)
b144 Memory functions	2 (1.7%)	0 (0%)	2 (1.3%)
b117 Intellectual functions	1 (0.8%)	0 (0%)	1 (0.6%)
b122 Global psychosocial functions	1 (0.8%)	0 (0%)	1 (0.6%)
b156 Perceptual functions	1 (0.8%)	0 (0%)	1 (0.6%)
b260 Proprioceptive function	1 (0.8%)	0 (0%)	1 (0.6%)

We identified 25 WHO ICF outcome domains from the 'body functions' component. These were based on 386 unique outcomes.

3.4.3.2 'Activities and participation' component outcome domains

Table 3-5 Outcome domains within the 'activities and participation' component

Outcome domain	DRF studies (n=121)	Non-DRF studies (n=39)	Total (n=160)
d850 Remunerative employment	84 (69%)	31 (80%)	115 (72%)
d440 Fine hand use	85 (70%)	29 (74%)	114 (71%)
d920 Recreation and leisure	84 (69%)	29 (74%)	113 (71%)
d510 Washing oneself	83 (69%)	28 (72%)	111 (69%)
d430 Lifting and carrying objects	82 (68%)	29 (74%)	111 (69%)
d640 Doing housework	82 (68%)	28 (72%)	110 (69%)
d445 Hand and arm use	81 (67%)	26 (67%)	107 (67%)
d550 Eating	79 (65%)	26 (67%)	105 (66%)
d230 Carrying out daily routine	73 (60%)	26 (67%)	99 (62%)
d540 Dressing	77 (64%)	18 (46%)	95 (59%)
d750 Informal social relationships	63 (52%)	21 (54%)	84 (53%)
d760 Family relationships	62 (51%)	21 (54%)	83 (52%)
d470 Using transportation	48 (40%)	11 (28%)	59 (37%)
d650 Caring for household objects	46 (38%)	11 (28%)	57 (36%)
d410 Changing basic body position	40 (33%)	4 (10%)	44 (28%)
d530 Toileting	35 (29%)	5 (13%)	40 (25%)
d450 Walking	21 (17%)	2 (5.1%)	23 (14%)
d455 Moving around	12 (9.9%)	1 (2.6%)	13 (8.1%)
d839 Education unspecified	7 (5.8%)	2 (5.1%)	9 (5.6%)
d520 Caring for body parts	7 (5.8%)	1 (2.6%)	8 (5.0%)
d630 Preparing meals	5 (4.1%)	3 (7.7%)	8 (5.0%)
d560 Drinking	4 (3.3%)	1 (2.6%)	5 (3.1%)

Table 3-5 Outcome domains within the 'activities and participation' component (cont'd)

Outcome domain	DRF studies (n=121)	Non-DRF studies (n=39)	Total (n=160)
d620 Acquisition of goods and services	4 (3.3%)	1 (2.6%)	5 (3.1%)
d570 Looking after one's health	3 (2.5%)	0 (0%)	3 (1.9%)
d420 Transferring oneself	2 (1.7%)	1 (2.6%)	3 (1.9%)
d460 Moving around in different locations	2 (1.7%)	1 (2.6%)	3 (1.9%)
d855 Non-remunerative employment	2 (1.7%)	1 (2.6%)	3 (1.9%)
d870 Economic self-sufficiency	2 (1.7%)	1 (2.6%)	3 (1.9%)
d845 Acquiring, keeping and terminating a job	1 (0.8%)	2 (5.1%)	3 (1.9%)
d475 Driving	2 (1.7%)	0 (0%)	2 (1.3%)
d710 Basic interpersonal interactions	1 (0.8%)	0 (0%)	1 (0.6%)
d720 Complex interpersonal interactions	1 (0.8%)	0 (0%)	1 (0.6%)
d860 Basic economic transactions	1 (0.8%)	0 (0%)	1 (0.6%)
d910 Community life	1 (0.8%)	0 (0%)	1 (0.6%)

We identified 34 WHO ICF outcome domains from the ‘activities and participation’ component. These were based on 107 unique outcomes.

3.4.3.3 'Body structures' component outcome domains

Table 3-6 Outcome domains within the 'body structures' component

Outcome domain	**DRF studies (n=121)**	**Non-DRF studies (n=39)**	**Total (n=160)**
s730 Structure of upper extremity	74 (61%)	18 (46%)	92 (58%)
s810 Structure of areas of skin	5 (4.1%)	1 (2.6%)	6 (3.8%)

We identified two WHO ICF outcome domains from the 'body structures' component, which were based on 108 unique outcomes.

3.4.3.4 'Environmental factors' component outcome domains

Table 3-7 Outcome domains within the 'environmental factors' component

Outcome domain	**DRF studies (n=121)**	**Non-DRF studies (n=39)**	**Total (n=160)**
e580 Health services, systems and policies	28 (23%)	8 (21%)	36 (23%)
e565 Economic services, systems and policies	1 (0.8%)	1 (2.6%)	2 (1.3%)

These two WHO ICF outcome domains within the 'body structures' component were based on 36 unique outcomes.

3.4.3.5 Outcome domains outside WHO ICF framework

Table 3-8 Outcome domains not classified or definable within the WHO ICF framework

Outcome domain	**DRF studies (n=121)**	**Non-DRF studies (n=39)**	**Total (n=160)**
nc-Complications/Adverse events	73 (60%)	20 (51%)	93 (58%)
nc-Overall satisfaction	27 (22%)	11 (28%)	38 (24%)
nc-Bone healing	20 (17%)	16 (41%)	36 (23%)
nd-gh (general health)	18 (15%)	3 (7.7%)	21 (13%)
nd-ph (physical health)	17 (14%)	1 (2.6%)	18 (11%)
nc-Bone healing time	7 (5.8%)	9 (23%)	16 (10%)
nc-Technical (related to intervention)	11 (9.1%)	1 (2.6%)	12 (7.5%)
nc-Satisfaction with intervention	2 (1.7%)	1 (2.6%)	3 (1.9%)
nc-Blood tests	1 (0.8%)	0 (0%)	1 (0.6%)
nc-Individualised rating scale	1 (0.8%)	0 (0%)	1 (0.6%)
nd-Patient adherence to treatment	0 (0%)	1 (2.6%)	1 (0.6%)

WHO ICF – World Health Organization International Classification of Functioning, Disability and Health

nd – not definable within WHO ICF

nc – not covered within WHO ICF

There were 11 outcome domains which could not be defined or classified within the WHO ICF framework. These were based on 110 unique outcomes.

We identified 29 distinct multi-domain PROMs (i.e. not including single-domain PROMs such as ‘pain visual analogue score’). In addition, some studies used only a subscale/component of the PROM as a ‘unique outcome’ (e.g. the pain subscale of the Patient-Rated Wrist Evaluation questionnaire) effectively reducing it to a single-domain PROM.

The most frequently used multi-domain PROMs were the Disabilities of the Arm, Shoulder and Hand (DASH) (Hudak et al. 1996) questionnaire (57/160 studies, 35.6%), the Patient-Rated Wrist Evaluation (PRWE) (MacDermid et al. 1998) questionnaire (30/160 studies, 18.8%) and the abbreviated version of the DASH questionnaire called QuickDASH (Beaton et al. 2005) (29/160 studies, 18.1%).

A summary of PROMs as well as the linked WHO ICF outcome domains and frequency of use across the included studies is provided in Appendix 3.

3.4.4 Comparison of distal radius fracture (DRF) and non-DRF studies

I performed a comparison between studies involving either a majority or solely distal radius fractures (DRF) and those involving the other hand fractures and joint injuries (non-DRF). Of the 160 included studies, 121 (75.6%) involved mainly patients with distal radius fractures.

Table 3-9 summarises the age and gender distribution of participants in DRF studies as compared to non-DRF studies. There were some DRF studies which specified an inclusion factor of a minimum age well within the ‘adult’ age group (e.g. 55 years old) and would therefore have slightly skewed the above demographic data but this was infrequent and excluding these studies yielded a weighted mean age of 56.9 years (±9.6 years) for the remaining DRF studies.

Table 3-9 Comparison of participant age and gender distribution between DRF and non-DRF studies

	Type of study	
	DRF study	**Non-DRF study**
Number of studies	121/160	39/160
Age distribution of participants	(n = 113)	years (n = 33)
Range of mean age (years)	32.2 – 77.1	26.0 – 50.0
Weighted mean ± SD (years)	58.2 ± 10.4	38.5 ± 6.0
Gender distribution of participants	(n = 112)	(n = 38)
Range (% female)	12.0 – 100.0	• 0.0 – 59.0
Weighted mean ± SD (% female)	71.8 ± 18.7	32.9 ± 18.3

Table 3-10 summarises the five commonest PROMs and ten commonest outcome domains across all studies. These were also the commonest PROMs and outcome domains across each of the DRF and non-DRF study groups.

Table 3-10 Top five PROMs and ten outcome domains most commonly used across all study groups and overall

	DRF studies (n=121)	Non-DRF studies (n=39)	Total (n=160)
Patient reported outcome measure			
Visual analogue scale for pain	51 (42%)	15 (38%)	66 (41%)
Disabilities of the Arm, Shoulder and Hand questionnaire (DASH)	46 (38%)	11 (28%)	57 (36%)
Patient-Rated Wrist Evaluation (PRWE)	28 (23%)	2 (5.1%)	30 (19%)
Abbreviated Disabilities of the Arm, Shoulder and Hand questionnaire (QuickDASH)	17 (14%)	12 (31%)	29 (18%)
European Quality of Life Scale (EQ-5D-3L)	9 (7.4%)	4 (10%)	13 (8.1%)
Outcome domain			
b280 Sensation of pain	108 (89%)	39 (100%)	147 (92%)
b710 Mobility of joint functions	102 (84%)	35 (90%)	137 (86%)
b730 Muscle power functions	94 (78%)	29 (74%)	123 (77%)
d850 Remunerative employment	84 (69%)	31 (79%)	115 (72%)
d440 Fine hand use	85 (70%)	29 (74%)	114 (71%)
d920 Recreation and leisure	84 (69%)	29 (74%)	113 (71%)
d510 Washing oneself	83 (69%)	28 (72%)	111 (69%)
d430 Lifting and carrying objects	82 (68%)	29 (74%)	111 (69%)
d640 Doing housework	82 (68%)	28 (72%)	110 (69%)
d445 Hand and arm use	81 (67%)	26 (67%)	107 (67%)

An alternative way of comparing outcome domain selection between DRF and non-DRF studies is to examine which outcome domains were only found in one of the two study groups. Table 3-11 demonstrates these findings.

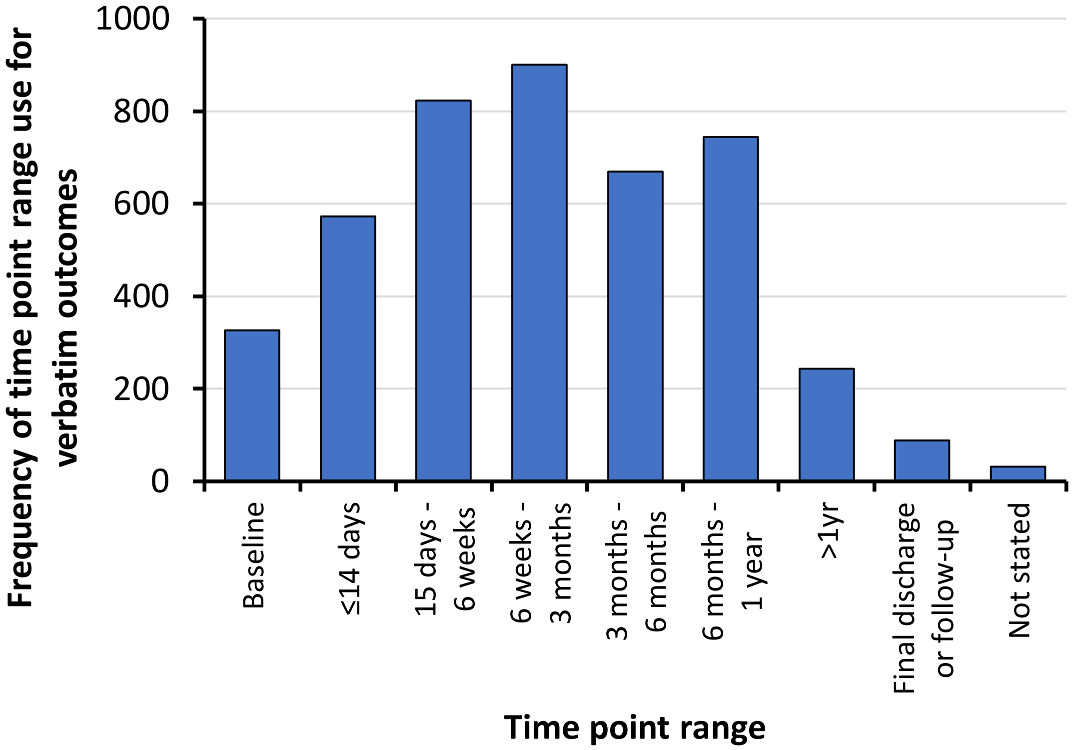
Table 3-11 Outcome domains unique to either DRF or non-DRF studies, ordered by frequency of reporting across all studies

Outcome domain	DRF studies (n=121)	Non-DRF studies (n=39)	Total (n=160)
b164 Higher-level cognitive functions	7 (5.8%)	0 (0%)	7 (4.4%)
b140 Attention functions	3 (2.5%)	0 (0%)	3 (1.9%)
b160 Thought functions	3 (2.5%)	0 (0%)	3 (1.9%)
b144 Memory functions	2 (1.7%)	0 (0%)	2 (1.3%)
b117 Intellectual functions	1 (0.8%)	0 (0%)	1 (0.6%)
b122 Global psychosocial functions	1 (0.8%)	0 (0%)	1 (0.6%)
b156 Perceptual functions	1 (0.8%)	0 (0%)	1 (0.6%)
b260 Proprioceptive function	1 (0.8%)	0 (0%)	1 (0.6%)
d570 Looking after one's health	3 (2.5%)	0 (0%)	3 (1.9%)
d410 Changing basic body position	2 (1.7%)	0 (0%)	2 (1.3%)
d475 Driving	1 (0.8%)	0 (0%)	1 (0.6%)
d710 Basic interpersonal interactions	1 (0.8%)	0 (0%)	1 (0.6%)
d720 Complex interpersonal interactions	1 (0.8%)	0 (0%)	1 (0.6%)
d750 Informal social relationships	1 (0.8%)	0 (0%)	1 (0.6%)
d860 Basic economic transactions	1 (0.8%)	0 (0%)	1 (0.6%)
d910 Community life	1 (0.8%)	0 (0%)	1 (0.6%)
nc-Patient-selected score	1 (0.8%)	0 (0%)	1 (0.6%)
nc-Patient adherence to treatment	0 (0%)	1 (2.6%)	1 (0.6%)
nc-Blood tests	1 (0.8%)	0 (0%)	1 (0.6%)
nc-Satisfaction with intervention	1 (0.8%)	0 (0%)	1 (0.6%)
nd-qol (quality of life)	1 (0.8%)	0 (0%)	1 (0.6%)

3.4.5 Outcome time point selection

Time points of outcomes are often heterogeneous. To determine meaningful heterogeneity resulting from use of multiple and varying time points for outcome assessment, I created time point ‘ranges’ post-intervention, representing typical follow-up windows. We then categorised the verbatim outcomes according to these ranges, with the frequency of use of the time point ranges demonstrated in Figure 3.2.

Figure 3.2 Frequency of use of time point ranges for the verbatim outcomes



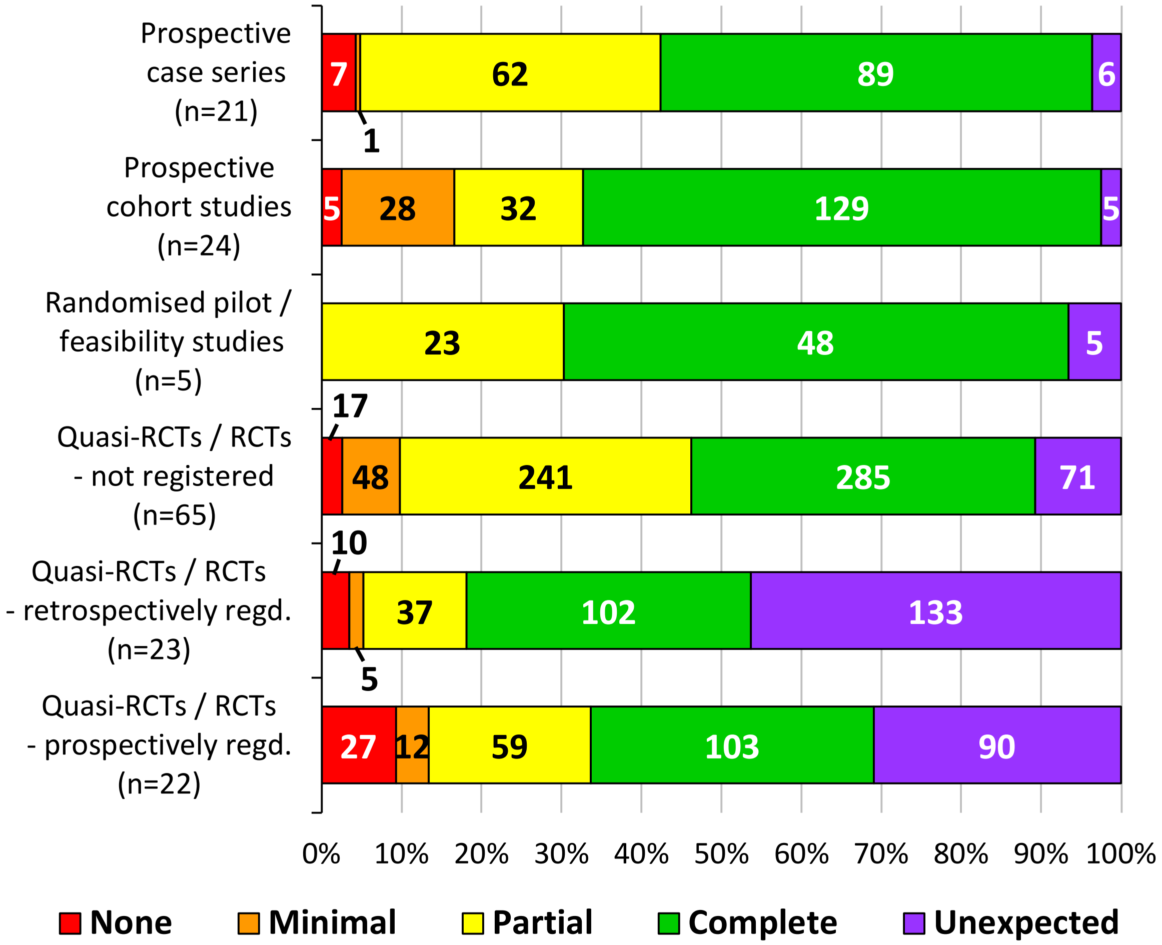
In terms of the overall heterogeneity added by time point ranges however, it is more useful to consider the unique outcomes as defined earlier. When factoring in the time point range in which outcomes were assessed or measured, the 639 unique outcomes correspond to 1,917 distinct outcomes.

3.4.6 Outcome reporting bias

Outcome reporting bias was assessed using the tool described in Section 3.3.4. Of the 110 randomised/quasi-randomised controlled trials, only 22 (20%) had prospective registration completed while a further 23 (21%) had retrospective registration. The remaining 65 (59%) were not registered.

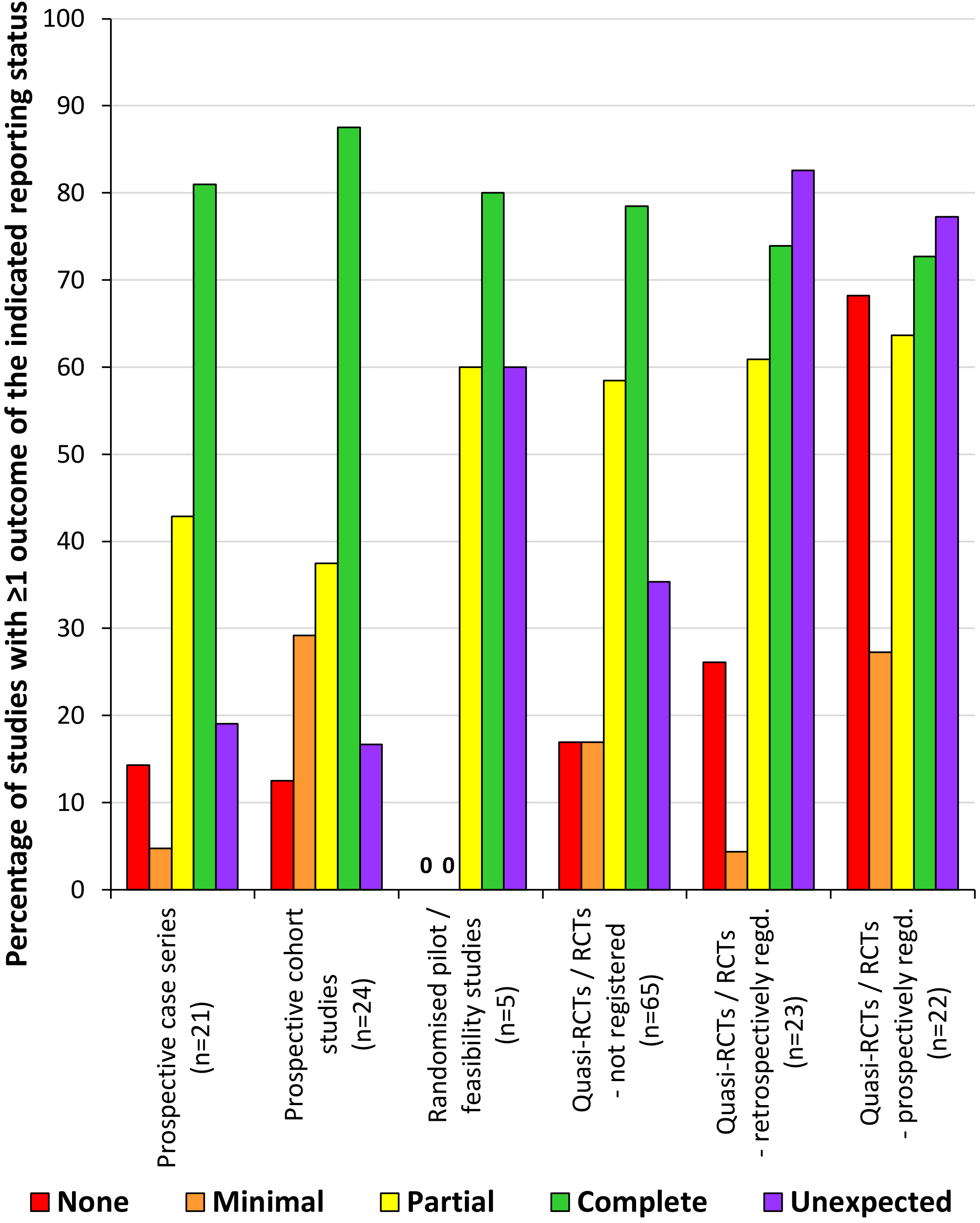
Figure 3.3 shows the outcome reporting bias across the different study types, with quasi-randomised and randomised controlled trials subdivided based on trial registration status. It summarises the categorisation of all verbatim outcomes based on reporting status and whether they were pre-specified or not (the latter being the 'unexpected' outcomes).

Figure 3.3 Cumulative bar chart showing percentage and number of verbatim outcomes within each reporting status category across study types



While the above figure conveys the quality of outcome reporting, it does not clarify whether status categories were clustered in particular studies. Figure 3.4 demonstrates the proportion of studies in each of the study type highlighted above which had at least one outcome in the specified outcome reporting status categories.

Figure 3.4 Cluster chart showing percentage of studies per study type with one or more outcomes in each of the reporting status categories



There were 66 instances of non-reported verbatim outcomes, comprising 62 unique outcomes. Of the non-reported outcomes, 26/66 (39%) involved range of movement (WHO ICF domain ‘b710 mobility of joint functions’), 12/66 (18%) involved pain (b280 sensation of pain) and 10/66 (15%) involved radiographic alignment (s730 structure of upper extremity).

The 310 instances where verbatim outcomes were unexpectedly reported comprised 193 unique outcomes. Table 3-12 summarises the WHO ICF domains linked to 10% or more of the unexpected verbatim outcomes. Appendix 4 shows a list of all 74 WHO ICF outcome domains and how frequently they arose in the verbatim outcomes per reporting status category.

Table 3-12 Commonly linked WHO ICF outcome domains for unexpected outcomes

WHO ICF outcome domain	Number (%) of unexpected outcomes (n=310)
b710 Mobility of joint functions	83 (27%)
s730 Structure of upper extremity	56 (18%)
b280 Sensation of pain	47 (15%)
d850 Remunerative employment	37 (12%)
e580 health services, systems and policies	33 (11%)
b730 Muscle power functions	33 (11%)

3.5 Discussion

3.5.1 Heterogeneity in outcome selection

COSs are developed to address the issues that afflict clinical research in fields where there is significant heterogeneity in outcome selection. This systematic review demonstrates that a wide range of outcomes were reported in the recent published research on hand fractures and joint injuries. This is reflected in the heterogeneity of selected outcomes, outcome measures and the specified outcome time points. Such variation hinders meta-analysis and predisposes to 'research waste' (Ioannidis et al. 2014; Yordanov et al. 2018).

The broad scope of injuries being covered does partially explain the high number of verbatim and unique outcomes found. For example, radiographic assessments and performance-based measures such as range of movement typically specify an anatomic location. Anatomic site will vary between injuries and results in distinct verbatim and unique outcomes even though the fundamental outcome domain might be the same. However, other factors which impact on pooling of data in meta-analysis have also contributed to the high numbers presented. For example, how the outcome is measured can have such an impact and different methods are used even for commonly reported outcome domains such as range of movement (e.g. measuring absolute degrees versus a percentage compared to the contralateral, uninjured side).

In an attempt to determine the degree of more meaningful variation in outcomes, the 639 unique outcomes were categorised into fundamental WHO ICF outcome domains. Despite neutralising factors such as anatomical site or method of measurement, there were still 74 distinct domains. Only three domains were reported in over 75% of studies; 'sensation of pain', 'mobility of joint functions' (range of movement) and 'muscle power function' (grip/pinch strength, performing certain actions). Considering the variation introduced by selection of different methods of measurement and of various time points, it is clear that much of the recently published outcomes data might prove difficult or inappropriate to meta-analyse.

3.5.2 Comparison of outcome selection for studies on distal radius fracture versus other hand fractures and joint injuries

I decided to compare these two groups of studies to check whether there were any key differences in the outcomes selected. The information from this review goes on to contribute towards a longlist of outcome domains considered in ensuing consensus work and as such the frequency of selection of outcome domains was not of direct concern to this purpose. All outcome domains, however frequently they were found to be used, were considered when developing the subsequent longlist. Nevertheless, differences in the outcomes selected could indicate that these two injury groups might differ in some way that is meaningful to researchers and clinicians in relation to the pertinent outcomes to be considered. If so this could be an indication that, at least from the professional perspective, a single COS encompassing all these injuries might not be as valid as two separate ones.

Differences in outcome selection could arise for a number of reasons, including not only a difference in the region of injury (i.e. wrist versus the rest of the hand) but also the fact that there tends to be a demographic disparity between these injury groups in both age and sex. Indeed, the results shown in Table 3-9 confirmed the anticipated demographic disparity between study participants of DRF versus non-DRF studies, with the former having a higher weighted mean age and proportion of females than the latter. It might follow that those designing the reviewed studies would tend to select different outcomes on the basis that the typical patient populations under study might have different priorities for outcomes.

However, on comparing outcome selection in DRF and non-DRF studies I found considerable overlap. Though the rank order of commonest PROMs and outcome domains varied slightly, the top five PROMs and top ten outcome domains were the same (Table 3-10). Differences in terms of outcome domains which are exclusive to one group or the other are quite infrequent (Table 3-11). In most cases these were outcome domains that were used by only one study, except for some of the 'specific mental function' WHO ICF outcome domains which typically involved PROMs assessing patients' perception of their injury or health.

Hence it appears that researchers consider similar outcome domains relevant to both populations despite any demographic differences in the participant populations of these studies. One proviso to drawing this conclusion is that most PROMs reflect multiple domains – this will give rise to greater apparent overlap in the observed outcome domains. The commonest multi-domain PROMs used were DASH (Hudak et al. 1996), PRWE (MacDermid et al. 1998) and QuickDASH (Beaton et al. 2005). DASH captures all of the ten commonest outcome domains, while PRWE and QuickDASH each capture eight of the ten commonest domains used (except for 'mobility of joint functions' and 'muscle power functions').

3.5.3 Outcome reporting bias

This review was primarily concerned with determining the outcomes prior studies have assessed/measured and not the quantitative results obtained. Therefore I believed that a general risk assessment of bias on overall aspects of study design and reporting was unnecessary. However, I felt that a risk assessment focused on outcome reporting bias would be worthwhile as this might reflect the relative value study authors placed on particular outcomes. The assessment tool described in Section 3.3.4 accounted for outcomes that were intended to be reported, but also which ones were subsequently actually reported. This provided a sense of the outcome reporting bias in the included studies. This review did not aim to delve into the depth of analysis that is possible through the Outcome Reporting Bias In Trials (ORBIT) classification systems (Kirkham et al. 2018). Such depth would be of relevance in systematic reviews aiming to reach a conclusion about the risk of bias in individual studies where the quantitative results reported were an important part of the data analysis. This did not pertain to the SO-HANDI review.

We found marked outcome reporting bias in agreement with previous studies of different populations. There were 66 instances of outcomes not being reported at all despite being specified in the publication or trial registration. Perhaps unsurprisingly the majority of these corresponded with the most commonly arising WHO ICF outcome domains across all verbatim outcomes – relating to range of movement, pain and radiographic alignment. Even more were reported incompletely, with only a brief comment or lacking sufficient detail for meta-analysis. All represent non-adherence to reporting standards. There were many ‘unexpected’ outcomes, with return to work and muscle power outcomes each featuring in over 10% of the unexpected outcome pool. One particularly interesting finding was that the ‘e580 Health services, systems and policies’ WHO ICF domain was involved in only 66/1,777 verbatim outcomes, and for half of these it was an unexpected outcome. Much of the time, this domain was linked to an outcome relating to ‘duration of surgery’ or similar, and I suspect that this outcome was being reported out of convenience since the data were readily available but had not been pre-specified as it had not been thought to be of importance at the study design phase.

Figure 3.3 shows that even in studies with retrospective registration there were unexpected outcomes (i.e. not listed in trial registration). Prospectively registered studies had a lower proportion of ‘unexpected’ outcomes as compared to retrospectively registered trials. It is possible that prospective registration correlates with a higher methodological quality in general, which is being reflected in this marker of outcome reporting bias. A formal assessment of overall study design and risk of bias across all domains was beyond the primary scope of this study.

I believed it would be informative to check whether the non-reported and unexpectedly reported outcomes tended to arise from only a small subset of studies. For trials with prospective registration or protocol, which was the most stringently assessed study type, the non-reported and unexpected outcomes were spread amongst the majority of studies – i.e. the majority of trials with prospective registration had at least one outcome not reported and unexpectedly reported (Figure 3.4).

These data demonstrate the significant issue of inconsistency in outcome reporting in clinical research on hand fractures and joint injuries, which can potentially be reduced by the use of a universally accepted COS. This is not a new issue, nor one isolated to clinical research in this field as highlighted by previous work by Dwan et al (2011) for randomised controlled trials, and Page et al (2014) for systematic reviews and meta-analyses. The importance of avoiding or minimising discrepancies in outcome reporting between pre-specified outcomes and final publication has been stressed by these authors, with the implication that post-hoc decisions on selective inclusion of outcomes presents a biased analysis.

3.5.4 Summary of findings of related reviews of outcomes

A prior study limited to a small selection of journals found that 'objective clinical measures' (e.g. grip strength, range of motion, functional status), 'quality of life' and morbidity were the commonest outcomes assessed (Chung et al. 2006). Weinstock-Zlotnick and Mehta (2016) reported on outcomes for wrist fractures and ligament injuries from RCTs between 2005 and 2015. Though lacking details in terms of WHO ICF outcome domains, they found 'range of movement', 'grip strength' and 'pain' were the commonest physical outcome measures used, while DASH and PRWE were the commonest PROMs. Their findings are in broad agreement with this review, indicating that priorities in outcome selection for studies preceding this search window were similar.

Goldhahn et al. (2014) undertook a literature review as part of a process which aimed to establish a core set for DRF. Though highlighting some commonly used outcomes, they did not present detail on the heterogeneity of outcomes identified. They found that 'radiological outcomes' (e.g. healing and alignment), 'grip strength', 'range of motion' and 'pain' were commonest, present in 68%, 49%, 49% and 38% of studies respectively. The 'pain' outcome was used much less commonly than the near-universal use found in this review. Conversely, their reported proportion of studies using 'radiological outcomes' was higher than that found by us. Such discrepancies are probably attributed to differences in the classification of outcome domains. For example, we identified that many PROMs incorporate pain as an outcome domain through the WHO ICF linking method, whereas Goldhahn et al. (2014) only considered pain when it was a distinct outcome instrument (i.e. not part of a PROM). Conversely, for radiological outcomes they combined some that we considered distinct, such as domains of 'healing' (bone healing) and 'alignment' (structure of upper extremity).

These other reviews of outcomes offer some reassurance in that the commonly prioritised outcome domains in clinical research on hand fractures and joint injuries are broadly similar to the findings of the SO-HANDI review.

3.5.5 Strengths and limitations

Although past reviews of hand fractures and joint injuries have highlighted issues of “inadequate outcome assessment” and “large variation in reported outcomes” (Handoll and Vaghela 2004; Poolman et al. 2006; Verver et al. 2017), this review specifically quantifies the magnitude of the problem. This review has aimed to be comprehensive by not restricting the search to only randomised controlled trials. However including all clinical research (in particular retrospective studies) was expected to be too resource-intensive and impractical. Randomised studies or large prospective studies with ≥100 participants were selected for inclusion as a balance between pragmatism and optimised sensitivity in finding the outcomes used in clinical research. In addition, these types of studies have historically been highlighted as having a higher chance of publication (Begg and Berlin 1989). Publication bias in favour of such studies could mean that they are more likely to influence clinical practice, which makes the outcome domains identified as being used in such studies all the more meaningful to identify.

The methodology of this review made use of the WHO ICF outcome domains as this is a recognised framework with a high level of granularity available. It has allowed us to categorise a vast array of unique outcomes into the 74 outcome domains identified. It does however have its own limitations, as highlighted by the need to create outcome domain labels when we felt they were not covered or defined by WHO ICF. I also realised that there would need to be a synthesis of these domains with those identified in the subsequent qualitative interview and focus group work with patients, as many of the WHO ICF outcome domains use technical terms or have a definition which seems unintuitive based on the outcome domain label alone.

Initial scoping searches demonstrated that there would be a large number of published studies that fit the inclusion criteria. A systematic review attempting to answer a specific question on treatment effectiveness would typically collate all existing relevant data. The aim of this review was to determine potential outcome domains for inclusion in a longlist that would then undergo a consensus process to work towards a COS. I therefore believed it reasonable to not expand the search through personal contact with authors of identified studies, reference lists or grey literature.

A limitation of this review was the exclusion of studies for which a publication in English could not be obtained (n=22, Figure 3.1). However, for almost every country of origin where this occurred there were other studies with an English publication available, maintaining some representation of these countries in the review. A theoretical limitation was the date range used, but I made this choice to focus on outcomes used in the more recent literature through a ‘staged search’ approach, as recommended by the COMET Initiative where a large number on included studies is expected (Williamson et al. 2017). By performing an initial search, data extraction and analysis for a pre-specified period of time and then extending the process for a further interval of time to ensure no new important outcomes were missed by the initial search window, then it can be deemed that ‘outcome saturation’ is reached. It is recommended that a minimum of two years is reviewed initially.

It is perhaps reasonable to speculate that this staged search allows outcomes being used in the more recent literature in this field to be found while potentially missing others which were used in the past but have fallen out of favour with researchers. It may seem that one should aim to find out all outcomes used in published research to be as comprehensive as possible. However this systematic review was used to identify ‘existing knowledge’ in terms of outcome domains to be measured in clinical studies from the perspective of clinicians and researchers – it stands to reason to focus on those which have been used in more recent publications as these are precisely what these stakeholder groups have prioritised. COMET recommends a balance between pragmatism and diminishing returns for resources invested at this stage, and on this basis we went beyond the minimum initial period of two years and instead reviewed just over four years in the initial stage before reviewing a further year. Further work described later in the thesis also afforded an opportunity for the professional stakeholder groups to suggest additional outcome domains, such as in Round 1 of the Delphi. There was therefore ample opportunity to capture additional domains that this staged search might have neglected.

The assessment of outcome reporting bias had a limitation in terms of the availability of information on the original outcomes selected by researchers in many studies. While several randomised/quasi-randomised controlled trials did have a trial registration completed, some of these were retrospective. The majority of these studies were not registered. They, alongside the other study types included in this review, underwent an assessment of risk of bias on the basis of analysis of intended outcomes reported in the final publication (prior to the ‘Results’ section). This most likely explains the seemingly ‘better’ completion status of those studies as compared to the trials with a registration or published protocol.

3.6 Conclusion

This review contributed to a longlist of outcome domains, laying the foundations for COS development. By determining the WHO ICF outcome domains captured by recent published literature in this field I believe that we have comprehensively accounted for the potential range of outcome domains pertinent to hand fractures and joint injuries from the perspective of professionals, including researchers and clinicians involved in research design. It has also served to reinforce the message that there is a problem in terms of outcome heterogeneity in this field, and a COS would be of benefit.

The next step was to address the patients' perspective, via formal and extensive exploration through interviews and focus groups with those who have first-hand experience of these injuries. This would also be an opportunity to establish whether different outcome domains were relevant to those with distal radius fractures as opposed to the other hand fractures and joint injuries.

The output of this systematic review is revisited in Chapter 5, which describes the Delphi consensus study.

Chapter 4 Qualitative Research – Interviews and focus groups

4.1 Introduction

Involvement of patients in COS development is important because it can lead to inclusion of outcomes which might have otherwise been neglected when relying only on specialist (e.g. clinician or healthcare professional) and researcher input. This was exemplified in the OMERACT group's 6th meeting, when patients highlighted 'fatigue' and 'disturbed sleep' as significant issues which were not accounted for in the COS for rheumatoid arthritis that had been in place for approximately 10 years (Kirwan et al. 2003). This led to a change in the COS to include 'fatigue' as a new outcome domain.

COMET strongly advocates exploration of the patient perspective during COS development (Williamson et al. 2017) and the COS-STAD recommendations include considering the patients' views when developing an initial list of outcomes (Kirkham et al. 2017). Based on a survey of COS developers (of both completed and ongoing COS projects), as part of their development process 64% (90/140) of COS studies had involved patient participation through individual interviews, focus groups or a combination of the two (Biggane et al. 2018).

While the systematic review shed light on the clinician and researcher perspective on outcomes of relevance to hand fractures and joint injuries, the patient voice is also important in the development of a longlist of outcome domains for use in the Delphi study. The output from this qualitative study was essential in providing this patient voice perspective. This study identified outcome domains based on the input of patients with direct experience of hand fractures and joint injuries. It was also an opportunity to help contextualise and understand why these outcomes are important from the patient perspective (Keeley et al. 2016).

4.2 Aims

The purpose of this step was to support COS development by identifying patient stakeholder opinions about pertinent outcomes of hand fractures and joint injuries and their treatment.

Primary objective: To identify a comprehensive list of outcomes domains relevant to the treatment of hand fractures and joint injuries by considering the patient stakeholder experience of hand fractures and joint injuries and their expectations of treatment and recovery.

Secondary objective: To consider variation therein, between patients who have sustained distal radius fractures and those who sustained other forms of hand fractures and joint injuries.

4.3 Methods

This was a qualitative study exploring individual experiences of hand fractures and joint injuries with focus upon the injury itself, the treatment and the recovery process. It involved sampling of patients with experience of an injury within the scope of the COS as specified in Section 2.1. This essentially includes fractures of any of the bones distal to and including the distal radius and/or ulna, as well as joint injuries (e.g. subluxation, dislocation, ligament sprain/rupture, mallet finger). Qualitative methods are well-established in the medical research world (Sbaraini et al. 2011). Alongside quantitative methods, they provide complementary insight into clinical experiences by enabling exploration and better understanding of context and implications (Malterud 2001). Both one-to-one interviews and focus groups were used in this study.

4.3.1 Interviews and focus groups

Individual interviews are the commonest data collection strategy in qualitative research (Nunkoosing 2005). Interviews allow us to find out that which cannot be directly observed, such as people's feelings, thoughts, past behaviours (Patton 2002). There are several forms of one-to-one interview (Bryman 2004):

- Structured or standardised interview – this involves an interviewer administrating an interview schedule, aiming to have uniform context of questioning and recording of answers
- Unstructured interview – this involves a typically informal approach, with the interviewer having a range of issues or topics that they typically cover but with phrasing and sequence of questions varying between interviews
- Semi-structured interview – this encapsulates a broad spectrum between the structured and unstructured interview formats. The interviewer asks a series of questions for a certain context, working through a topic guide or interview schedule but able to alter the sequence of questions. These questions will usually be more open-ended than those of a structured interview. Additional questions can also be posed based on interviewee responses

I selected the semi-structured interview format for this study as it allowed for the exploration of individual experiences (Flick 2018).

Focus groups are also commonly used in health research (Tausch and Menold 2016).

Focus groups enable within-group interaction, which can develop insights that cannot be reached through one-to-one interviews (Kitzinger 1995). For instance, it can be a way in which one can start to understand ‘why’ people feel what they do, as participants can probe and even challenge one another’s reasons for certain points of view (Bryman 2004). However, if used alone they can result in fewer outcomes being identified and difficulty in understanding the ‘patient journey’ as a by-product of each individual participant having less ‘speaking time’ than in a one-to-one interview (Keeley et al. 2016).

Given the various strengths and limitations of the data collection methods outlined above, I believed that the strongest approach for this study would be a mix of semi-structured interviews and focus groups.

4.3.2 Medium of interviews and focus groups

The face-to-face medium is generally regarded as the ‘gold standard’ for interviews (McCoyd and Kerson 2006). This enables gathering of information in the form of non-verbal cues and may also allow a building of rapport which is difficult to achieve without direct physical interaction (Flick 2018). Advancing technology has resulted in increased accessibility to equipment suitable for video conferencing amongst the wider population. Virtual interviews and focus groups allow combined video and audio communication between physically separated researchers and participants.

There are two broad categories of interviews as a whole – the asynchronous and synchronous interviews:

- Asynchronous interviews are not conducted in real-time, with researcher and participant taking turns to communicate. They typically involve text-only communication such as forum posts, emails or letters. Such a medium does not convey non-verbal cues. It can however reduce 'interviewer effect' wherein a perceived status difference between interviewer and interviewee might lead to an alteration of the information provided by the interviewee (Selwyn and Robson 1998). It might also enable shy participants to say more, or provide some participants time to carefully consider and word their responses (McCoyd and Kerson 2006). It has been commented that when participants are separated from researchers by significant distance then asynchronous interviews can save on cost compared to international telephone calls (Ratislavová and Ratislav 2014) but this is probably no longer as significant an advantage if participants are able to access the internet. Time can be saved in terms of avoiding transcription if an electronic text medium is used though, as data can simply be copied and pasted into interview transcripts (Ratislavová and Ratislav 2014).
- Synchronous interviews are conducted in real-time. Face-to-face interviews are synchronous, but alternatives range from pure text-based approaches, to telephone interviews, and finally videoconferencing 'virtual' interviews. Telephone interviews share many of the pros and cons of the asynchronous interview methods already outlined (Holt 2010). However, virtual interview is perhaps the medium which comes closest to replicating the face-to-face interview experience while still providing benefit in terms of cost-saving, convenience and accessibility (Hanna 2012; Iacono et al. 2016). There are potential downsides though – technical issues in audio or video quality might

impact on interaction and accessibility relies on participants having appropriate equipment and being able to take part in an online interview (Seitz 2015).

While many of the points summarised above for interview media carry over to focus groups, focus groups in particular aim to have group dynamics and interaction as part of the exploratory process (Kitzinger 1995). Comparison of online forum focus groups and face-to-face ones suggests that there is reasonable overlap in the data quality (Woodyatt et al. 2016). However, this remains a relatively unexplored area of research methodology, particularly within the setting of COS development.

The original plan for this study was to try to achieve face-to-face data collection where possible, with telephone interviews available as an option to be used sparingly when participants requested this. This option could facilitate participation in cases where some patients might have otherwise found it too inconvenient to take part (Heath et al. 2018). However, restrictions due to COVID-19 necessitated an amendment to allow virtual interviews and focus groups.

4.3.3 Recruitment

Recruitment was by purposive sampling (Lingard et al. 2008) from fracture clinics at Queen's Medical Centre, Nottingham. The fundamental eligibility criteria were:

- Inclusion criteria:
    - Have experience of a 'condition of interest' (detailed previously)
    - Adults (age ≥18 years)
    - Able to converse fluently in English

- Exclusion criteria:
    - Inability to give informed consent

Purposive sampling involved selection of patients with a range of demographics (both men and women, of various ages and occupations), injuries and treatment. I constructed sampling frames (Table 4-1 for interviews and Table 4-2 for focus groups) to broadly guide participant selection for each of two injury groups (distal radius fractures, and other hand fractures and joint injuries). I deliberately pre-defined injury groups within the sampling frames in order to enable comparison between these in the analysis. I felt this was important because distal radius fractures occur more frequently in older women, while the other hand fractures and joint injuries typically occur in younger men. There is a sex and age disparity which could affect the outcome domains considered relevant by each of the two injury groups. In line with the epidemiological trend in sex and age for these injuries as described in Chapter 2, the target numbers in the sampling frames prioritised older females with distal radius fractures, and younger males with other hand injuries. Despite this, the frames also attempted to guide recruitment of some interviewees and focus group participants from combinations such as males with distal radius fracture or older patients with other hand fractures or joint injuries, as these injury types are not exclusive to the highlighted demographic groups.

The frames also distinguished between conservative and surgical treatment as I believed this could be an important factor in the patient experience of these injuries. The target numbers for this category were kept approximately even.

Table 4-1 Target sample for interviews

Injury group	**Sex (n)**		**Age (n)**		**Treatment (n)**	
	Female	**Male**	**≤50 years**	**>50 years**	**Conservative**	**Surgical**
Distal radius fracture	7	3	3	7	5	5
Other hand fractures and joint injuries	4	11	11	4	8	7
Total	**11**	**14**	**14**	**11**	**13**	**12**

Table 4-2 Target sample for focus groups

Injury group	**Sex (n)**		**Age (n)**		**Treatment (n)**	
	Female	**Male**	**≤50 years**	**>50 years**	**Conservative**	**Surgical**
Distal radius fracture	7	3	3	7	5	5
Other hand fractures and joint injuries	3	7	7	3	5	5
Total	**10**	**10**	**10**	**10**	**10**	**10**

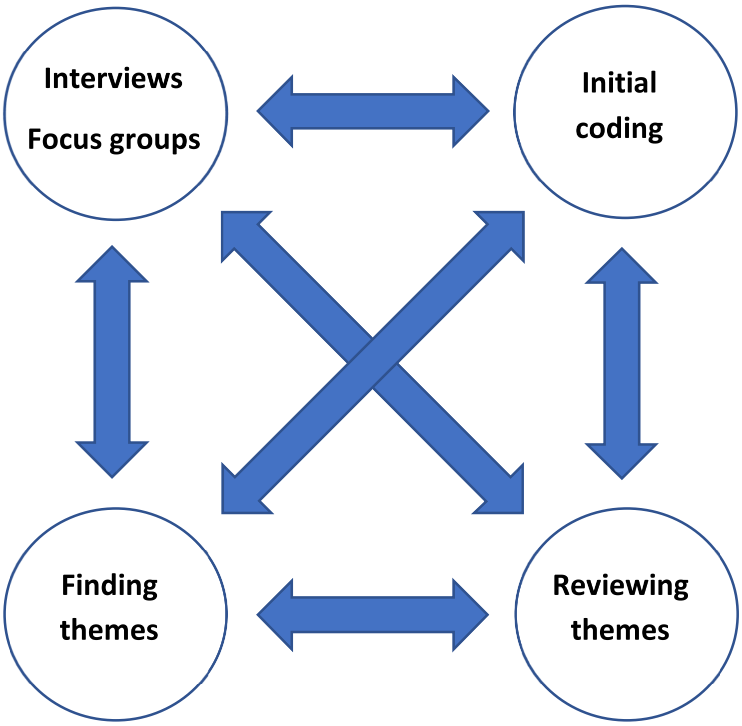
Despite guidance offered through the sampling frames the most important goal during recruitment was data saturation (the point at which no additional themes are generated from additional data collection and analysis). There has been some work exploring sample size requirements to reach saturation in interviews. Francis et al (2010) found that the number of new beliefs among interviewees on a couple of set topics started to plateau after about 6 interviews though to reach data saturation they suggested a minimum of 13. Guest et al (2006) found that after the first 12 interviews analysed, they had 88% of data saturation (defined as the output of analysing all 60 interviews). However, their group was fairly homogeneous. Given the various important factors incorporated in the SO-HANDI sampling frames (sex, age and injury group) I predicted that more interviews would be needed for the study and that after 20-30 interviews data saturation would be reached. This is in keeping with the range of sample sizes suggested according to a collection of expert opinions (Baker and Edwards 2012).

Estimating participant numbers for focus groups is more complex because there is an additional variable in the form of the specific mix of participants involved in each group. An empirical study found that over 80% of themes were found based on 2-3 focus groups, with three focus groups able to identify all of the most common themes within the data set (Guest et al. 2016). In that study focus groups consisted of 6-8 participants. As such, I set an approximate goal of a similar three focus groups, involving 6-8 participants each. The sampling frame was constructed with a target of 20 participants in total. In the focus groups the emphasis was not simply on generation of novel themes, but to clarify any uncertainties arising from the interviews.

4.3.4 Data collection and analysis

I utilised a constant comparison method with iterative cycles of data collection and analysis (Lingard et al. 2008), and refinement of emerging concepts through thematic analysis (Braun and Clarke 2006). Interview and focus group transcripts were coded and themes were found based on the initial coding. I then reviewed these themes to focus on outcomes from the patient perspective. The analysis of initial transcripts fed back into coding and even data collection, as sampling of further participants was shaped by the analysis (Figure 4.1). Initial individual interviews and focus groups therefore influenced subsequent data collection and recruitment as concepts and themes were explored (Lingard et al. 2008).

Figure 4.1 Constant comparison method incorporating thematic analysis



The interviews formed the main exploratory component of the work: to develop an in-depth understanding of multiple ‘patient journeys’ in the context of their injury, treatment and recovery. Focus group sessions functioned as ‘checkpoints’ to both validate outcome domains constructed from data from the interviews, as well as to explore uncertainties arising from them. They also enabled exploration of outcome domain wording, for use in subsequent consensus study work.

Whilst not planning a strict timeline, I expected to perform approximately five to ten interviews (to have data from participants across a range of injuries and demographics), then begin coding and analysis on these. I would then conduct further interviews, aiming for repeat review after roughly every five interviews. Focus group recruitment was planned to begin once the initial interview review was completed, but was delayed by the change to an online-only format.

4.3.4.1 Data collection

Interviews and focus groups were audio recorded, with subsequent production of anonymised transcripts for analysis purposes.

Interviews were semi-structured one-to-one discussions. A general topic guide outline was followed, organised in three parts: 1) experience of the injury and its initial impact, 2) experience of treatment (including recovery phase) and, 3) patients' thoughts on important outcomes that should be considered. Each interview was expected to last approximately 30-45 minutes. Participants were encouraged to introduce or focus on topics which they felt were particularly pertinent. The interview transcripts coding and analysis shaped subsequent data collection via adjustment in the prompts used or how questions were phrased, though remaining within the broad outline of the pre-determined topic guide.

Focus groups were intended to each include six to eight patients. After a brief presentation to provide study context, participants introduced themselves and were invited to state the impact of their injuries and experiences of treatment and recovery (in line with the interview topic guide components). Further discussion was shaped around these points, but also clarification of themes generated from analysis of previous interview and focus group transcripts. Each focus group discussion was expected to last 90-120 minutes. The focus group sessions were also an opportunity to generate a sense of simple consensus around wording of outcome domains.

Interviews were face-to-face, via telephone or virtual. All focus groups ran virtually.

4.3.4.2 Data analysis

NVivo v12 software was utilised in the analysis of transcripts. The overall methodological approach was an inductive, thematic one as described by Braun and Clarke (2006) – this consists of six stages which involve exploring the data 'as is' and then working in a systematic way to identify themes:

- Stage 1: Familiarisation with data – transcribing (some or all), reading and re-reading data

All interviews were conducted by a single researcher, who therefore gained an inherent basic familiarity with the data. Only a minority of transcription was outsourced, allowing first-hand familiarisation of the majority of data simply through transcribing. All transcripts were read and re-read to ensure sufficient familiarity with the complete data set.

- Stage 2: Generating initial codes – work systematically to code entire data set

Inductive coding was completed by the study researcher, with a supervisor experienced in qualitative research methodology ensuring validity and consistency of the coding.

- Stage 3: Searching for themes – collation of codes into potential themes

Initial codes were developed into themes reflecting the patient journey, singly in a few cases but more frequently through merging or grouping. The patient journey summarised the patient story gained through use of the topic guide, but was recognised to be a broad exploration of the injury, treatment and recovery.

- Stage 4: Reviewing themes – checking if themes work with coded extracts and across whole data set

The 'reviewing themes' phase was used as an opportunity to re-work the patient journey themes with a focus on outcomes. Some of the patient journey themes were of substantially less relevance to the overall study goal and only specific codes within these were retained. In some cases, reviewing of themes through the lens of outcomes clarified the need for further coding. This review stage culminated in the generation of a range of subthemes: in focusing on outcomes, these were synonymous with 'outcome domains'.

- Stage 5: Defining and naming themes – further analysis to refine, define and name themes

The subthemes were then grouped into themes which were subsequently defined and named. I also formed a short narrative description to explain the unifying concept within each theme.

- Stage 6: Produce report – selection of extract examples relating back to research question and reporting the analysis

The final stage is that of producing the report, while incorporating ongoing interpretation. The final narrative was validated by review by the research group.

As part of fulfilling the secondary aim of this study, ongoing analysis was used to check whether there were substantial differences in outcomes of interest to patients with distal radius fractures compared to those with other hand fractures and joint injuries. Some sense of this was achieved through initial coding and theme generation for the patient journey (Stages 1-3), but my focus was on whether different outcome domains (Stage 4) would be generated. I therefore compared outcome domains associated with the first few interviews from both injury groups. If major differences were identified, then the constant comparative approach would allow for subsequent data collection to evolve and explore these.

4.3.5 Ethics approval and patient and public involvement

The study protocol was reviewed and given favourable opinion by South Central (Berkshire B) Research Ethics Committee. A minor amendment was subsequently made to allow virtual interviews and focus groups due to the COVID-19 pandemic precipitating a lockdown in the UK.

In developing the study protocol and applying for ethics approval, patient and public involvement (PPI) work was completed to develop and refine both the overall study design and patient-facing documents (patient information sheets, consent forms, poster summarising the study). The interview/focus group topic guide was also discussed.

Feedback from patients was that this seemed to be an interesting study as it sought the patient’s viewpoint on the experience. Patients felt that the planned durations for interviews and focus groups were practical and should be feasible, although patients with occupations did mention that taking part once recovered to the point of being back at work might be a challenge. The option of a telephone or virtual interview was felt to be a reasonable workaround.

4.4 Results

4.4.1 Recruitment

Overall, 25 patients took part in the interviews (Table 4-3) and 21 participated in the focus groups (Table 4-4). Of these, 11 participated in both activities. In total five focus groups were conducted, with a range of three to six participants per group.

Table 4-3 Recruited sample for interviews

Injury group	**Sex (n)**		**Age (n)**		**Treatment (n)**	
	Female	**Male**	**≤50 years**	**>50 years**	**Conservative**	**Surgical**
Distal radius fracture	8	2	2	8	6*	5*
Other hand fractures and joint injuries	4	11	10	5	9	6
Total	**12**	**13**	**12**	**13**	**15***	**11***

*One patient had bilateral wrist fractures, experiencing conservative treatment of one side and surgical for the other. They are counted twice in the indicated numbers for treatment.

Table 4-4 Recruited sample for focus groups

Injury group	**Sex (n)**		**Age (n)**		**Treatment (n)**	
	Female	**Male**	**≤50 years**	**>50 years**	**Conservative**	**Surgical**
Distal radius fracture	8	1	2	8	5*	5*
Other hand fractures and joint injuries	6	6	7	5	9	3
Total	**14**	**7**	**9**	**13**	**14***	**8***

*One patient had bilateral wrist fractures, experiencing conservative treatment of one side and surgical for the other. They are counted twice in the indicated numbers for treatment.

4.4.2 Demographics

A summary of demographic characteristics of interview (Table 4-5) and focus group (Table 4-6) participants is shown below. These are divided in terms of participants with distal radius fracture (DRF) or other hand fractures and joint injuries (non-DRF) for each data collection activity. In terms of demographics, beyond the key characteristics of the sampling frames I also show the ‘work type’; during data collection and analysis it became apparent that this could influence the impact of the injury on participants.

Table 4-5 Demographics of interview participants by injury group

Characteristic	**DRF Injury (n=10)**	**Non-DRF Injury (n=15)**
Age	Mean: 58 years	Mean: 42 years
	Range: 22-77 years	Range: 21-68 years
Gender	8 F : 2 M	4 F : 11 M
Occupation type	None/retired: 5	None/retired: 2
	Office/non-manual: 4	Office/non-manual: 5
	Light manual: 1	Light manual: 6
	Heavy manual: 0	Heavy manual: 2
Intervention	Conservative Rx: 5	Conservative Rx: 9
	Surgery: 4	Surgery: 6
	Both: 1 (bilateral DRF, one treated conservatively and the other surgically)	Both: 0
Duration between injury and interview	Median: 6.5 weeks	Median: 6 weeks
	Range: 3-9 weeks	Range: 2-60 weeks
Interview medium	Face-to-face: 3	Face-to-face: 9
	Telephone: 6	Telephone: 4
	Virtual: 1	Virtual: 2
COVID-19 lockdown phase	Pre-lockdown: 3	Pre-lockdown: 9
	During: 7	During: 3
	Post-lockdown: 0	Post-lockdown: 3

The time interval between injury and interview had similar medians for the two injury groups. The mean duration of interviews was 45 minutes, ranging from 20 to 79 minutes.

Table 4-6 Demographics of focus group participants by injury group

Characteristic	**DRF Injury (n=9)**	**Non-DRF Injury (n=12)**
Age	Mean: 62 years Range: 22-75 years	Mean: 45 years Range: 21-68 years
Gender	8 F : 1 M	6 F : 6 M
Occupation type	None/retired: 5 Office/non-manual: 3 Light manual: 1 Heavy manual: 0	None/retired: 2 Office/non-manual: 4 Light manual: 6 Heavy manual: 0
Intervention	Conservative Rx: 4 Surgery: 4 Both: 1 (bilateral DRF, one treated conservatively and the other surgically)	Conservative Rx: 9 Surgery: 3 Both: 0
Duration between injury and interview	Median: 8 weeks Range: 3-23 weeks	Median: 5.5 weeks Range: 2-31 weeks

Mean interval between injury and focus group session was 11.1 weeks for the DRF participants and 12.8 weeks for the non-DRF participants.

The mean duration of focus groups was 91 minutes, ranging from 78 to 110 minutes.

In the rest of this chapter, the convention for participant identification follows the format of: distal radius fracture (R) or other hand fractures and joint injuries (H) + participant ID number, sex, age group (≤50 or >50 years), interview or focus group.

Where examples of codes are cited, quantitative data are not provided because they did not allow for meaningful interpretation for the following reasons:

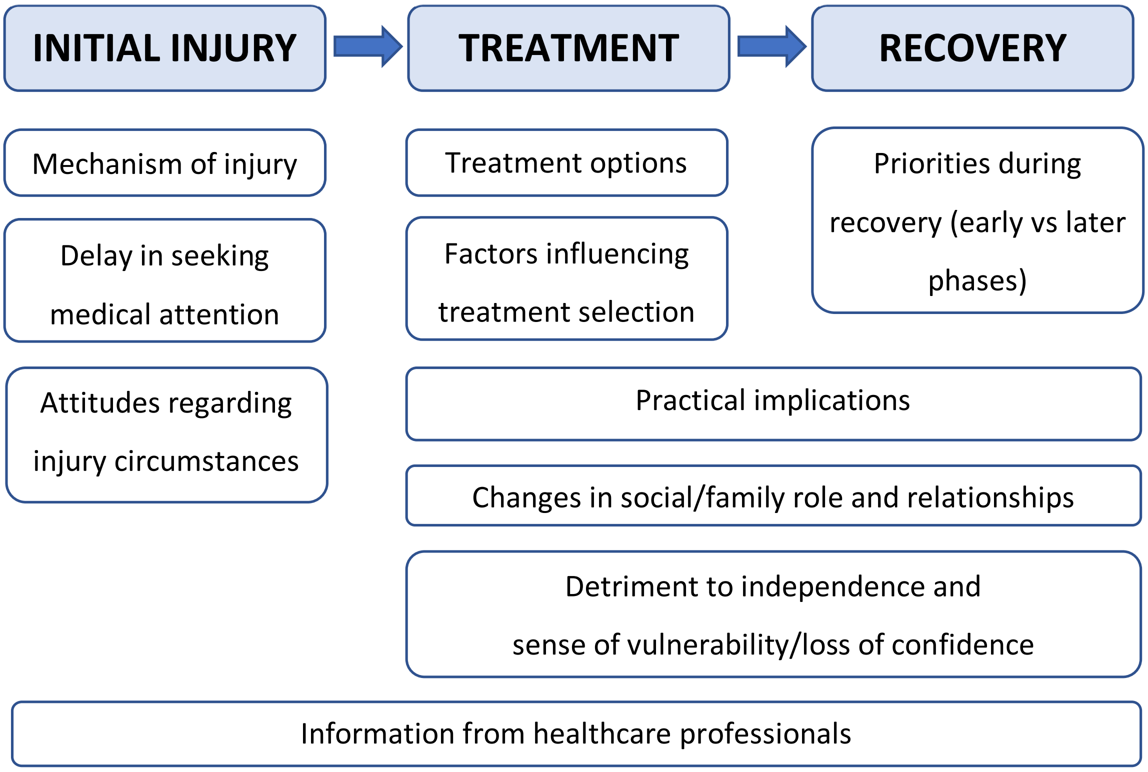
1. Raw frequency of code occurrence might be misleading as there could be multiple data extracts for a given code from a given participant (whether in an interview/focus group session).
2. Number of participants in which codes occurred was not necessarily useful as coding was done across the data set, combining interviews and focus groups; in groups the discussion of certain codes could have been prompted by the facilitator or one participant and led to a discussion involving other participants who might not have spontaneously raised the topic themselves. Also, some participants took part in both activities.
3. The primary aim of this study was to identify a longlist (i.e. breadth) of outcome domains. Therefore, I believe it was important not to succumb to a positivistic stance which concerns itself with the quantitative data related to code occurrence. Instead, the narrative offers a semblance of this through use of adjectives such as ‘some’, ‘many’, etc, while highlighting instances where a finding occurred ‘rarely’.

4.4.3 Descriptive analysis: The patient journey

This section summarises the key themes initially identified as relating to the participants' overall lived experiences of sustaining, being treated for, and recovering from, their hand fractures and joint injuries. The analysis corresponds to Braun and Clarke (2006) thematic analysis stages 1, 2 and 3.

This phase of finding themes was structured around three broad topics: initial injury, treatment and recovery. I identified 10 themes relating to the patient journey, and the topic(s) under which they fell is shown in Figure 4.2.

Figure 4.2 Broad patient journey topics and associated themes



The themes along with example codes are shown in Table 4-7.

Table 4-7 Patient journey themes and examples of associated codes

Patient journey theme	**Examples of initial codes**
Mechanism of injury	Slip or trip; Football; Hockey; Snowboarding; Carrying heavy objects; Cycling
Delay in seeking medical attention	Minimisation of the injury; Pain; Use of healthcare-resources; Coronavirus-related
Attitudes regarding injury circumstances	Regret; Frightened or scared of returning to area or activity of injury
Information from healthcare professionals	Diagnostic process; Accessibility to further information; Movement range and stiffness; Conflicting information; Concern over worsening the injury
Treatment options	Treatment – surgery; Treatment – splint or cast; Treatment side effects or complications
Factors influencing treatment selection	Treatment side effects or complications; Post-traumatic arthritis; Pain; Time to return to work; Deformity; Coronavirus-related
Practical implications	General household chores; Driving; Eating and drinking; Dressing or undressing yourself; Holding light objects; Work impact; Splint or cast hindering function
Changes in social/family role and relationships	Spouse or partner emotional impact; Children emotional impact; Reliance on others; Sexual activity
Detriment to independence and sense of vulnerability/loss of confidence	Reliance on others; Time to return to normal day-to-day life; Washing yourself, Toileting; vulnerability or loss of control; Desire to maintain or regain independence
Priorities during recovery (early vs later phases)	Change in priorities at differing phases of recovery

4.4.3.1 Mechanism of injury

Learning how the injury was sustained provided context for each participant's story and highlighted activities they had been engaged in and might therefore place some importance in returning to.

Some participants sustained their injury in the process of bracing against a fall whilst walking: *"I slipped. My foot went from under me, fell backwards, put my arm down"* (H051, female, ≤50 years, interview).

Others were injured during work-related activity: *"[O]ne of my jobs at the end of the night is to empty the bottles, which are in a standard crate…the crate slipped and I got my thumb stuck under the rim"* (H004, female, >50 years, interview); *"I'm a police officer. We were responding to a job and we're in convoys…Essentially, the car in front has gone to take a right-hand turn… and we've basically collided with the other vehicle"* (H041, male, ≤50 years, interview).

For most participants though, the injury occurred as a result of recreational activity: *"I'm a keen runner so I was out running and just tripped"* (H049, female, >50 years, interview); *"I broke my wrist…at the gym. I was in a class and I fell off the box"* (R028, female, ≤50 years, interview); *"I came off a skateboard. So basically all my weight just landed on my right hand as I fell backwards"* (R044, male ≤50 years, interview).

Non-DRF injuries were sometimes the result of intentional violent activity: "*I pun*ched *a wall…and noticed straightaway that I was missing a knuckle"* (H012, male ≤50 years, interview). These instances all occurred in young, male participants.

4.4.3.2 *Delay in seeking medical attention*

Some participants with non-DRF injuries deferred seeking medical attention, initially believing their injury to be fairly trivial. A few even went on to complete the activity that had led to their injury:

> *"Field hockey…I got hit on my right thumb by the ball…carried on playing. It hurt a bit and then it swelled up during the game…the following morning…sought medical treatment"*
>
> (H015, male ≤50 years, interview)

A few assumed symptoms would subside of their own accord: *"2 or 3 weeks I left it, because I thought it would just go away"* (H003, male, ≤50 years, interview); *"I didn't go to the hospital until four days afterwards. I just thought it was badly bruised"* (H132, female, >50 years, focus group).

Others cited concerns about placing unnecessary burden on A&E: *"I thought it was just bruised so I didn't go to A&E. You're always being told not to go to A&E unless it's life-threatening aren't you?"* (H004, female, >50 years, interview).

All patients with DRFs sought medical attention within hours of their injury: *"I phoned my daughter…[she] was able to come from school and take me to A&E"* (R018, female, >50 years, interview); *"I live on my own so I shouted over the fence to my neighbours and they got into full gear straightaway. Phoned 111"* (R047, female, >50 years, interview).

4.4.3.3 *Attitudes regarding injury circumstances*

Many participants indicated a degree of regret about their actions leading up to the injury: *"I'm currently having lots of work done in my garden and…there was some work that I* *wanted to do. Which clearly now I wish I hadn't."* (H009, m*ale,* >50 years, interview); *"I broke my wrist, it's a stupid accident which I could kick myself for, and then I've had to go through all this because I didn't take care"* (R112, female, >50 years, focus group).

Some participants expressed fear or anxiety about being in a similar scenario, or performing the same activity, which resulted in their injury: *"I'm not very confident with going back cycling until I'm quite safe and healthy so I could handle what happens during cycling"* (H051, female, ≤50 years, interview); *"The place where I slipped, I do not want to go on to that part of that area again"* (R105, female, >50 years, focus group).

Others felt it important to confront their fears. Perhaps the most striking example of this was a trainee police officer injured in a road traffic accident:

> *"Obviously driving was quite a biggie for me…I have started driving now…So I'm slightly happy about that because that's a big thing. Me thinking 'I'm never gonna be able to drive again' to actually doing it…I've been back to where the accident happened, just so I could not have an issue with that…I wasn't going to have some sort of breakdown or whatever, or get upset."*
>
> (H041, male, ≤50 years, interview)

It was clear that for some participants the outcomes they considered relevant or a priority were shaped from the moment the injury occurred, such as through the injury circumstances.

4.4.3.4 Information from healthcare professionals

A number of healthcare staff typically interact with patients through their journey, including healthcare assistants, nurses, doctors, physiotherapists and occupational therapists. There were many facets to the information provided by healthcare professionals, from diagnosis to treatment to recovery process: *“I was taken to the walk-in centre on the Sunday…I was told the extent of the injury that I had and…that I’d got to go to the fracture clinic on the following morning to have it put in a cast”* (R118, female, >50 years, focus group).

Many patients felt that they had adequate information from healthcare professionals regarding treatment, recovery and healing outcomes: *“I was quite happy with the treatment that I got and the advice that I got on how to create a recovery and the healing”* (H101, male, >50 years, focus group); *“They told me everything that they were gonna do”* (R038, male, ≤50 years, interview).

However, some felt that information was somewhat lacking in specific areas such as impact to healing and what progress towards functional recovery they could expect:

> *“It might have been helpful to have those expectations initially of…‘in immobilising the digit in question it is likely to have limited range of motion when you first come out from the splint’ and that ‘it’s likely to remain painful for at least the same period again as long as you’ve been in the cast’.”*

(H015, male, ≤50 years, interview)

Participant H015 also presents an example where outcomes being prioritised were revealed by the specific queries raised by participants:

> *“The consultant pretty much said, ‘a couple of weeks after you’ve had the cast off’, that I could do pretty much anything as long as I was being careful. So hockey I was advised to wear a protective glove in future.”*

(H015, male, ≤50 years, interview)

4.4.3.5 *Treatment options*

In the majority of cases participants stated that there was in fact a strong recommendation from the doctor for one option over another: *"I was told…by a consultant that it's not going to heal. Absolutely not going to heal…without surgery"* (H003 male, ≤50 years, interview); *"they said it looked like it would heal quite well but if it didn't they would have to…put a metal plate in"* (R028, female, ≤50 years, interview).

It is noteworthy that some participants preferred not having to decide between options: *"'This is what will happen!'...There was no discussion about how this might affect my life. I was told what was going to happen and I was quite happy with all of that"* (H017, female, >50 years, interview) and:

> *"It was basically, 'this is what we're doing next, this is what will happen after that'. I'm glad they didn't give me options. Just tell me what I need to do, where I need to go, what you have to do. It makes me feel so much better cause I wouldn't be able to choose."*
>
> (R028, female, ≤50 years, interview)

4.4.3.6 *Factors influencing treatment selection*

For many participants a strong recommendation from a healthcare professional proved decisive in influencing treatment selection: *"If I was told, 'you need it operating on', then yes I would accept that"* (H049, female, >50 years, interview).

Several participants were concerned by potential long-term consequences such as post-traumatic arthritis and long-term pain:

> *"For me the important outcome in this was getting it fixed and getting it fixed properly and to get the union…I would be OK with the treatment taking as long as the treatment needs to take, as long as the outcome is 'I get a union and I don't get arthritis'"*
>
> (H130, male, ≤50 years, focus group)

For others, return to hobbies motivated their treatment selection: *"So the factors I found really important supporting that decision: number one is to get my finger back to normal…I take class for piano. I would like to go back to piano practice in the future"* (H051, female, ≤50 years, interview).

The time taken to return to function was a priority for several participants: *"Recovery time would have been a deciding factor. And I'd have gone with whichever one was shorter"* (R029, female, ≤50 years, interview); and:

> *"**Interviewer:** Are you saying that if there was one treatment that got you back to doing things sooner, but maybe had some long-term problems in your later life, you would take that treatment?*
>
> ***H103:** That sounds bad! Possibly…it's possibly a risk you take isn't it because it might not have an effect in later life? It's just one of those chances."*
>
> (H103, male, ≤50 years, focus group)

Participants described considering risks and complications, and potential impact of the options on their lives in the process of decision-making: *"I think obviously the first thing is to know…exactly what the two options are. And what the impact of those options are. And the risks…So it's to do with risks, the healing and impact"* (R019, female, >50 years, interview).

Only a small minority of participants mentioned that the final appearance of the injured body part influenced their choice of treatment: *"He said it would be fine either way but the cast would probably…make it a bit straighter…which I said, 'I may as well'."* (H012, male, ≤50 years, interview).

Finally, the COVID-19 pandemic was a factor in some participants' treatment selection. One explained reluctance for surgery as this would mean yet another visit to hospital:

> *"The problem for me was that this all hit me right at the start when the epidemic was getting bad. So for me it was extremely stressful because I didn't want to go into hospital, I didn't want to go out. I certainly didn't want to go and have surgery."*
>
> (R109, female, >50 years, focus group)

4.4.3.7 Practical implications

Participants described a vast range of function-related issues arising from their injuries and the treatment they underwent. Many noted difficulty in accomplishing daily tasks such as household chores, meal preparation, eating and drinking and personal care (e.g. washing, dressing): *"I can't do much…housework and all that stuff has been out the window for the past two months"* (H050, male, ≤50 years, interview); *"at the minute we've just been eating like ready meals because it's so convenient because I can't like chop food…I can't do onions or anything"* (H117, female, ≤50 years, focus group); *"food-wise I'd kind of been on a restricted diet if you like…whatever you can just use one hand to eat"* (R044, male, ≤50 years, interview) and:

> *"The male dress code of needing to do shirt buttons was somewhat difficult. Yeah,*

tying one's own shoelaces became an interesting task"

(H015, male, ≤50 years, interview)

Beyond what might be considered 'basic activities', many also highlighted driving as an issue. The inability to drive had numerous downstream impacts affecting work, socialising and shopping: *"If we hadn't have been in lockdown I wouldn't have been able to work, because I wouldn't have been able to drive…and I have to drive for my job"* (R110, female, >50 years, focus group).

Hobbies and recreational activities were impacted for many participants: *"I wanted to go back to the gym where I broke it, but couldn't do that…I can't do any sort of like planks or push-ups cause I can't hold myself up. I wanted to get back to yoga but I can't"* (R107, female, ≤50 years, focus group); *"going to the gym and kettle-bell and my sports stuff is out the window...can't…go in the sauna or swim in the swimming pool"* (H002, male, ≤50 years, interview).

Several participants treated with a cast or splint expressed how this was a hindrance to their function through its bulky nature, restriction of movement or due to having to avoid getting it wet: *"I've definitely fallen behind at work because I cannot complete my jobs at the normal rate…the cast is quite bulky. Gets in the way a lot and it limits the movement of my wrist"* (H003, male, ≤50 years, interview); *"every time you have a shower, having to put the black plastic bag over it makes life difficult obviously"* (R044, male, ≤50 years, interview).

Dressings, casts and splints also required maintenance: *"The only time I've called on assistance from my wife…is the removal of this <indicates splint> and the cleaning of it and that sort of thing"* (H009, male, >50 years, interview).

Inconvenience or difficulty with maintaining their splint even led a few participants to deviate from their treatment regimen:

> *"As the weeks progressed it just got worse and worse. And then the tape was getting loose…three days before the appointment that I was due back…I'd taken it off myself because it had come loose"*
>
> (H016, male, ≤50 years, interview)

These various practical limitations engendered a variety of emotional impacts. The majority reported a sense of frustration: *"It's just a bit frustrating and it's limiting what you can and can't do"* (R125, male, ≤50 years, focus group); *"It's been really frustrating. I mean really frustrating 'cause there's nothing else to do and this is a time when I could so be spending a lot of time doing my hobbies."* (R047, female, >50 years, interview) while some felt low mood: *"feeling slightly down because I've got things I can't do"* (H113, male, >50 years, focus group) and, "I have good days and some really bad days I think. Some days I don't want to get out of bed cause I've got no motivation" (H041, male, ≤50 years, interview).

4.4.3.8 Changes in social/family role and relationships

For many participants, there was significant impact to life roles relating to immediate family and work: *"The worst was not being able to help out with my son as much as I normally would"* (H002, male, ≤50 years, interview); *"I had to…help [mum] off the bed and…onto a commode at one point…I managed to help her but it was a strain"* (R019, female, >50 years, interview); *"I got to the stage where I missed work. Who'd have thought of that!"* (R036, female, >50 years, interview).

Social roles relating to wider family and friends were also affected for some:

> *"I've got grandchildren…so normally I would drive up to see them quite often, which I couldn't do cause I couldn't drive and I wouldn't be able to hold them and things…meet up with friends and everything which…I couldn't drive anywhere to meet up with anybody."*
>
> (R036, female, >50 years, interview)

Although for others the key issue impacting on this was the COVID-19 lockdown preventing socialisation:

> *"I'd be out and about with my friends, I'd be going to garden centres. Play days, go to National Trust properties. I normally play bridge…twice a week. I would also go to art club at least once a week. But I can't actually blame that on my wrist…I'm like everybody else, you know we can't do it at the moment."*
>
> (R045, female, >50 years, interview)

Effects on their life roles precipitated a range of emotional impact, perhaps most striking in those who had a carer role that they were struggling to fulfil: *"[M]y daughter [is] 16 weeks old and I don't really have much to do with her…it's bad that I'm her father and I don't really hold her much…it depresses me"* (H042, male, ≤50 years, interview).

Some participants raised the prospect of an emotional impact on others, such as friends and family. One mother described shielding her young child from anxiety:

> *"I didn't get upset and I didn't complain about it hurting…I was conscious of not doing that in front of [my daughter]. I didn't cry or anything like that…not make a big fuss…I'm conscious not to project any anxiety onto her."*
>
> (R029, female, ≤50 years, interview)

Some participants felt that their relationships with others had an altered dynamic. This was negative for some: *"The first few days were slightly, slightly frustrating for both of us. Because I'm saying, 'no you need to hold it like that', and I couldn't show him because I couldn't hold it with my right hand"* (R018, female, >50 years, interview).

However, for others it was positive: *"Relationship-wise it's probably brought us closer together rather than pushing each other away I think"* (H041, male, ≤50 years, interview); *"I've heard from them more. I've got two boys…when I was really low and really fed up, they were taking turns to ring every day"* (R045, female, >50 years, interview).

Any effect on intimate relationships was brought up only rarely, and by younger participants:

> *"I will say just as an aside, as a self-esteem issue and an impact during the treatment, the ability to maintain hygiene around a cast is a thing…If you think about intimate personal relationships…when you've got a big stinking cast on your arm, you don't become a very appealing person to be around do you?"*
>
> (H130, male, ≤50 years, focus group)

While most participants did not mention any concern about the long-term appearance of their injured hand or wrist, there were a few for whom this was of some significance. This again tended to be the younger patients; there was only one DRF participant who mentioned a concern about the appearance, and she was the youngest in that injury group: *"The one thing…I have got this kind of lump here <indicates ulnar side of wrist> that's…still…that slight lump there"* (R028, female, ≤50 years, interview).

4.4.3.9 Detriment to independence and sense of vulnerability/loss of confidence

Many participants described examples of their reliance on others: *"I've been taken good care by my boyfriend…daily chores like doing the hoovering, washing and empty the bin. And carrying things while shopping"* (H051, female, ≤50 years, interview); *"until I had me pot off, every morning my wife had to virtually dress me…things like carrots, potatoes and particularly meat, my wife had to cut it up into squares"* (R038, male, >50 years, interview).

For some there was a sense of loss of independence which elicited significant emotion and impacted on their sense of dignity: *"We did go to a restaurant and [my husband] did have to cut my food up for me…which was a bit embarrassing"* (R029, female, ≤50 years, interview) and, "It's just…because you're independent…you wanna do it and the frustration kicks in when somebody asks to help. They only doing it very kindly. But that kindness sometimes wears a little thin" (H101, male, >50 years, focus group).

Several expressed a desire to maintain or regain their independence, with some persevering quite hard to try to achieve this goal:

> ***"H050:*** *You do adapt. You find new techniques and styles. Like I'm using my mouth a lot which isn't good at the moment because of the COVID stuff.*
>
> ***Interviewer:*** *Yeah…what do you mean? Using your mouth to do what?*
>
> ***H050:*** *Like pull your socks up or when you're ripping a letter or something. I didn't have the grip in my hand."*
>
> (H050, male, ≤50 years, interview)

Some noted an issue of loss of confidence, a few more specifically with mobilising: *"I couldn't walk the dog because I just didn't feel me balance in a way. I know it sounds daft but I thought if he pulls and I'm off…"* (R036, female, >50 years, interview).

Others commented on a reliance developed on the support provided by splints or casts, and the lack of confidence at the end of treatment when trying to function without this support: *"I spent all that time in the cast saying, 'oh my God I want this off, I want it off, I want it off'. The second I left the building I really wanted to go back into clinic and say, 'can you put one on please?'"* (H042, male, ≤50 years, interview).

Some participants highlighted the sense of vulnerability introduced simply by having sustained their injury: *"I've gone through 60-odd years without having to deal with anything else before this…get a wake-up call now and again…"* (H009, male, >50 years, interview).

4.4.3.10 Priorities during recovery (early vs later phases)

Across most of the interviews, it became apparent that in many cases patients' priorities shifted over the course of recovery. The following exchange was the most succinct example where the matter was explored (the specific priorities differed between interviewees). It also reflected the link between data collection and analysis in a constant comparison method, as data interpretation and analysis enabled discussing and testing of the theory:

> ***"Interviewer:*** *If you had to pick one or two things then, that you were really keen to get back to…what was the key thing for you?*
>
> ***R018:*** *Well the gardening…sawing tree branches and pruning with the big heavy pruners…it's gardening and it's cooking with lifting things in and out of the oven.*
>
> ***Interviewer:*** *These are your priorities as you are now…what about when you were at your worst in terms of how much you could do, so say 3, 4 weeks into it?*
>
> ***R018:*** *The priorities then were getting dressed, I struggled with that, and I couldn't put my bra on…So yes, sort of two or three weeks on it was the getting*

> *dressed…having a wash…Having a shower…I 've got long hair and I need two hands to get through my hair…*
>
> ***Interviewer:** OK a few days into having the cast on, what were your priorities then compared to the 2 or 3 weeks you just told me?*
>
> ***R018:** Well the priorities were meals, because I couldn't do them. Getting dressed, they were still the key things.*
>
> ***Interviewer:** They were still the key things? I wondered whether you would think pain was more of a priority in the first few days because it sounded like it was quite bad for you.*
>
> ***R018:** Oh pain, it was! I did have a lot of pain."*
>
> (R018, female, >50 years, interview)

This example also highlighted an issue when directly enquiring about participants' priorities – although having made it clear earlier in the interview that pain was a major issue initially, the participant did not volunteer 'pain' as a priority when asked specifically about priorities a few days into the treatment. However, on being reminded of her earlier statements, she agreed that 'pain' was a priority concern of hers early on – as it was for many others.

In some cases though, what might be only an early priority for many others can persist as a long-standing one. If 'pain' persists then there may not be progression to other priorities, as was the case with the interviewee at the longest follow-up in the study 60 weeks after his injury. For this participant the ongoing issue of pain impacted on his psychological wellbeing, to the point where he had at times considered extreme measures might be necessary:

> *"**H042:** When I've been going through everything and all that lot and I'm having pains and stuff, I've actually considered asking the doctor to take me hand off. I'll be honest with you.*
>
> ***Interviewer:** To do what sorry?*
>
> ***H042:** Take the hand off.*
>
> ***Interviewer:** Really?*
>
> ***H042:** Yeah, yeah…I want to keep me hand. But I do get to that point where I'm just like, 'you know I'd just be better off if it wasn't there'."*
>
> (H042, male, ≤50 years, interview)

4.4.3.11 Summary

In exploring and developing the patient journey themes I drew upon the experiences of many participants. As evidenced by the data presented, patients' journeys were complex and varied. They all began with an injury that deprived them, to varying degrees, of the ability to do what they would normally take for granted and this had corresponding psychosocial impacts. The tone of the ensuing treatment and recovery phases was set by the likelihood and time course of 'getting back to normal'. There was not a clear divide between patients based simply on demographics, with most of the various themes affecting patients regardless of age or sex. More telling perhaps were the basic functional needs as well as higher needs based on specific life roles that each individual considered their norm and wanted to be restored to.

Interestingly, the patient journeys were broadly similar for those with distal radius fractures as compared to the other hand fractures and joint injuries, with the key difference being in the theme of ‘delay in seeking medical attention’.

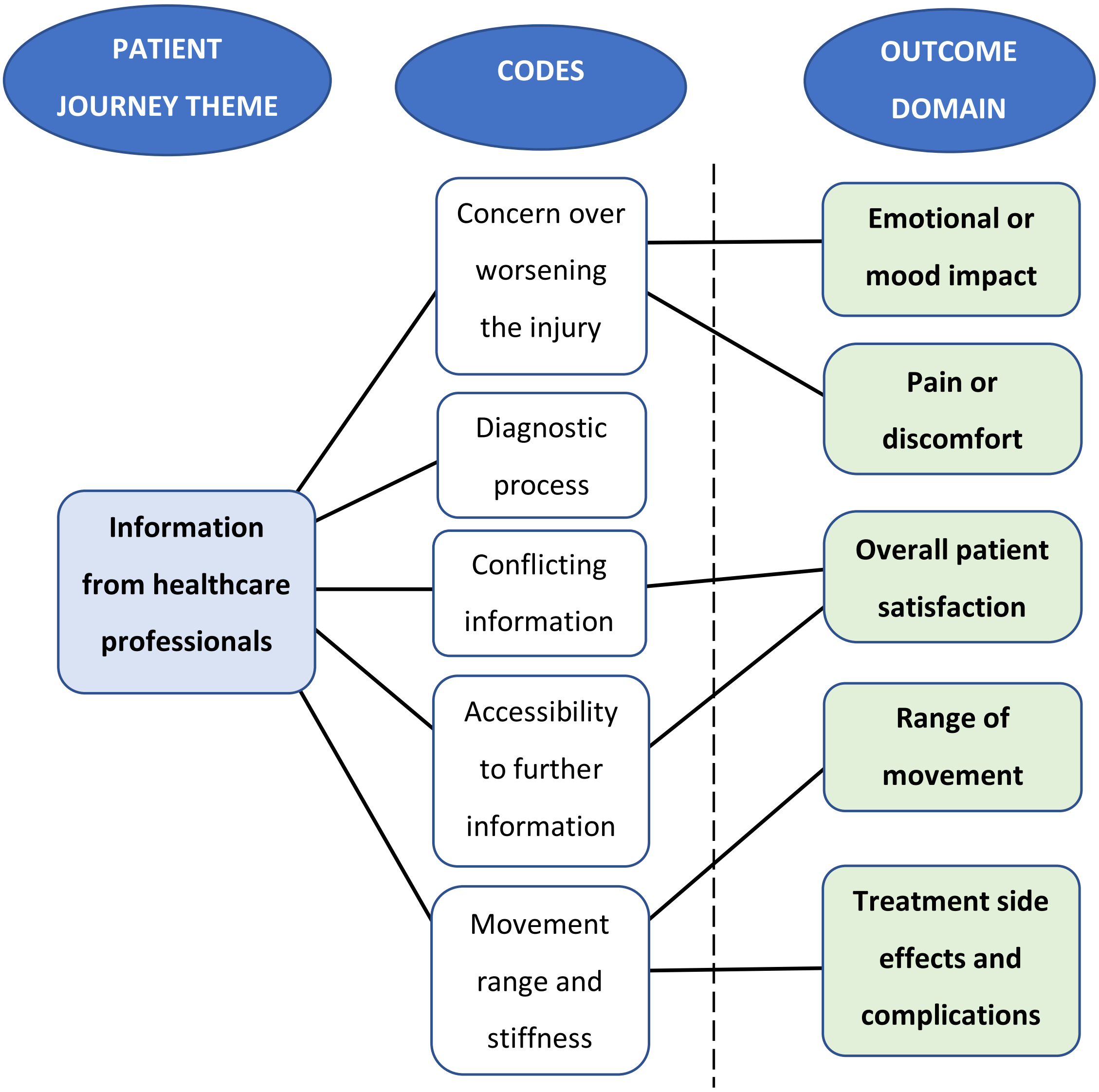
4.4.4 Interpretive analysis: Themes related to outcomes

Developing the patient journey themes allowed generation of a narrative that broadly captured the participants’ experiences of hand fractures and joint injuries. However, in order to focus on outcome domains relevant to these injuries further analysis and interpretation of the data was necessary. This process mainly corresponded to Braun and Clarke (2006) thematic analysis stages 4 and 5. I filtered the codes to focus on those with some relation to outcomes – those meeting this criterion were then collated within discrete concepts (i.e. outcome domains).

Figure 4.3 shows a very simplified example of the process with some of the codes from the ‘information from healthcare professionals’ patient journey theme. ‘Concern over worsening the injury’ was a code that overlapped with the codes ‘pain’ and ‘emotional impact’, as participants were seeking information from healthcare professionals to clarify whether their pain meant the injury was deteriorating in some way. Collation of codes (along with others not shown in the figure) generated outcome domains – in this case the data extract was linked to two separate outcome domains: ‘emotional or mood impact’ and ‘pain or discomfort’.

This process was repeated across the narrative and a summary of the links between patient journey themes to outcome domains is presented in Appendix 5. I generated 35 outcome domains which captured all outcomes-related coded data from the patient journey narrative.

Figure 4.3 Schematic depicting the process of transitioning from patient journey narrative themes to outcome domains



I grouped the outcome domains within themes, generating six outcome themes in total with discussion points raised in focus group sessions often helping to shape the process of outcome domain grouping. The outcome themes, and the outcome domains they encompass, are summarised in Figure 4.4 and described in the rest of this section.

Figure 4.4 Schematic summary of outcome themes and their associated outcome domains

Disruption to biology and physiology

- *Change in sensation*
- *Pain / discomfort*
- *Strength or muscle power*
- *Range of movement*
- *Confirmation of bone healing and alignment (for fractures)*
- *Confirmation of soft tissue healing*

Impact on self-care and independence

- *Eating and drinking*
- *Self-hygiene and personal care*
- *Cooking / preparing meals*
- *Grasping and moving light objects*
- *Walking / mobility*
- *Using hands to change body position*
- *General household chores*
- *Shopping for groceries*
- *Sleeping*
- *Dressing*

Return to wider social roles, engagement and responsibilities

- *Return to work / job*
- *Return to driving*
- *Fine hand use*
- *Hobbies and recreation*
- *Socialising / relationships with friends and family*
- *Carrying heavy objects*
- *Sexual activity*
- *Looking after dependents within household*
- *Keyboard and mouse / touchscreen use*
- *Fatigue / tiredness*

Emotional impact and psychological wellbeing

- *Emotional / mood impact to self*
- *Emotional / mood impact to friends and family*
- *Overall patient satisfaction*
- *Appearance*

Impact to personal and healthcare resources

- *Loss of income*
- *Added expenses due to injury*
- *Healthcare resources used*

Treatment factors and complications

- *Treatment convenience / comfort*
- *Treatment side effects or complications*

4.4.4.1 Outcome theme: Disruption to biology and physiology

Participants all experienced symptoms of a physical or organic nature as a result of their injury. Though the severity and nature of symptoms varied, descriptions and associated concerns coalesced around discrete outcome domains which were unified in having a physical manifestation. On a superficial level it might seem that some of these domains are so closely linked to function that they could simply be subsumed into one of the themes presented later in this chapter. However, it became evident that this category of outcome domains was seen as something quite distinct to ‘real world’ functional outcomes. As an example, one participant found value in measuring range of movement as an outcome distinct from whether or not the movement was used to ‘do something’:

> *“By seeing what a lady could do that she couldn't do, makes you realise that you can, by physio and patience, get to a certain point of recovery. And that’s a physical sign people can see, that you can think, ‘well, that's happened to them. It can happen to me if I just keep doing what I'm doing’…So being given certain measurements…just because you don’t use that movement every time you do something…It's a physical sign of how you're progressing.”*
>
> (R109, female, >50 years, focus group)

In a similar way, other outcome domains of a fundamental nature have been grouped within this theme of disruption to normal biology or physiology.

4.4.4.1.1 Outcome domain: Change in sensation

Descriptor: Change from normal feeling in the injured hand/wrist

Example codes: numbness; pins and needles; tingling; altered temperature sensation; hyperaesthesia

This domain includes all manner of change in the normal sensation provided by the hand or wrist. Affected participants were keen to learn what the long-term outcome would be: *"I had numbness around my finger…so I want to know what's going to happen"* (H111, female, ≤50 years, focus group); *"it's like…it doesn't feel natural…but I'm hoping that will get better over time"* (R036, female, >50 years, interview).

4.4.4.1.2 Outcome domain: Pain / discomfort

Descriptor: Discomfort or pain in the injured hand/wrist

Example codes: pain; dysaesthesia; discomfort or ache

This domain was drawn together from multiple points of the patient journey, from the time of injury through to ongoing pain during treatment and recovery phases. Participants experienced pain of varying severity and character, and pain was the underlying reason behind many instances of practical limitations. It also influenced treatment selection: *"I don't want that pain in like my whole hand in the long run so I'm hoping that they give me an option to re-x-ray, see if it's healing and if it's not healing then maybe surgery if it needs it"* (H117, female, ≤50 years, focus group).

Pain took an emotional toll on some – an extreme example was highlighted in Section 4.4.3.10 with participant H042 considering an amputation of his injured hand to resolve his pain issues.

4.4.4.1.3 Outcome domain: Strength or muscle power

Descriptor: Being able to exert force through the injured hand/wrist

Example codes: grip strength; pinch strength; pronation-supination strength

Participants recognised that their strength was impacted by their injury, and often through the immobilisation required to treat their injury. Regaining strength was a fundamental aspect of their recovery to enable general use of the hand or wrist: *"[M]y strength is just nothing like what it was before, so it's just a bit frustrating and it's limiting what you can and can't do"* (R125, male, ≤50 years, focus group); *"trying to use a pair of pliers to fix something and it was like…'really need that little finger to get that extra bit of power!'"* (H002, male, ≤50 years, interview).

4.4.4.1.4 Outcome domain: Range of movement

Descriptor: Amount of movement or stiffness in joints of the injured hand/wrist

Example codes: finger flexion-extension range or stiffness; wrist pronation-supination range or stiffness

Some saw recovery of range of movement as a key milestone, after which something like 'strength' could be considered: *"[I]t was mostly to do with bringing back range of motion, which was obviously limited. Rather than necessarily strengthening of anything. More trying to get that range of motion back so that I could use a pen and generally grip things as normal"* (H015, male, ≤50 years, interview).

4.4.4.1.5 Outcome domain: Confirmation of bone healing and alignment (for fractures)

Descriptor: Proper healing of the broken bones (including how well they are aligned)

Example codes: bone healing; x-ray appearance; malunion of fracture; non-union of fracture

Most participants with fractures cared about healing of the broken bone. In a basic sense, bone healing could be regarded as an indication of being on the path back to normalcy. Fracture healing and alignment was important to several patients, who wanted confirmation of healing despite other proxy clues of recovery and progress such as return to function: *"Worrying how the healing was going on. But that's something you don't know cause it's inside. You can't see it…it's just knowing that whether the bone's in line and doing what they should do"* (R105, female, >50 years, focus group).

For these patients, the potential long-term consequences of poor healing or alignment meant that any short-term return to function or diminished symptoms was not sufficiently reassuring:

> *"**H113:** I would worry about pain later in life. About being in the cold winter and having constant pain in the joint because it didn't heal straight the first time."*
>
> ***R109:** I was basically told if I didn't have the plate in, that would've been the situation I would be in."*
>
> (H113, male, >50 years & R109, female, >50 years, focus group)

4.4.4.1.6 Outcome domain: Confirmation of soft tissue healing

Descriptor: Healing of the soft tissues (e.g. tendons, ligaments, surgical wound)

Example codes: surgical wound; soft tissue healing

Some participants regarded healing of soft tissues as a distinct outcome domain, perhaps prompted by information from healthcare professionals: *“they said the tendon’s got less blood flow so I’ve got to keep it on for a long period of time to heal. And I’m also quite conscious and worried about taking the splint off. Because I don’t want to go back to square one…”* (H025, male, ≤50 years, interview).

For those treated surgically, sufficient healing of the wound was an important step to enable functions which required exposure to water: *“the wound is healed so it’s safe to touch water. So I start to use my left hand and left arm more to cover my daily activities”* (H051, female, ≤50 years, interview).

4.4.4.2 Outcome theme: Impact on self-care and independence

This theme involves outcome domains covering a range of ‘basic’ functional tasks that revolved around the concept of caring for oneself. It seemed that as the ‘symptomatic’ issues of the previous theme started to diminish, these were the first tasks which participants would tend to mention being able to do:

> *“I did start to figure ways around things. Towards the end I could put the bag on my arm myself. I was able to put certain jumpers on after a certain amount of time because it didn’t hurt my wrist when I had to pull the fabric over my sleeve. I was able to put jeans back on…got better with prepping food and that. I was alright chopping foods which I found really difficult at the start but I slowly progressed with it…I think it was just the mobility in my arm was coming back.”*
>
> (R028, female, ≤50 years, interview)

Being able to perform these fundamental tasks was seen as a marker of regaining independence and was clearly a priority for several participants: *"You take it for granted that you can get in the shower or the bath, but I don't think anybody wants to ask for help to be washed or anything"* (H135, female, >50 years, focus group) and:

> *"We've all mentioned…how important is it to be independent again. Cause independence and to be able to do things for yourself is really important…if somebody says what do you really want to do, the answer is, 'I really would like to be able to dry myself after a shower, wipe down the seat in the bathroom, chop up the food for my meal'."*
>
> (R109, female, >50 years, focus group)

Several of the outcome domains here arise directly from the examples of practical limitations highlighted in Section 4.4.3.7. However, where there are additional data extracts highlighting certain tasks as key outcome targets or milestones for participants they have been cited here under the relevant outcome domain.

4.4.4.2.1 Outcome domain: Eating and drinking

Descriptor: Being able to carry out the tasks need to feed oneself and quench thirst

Example codes: cutting food; holding a drink; general cutlery use to eat

Many participants mentioned an impact to this basic everyday activity, with some even setting it as a milestone on the path to recovery: *"To start off with I couldn't feed myself…I couldn't even lift a cup. Everything I drank, I drank with a straw. So to me progress was the day I was able to lift my cup up."* (R045, female, >50 years, interview).

4.4.4.2.2 Outcome domain: Self-hygiene / personal care

Descriptor: Being able to perform tasks related to cleansing oneself and maintaining personal hygiene

Example codes: washing oneself; brushing teeth; toileting

There was a general sense that washing oneself was a basic task and amongst the early functions that participants aimed to return to doing; *"initially, it's just those simple things and taking care of yourself. Having a shower or making some food"* (R125, male, ≤50 years, focus group).

4.4.4.2.3 Outcome domain: Cooking / preparing meals

Descriptor: Being able to complete the tasks involved in cooking or meal preparation

Example codes: cutting vegetables; eating ready meals; meal preparation

Some participants appeared strongly motivated to try to retain or regain this function, seeing it as a marker of independence:

> *"I was resting my elbow on a potato, trying to peel it…sat there in my uni kitchen and they'd just come in and be like, 'what are you doing?'. And I'm like, 'I'm peeling a potato on my own. I will do it'...Not fun. It just gets slippery and then your elbow smells like an unwashed potato, and it's really not fun."*
>
> (R107, female, ≤50 years, focus group)

4.4.4.2.4 Outcome domain: Grasping and moving light objects

Descriptor: Being able to hold onto and move/lift objects weighing less than ~1kg

Example codes: holding filled cup; opening containers; lifting pint of milk; opening or closing doors

This domain began as ‘grasping and moving objects’, and was refined on the basis of focus group discussion revealing that the difference from ‘carrying heavy objects’ was not clear enough to participants. The ~1kg threshold allowed for the sorts of examples patients cited for this function, involving relatively low levels of force necessary to perform the ‘basic’ task of manipulating light objects:

> *“I spilt a cup of tea yesterday evening, and I thought, ‘why did you try to pick that up?’ But it's down to the fact that you want to pick it up, you feel able to pick it up. But you can’t.”*
>
> H101, male, >50 years, focus group

4.4.4.2.5 Outcome domain: Walking / mobility

Descriptor: Being able to walk / mobilise after an injury, and have the confidence to do this

Example codes: confidence with mobilising; walking; general mobility impact

Even though the ability to mobilise was not directly related to their injury, several patients expressed an impact on this basic activity. This tended to relate to an almost subconscious loss of confidence and sense of vulnerability as walking, particularly outdoors, was seen as exposing themselves to risk of re-injury: *“Thinking about how much I've been out in the last couple of weeks, it's certainly been less, and I’m probably feeling less confident about going out…”* (H113, male, >50 years, focus group).

4.4.4.2.6 Outcome domain: Using hands to change body position

Descriptor: Being able to use your hands as needed to adjust body position, such as pushing up to stand from a seated position or turn in bed

Example codes: stand up from seated position; change in body position

This troubled some participants as a matter distinct from other outcome domains. For example, this participant was able to return to his hobby of gardening by using his uninjured upper limb but had persisting difficulty with the domain:

> *"I'm managing on a kneeler pad for example, and do it with one hand…The difficulty is when you get up from a special kneeling thing…at first, you use both hands to push up. I have to push with one hand and be very careful I don't topple over."*
>
> (R038, male, >50 years, interview)

4.4.4.2.7 Outcome domain: General household chores

Descriptor: Being able to do general tasks around the house

Example codes: household cleaning; washing dishes; household repair or maintenance

Housework involved a predictably broad range of activities depending on individual circumstances. A few participants seemed keen to return to performing these tasks, wanting to share the burden that was otherwise left to others in the household:

> *"I feel bad that I haven't been able to do housework or cook, or clean…all these things are slowly, slowly getting back. I think my girlfriend's happy that I could start doing the washing up! And help out around the house…"*
>
> (H041, male, ≤50 years, interview)

4.4.4.2.8 Outcome domain: Shopping for groceries

Descriptor: Being able to carry out the tasks involved in shopping for groceries (NOT the travel to and from shops)

Example codes: pushing shopping trolley; placing items into shopping basket/trolley

Grocery shopping was another domain which participants appeared to highlight as it was a task that, if not performed independently, necessitated reliance upon others:

> *"My sister helps take me shopping…she does my shopping for me, to put in the basket. I mean…when I'm out shopping I can't push a trolley. Cause once it starts getting heavy, going over bumps like on the ground, it rocks straight through me hand and it hurts. I can't do it."*
>
> (H042, male, ≤50 years, interview)

4.4.4.2.9 Outcome domain: Sleeping

Descriptor: Being able to sleep, including length and quality of sleep

Example codes: sleep disturbance due to pain; difficulty finding comfortable sleeping position; keeping arm elevated while sleeping

Difficulty sleeping, whether secondary to pain or lack of a comfortable resting position, was identified by several as an important impact on their overall experience:

> *"I'm having difficulty sleeping because I'd take whichever painkillers before I went to sleep, but three or four hours into the night I'd wake up in pain with it. Because they'd worn off…I'd take some more, but it would take an hour or so before they'd start to work, so I'd be awake for a while…So there's that side of it which you don't think about*

> *as well cause it's not just the injury itself. Cause if you're not sleeping very well, then that affects you anyway doesn't it?"*
>
> (R044, male, ≤50 years, interview)

4.4.4.2.10 Outcome domain: Dressing

Descriptor: Being able to put on and take off one's own clothing and footwear

Example codes: using buttons on clothing; fastening or unfastening bra; shoes and shoelaces

Even the simple act of dressing was an important task for some, who felt that relying on others for this activity impacted on their sense of dignity: *"I'm having to get my wife and my kids to button up my trousers like I'm about to go into an old people's home. You know, it was quite demeaning actually"* (H130, male, ≤50 years, focus group).

4.4.4.3 Outcome theme: Return to wider social roles, engagement and responsibilities

The interviews and focus groups involved participants at various phases of the patient journey and this appeared to be linked to the priorities of participants:

> *"You've got people here that are in three stages of…different stages of the injury and recovery. I think actually the timing of what stage you are in recovery is important because pain is very important to start off with cause it affects your mood and what you can do. But as you go through the recovery period that balance changes."*
>
> (R109, female, >50 years, focus group)

The previous theme focused on function related to self-care and independence. Beyond such issues there appeared to be a shift in the balance towards functions of a less ‘fundamental’ nature, related to roles and responsibilities involving others:

> *“The biggest impact is obviously not being able to carry out day-to-day duties, you know? Falling behind at work but also with my personal responsibilities. Mainly with my daughter and the knock-on effect that has with mine and my wife’s relationship. Even though she understands fully, it still has an effect.”*
>
> (H003, male, ≤50 years, interview)

The outcome domains in this theme are of a ‘higher’ functional level moving beyond meeting the basic needs of self-care.

4.4.4.3.1 Outcome domain: Return to work / job

Descriptor: Being able to return to work (includes unpaid carer role)

Example codes: work impact; time to return to work volunteer work; occupation

Many participants considered return to work an outcome domain distinct from the financial impact of not being able to work. There was a psychosocial element in not being able to maintain or engage in the ‘worker’ role: *“I was applying for the Navy so I've had to have a six month postponement of that…so, yeah, that was very disheartening”* (H106, male, ≤50 years, focus group); *“I feel guilty for being off work, I feel like should be back at work…it’s for others. Yeah, that's the mindset. It's just how I am”* (H049, female, >50 years, interview) and:

> *“I should have been signed off for four weeks, but I had a word with my employer and they just gave me light duties. And I think I needed that for my sanity…just need to*

> *get out of the house to focus on something other than look at my hand."*
>
> (H135, female, >50 years, focus group)

4.4.4.3.2 Outcome domain: Return to driving

Descriptor: Being able to drive

Example codes: driving child(ren); travel – work; travel – social events

For several participants, driving or transportation being impacted by their injury had consequences on various facets of life and therefore return to driving was an important outcome target: *"My priority was being able to drive. Because I found that very limiting…being able to go shopping and visiting other people…It's a sense of freedom as well isn't it?"* (R140, female, >50 years, focus group).

Overall though, this was a later priority:

> *"As you start to be able to do more things then you kind of expand to other parts of your life, like maybe thinking about when you can drive and that kind of thing. But initially it's just the simple things that you worry about."*
>
> (R125, male, ≤50 years, focus group).

4.4.4.3.3 Outcome domain: Fine hand use

Descriptor: Being able to do fine motor tasks with injured hand/wrist

Example codes: writing; drawing; applying make-up

Participants typically reported issues with fine motor skills related to writing or drawing, whether as part of work or routine tasks: *"I couldn't write anything out. We'd got a lot of birthdays, every time my daughter visited I had a pile of things for her to write out. 'Can you* *write this birthday card?', 'Can you write that envelope?'"* (R105, female, >50 years, focus group).

4.4.4.3.4 Outcome domain: Hobbies and recreation

Descriptor: Being able to do usual pastimes

Example codes: football; running; swimming; cycling; gardening; cross-stitching or sewing

While some participants found alternative hobbies to pursue even early on in recovery, for most re-engaging with pastimes was only a possibility at a later stage: *"it's only in the last two or three weeks…that I can actually garden and sew"* (R109, female, >50 years, focus group).

4.4.4.3.5 Outcome domain: Socialising / relationships with friends and family

Descriptor: Being able to socialise and maintain relationships with family and friends

Example codes: relationship with spouse/partner; relationship with wider family; socialising with non-family

Participants generally conveyed a sense of impact on relationships with others, with positive and negative experiences as summarised in Section 4.4.3.8. The impact on socialising and relationships with wider family or friends in particular was difficult to disentangle from that of lockdown for most, as the latter prevented meeting up even if otherwise possible. However, some could clearly discern an impact through the injury: *"I've got grandchildren…so normally I would drive up to see them quite often, which I couldn't do cause I couldn't drive and I wouldn't be able to hold them and things."* (R036, female, >50 years, interview).

4.4.4.3.6 Outcome domain: Carrying heavy objects

Descriptor: Being able to carry objects weighing more than ~5kg (i.e. more than a shopping bag with groceries)

Example codes: moving object – wheelbarrow; holding briefcase/suitcase/bag; holding filled saucepan

The domain brought together miscellaneous codes to do with lifting, carrying or moving heavy objects. A threshold was set for 'heavy' which conveyed the degree of force being considered by participants when discussing impact in this domain, and to distinguish it from carrying activities in the domains of 'grasping and moving light objects' and 'shopping for groceries'. Wider roles such as in work, recreation or caring for others tended to involve activity linked to this domain more commonly than the 'essential tasks' of the 'self-care and independence' theme. However, this domain could reasonably have been grouped with the latter theme.

4.4.4.3.7 Outcome domain: Sexual activity

Descriptor: Being able to accomplish one's usual sexual activity

Example code: sexual activity

Impact on sexual activity was brought up very rarely, and only by younger participants:

> *"Sex life for example…it is definitely something of consideration…it's something that obviously is affected…Yeah, I mean it's been affected by…the cast and the broken wrist."*
>
> (H003, male, ≤50 years, interview)

It is difficult to know whether this is because it was not a priority for most or whether there was general reluctance to discuss issues of this nature

4.4.4.3.8 Outcome domain: Looking after dependents within household

Descriptor: Being able to carry out the routine tasks involved in caring for dependents within the household

Example codes: bathing child(ren); feeding child(ren); dressing parent(s); walking dog

This domain summarised the various carer role codes that were identified. It was evident that the impact on participants' ability to look after others at home who relied upon them was important:

> *"I've…got a disabled daughter. She has got a feeding tube in her tummy and I can't physically fit the feeding tube or make the feeds or anything…My daughter's in a wheelchair and I can't push it and I can't even fasten the clips on it or fasten her in it. I can't even do up her nappy, which is really difficult, it's really, really hard."*
>
> (H117, female, ≤50 years, focus group)

4.4.4.3.9 Outcome domain: Keyboard and mouse / touchscreen use

Descriptor: Being able to make use of keyboard and mouse or touchscreens (such as is needed for computers, laptops, tablet computers or smartphones)

Example codes: keyboard use; use of computer mouse; touchscreen use

This domain arose out of a miscellany of codes which were related by virtue of being related in terms of how participants interact with IT equipment, usually for work and communication purposes. As such, it fit within this theme of wider roles in work and socialising:

> *"I think work assumed that I would probably be able to come out of the cast and day 1 go back to whatever it was that I was doing…So then to have to say…'I'm struggling. Even though I've got the cast off, I can't use a mouse right-handed so I'm still doing roughly what I was doing before whilst trying to build things up'. There's less understanding."*

(H015, male, ≤50 years, interview)

4.4.4.3.10 Outcome domain: Fatigue / tiredness

Descriptor: General feeling of low energy, tiredness or fatigue which restricts how much one can do (not due to issues with sleeping)

Example codes: tiredness or fatigue

A small number of patients noted a vague sense of fatigue that was distinct from any issue of difficulty with sleep, at times put down to the body compensating for functional impairment due to the injured hand or wrist:

> *"You don't realise how exhausting it is trying to manage those everyday things with just one hand…the amount of exhaustion and the effect it puts on other parts of your body compensating…"*

(R110, female, >50 years, focus group)

There was an impression that patients affected by such an issue persevered with basic activities of self-care but that it hindered wider roles of work, recreation, socialising, etc. The domain was therefore grouped within this theme.

4.4.4.4 Outcome theme: Emotional impact and psychological wellbeing

I found many outcome domains involved some impact on the mental health of participants, such as stress or frustration through functional impairment. I drew a distinction between such domains and others with a more direct link to emotional and psychological wellbeing. An example was described earlier, of the trainee police officer who was keen to return to driving after being injured in a road traffic accident: *"I do enjoy driving and the fear of not being able to drive…yeah, I felt 'was I gonna be worried driving?' and 'would it bring back bad memories?'"* (H041, male, ≤50 years, interview).

This participant was not simply concerned about the action of driving, but rather the emotional impact that the injury circumstances had on him. For H041, being able to drive again was more about overcoming his fear of re-engaging in the activity which had led to his injury.

4.4.4.4.1 Outcome domain: Emotional / mood impact to self

Descriptor: Effect on one's mood / emotions

Example codes: guilt over reliance on others; anxiety or stress; low mood; maintaining dignity; surprise or shock

There were many examples of the emotional impact of injury on participants. A common one was anxiety or even guilt due to reliance on others:

> *"I'm very worried about it because she's struggling to find a job. And she's given all this time up to me. And it means that she's not concentrating fully. She's trying to make sure she's there all of the time for me. That she doesn't leave me."*
>
> (R030, female, >50 years, interview)

4.4.4.4.2 Outcome domain: Emotional / mood impact to friends and family

Descriptor: Effect of one's injury on mood / emotions of their friends and family

Example codes: spouse or partner emotional impact; concern over angering spouse; children emotional impact

Some participants perceived an emotional impact on friends and family close to them:

> *"It's not something we've sat down as a family, or me and my wife, and discussed but I can probably gauge that. And how my injury and my treatment have had an impact on the family because they've had to step up or it's caused more friction."*
>
> (H130, male, ≤50 years, focus group)

4.4.4.4.3 Outcome domain: Overall patient satisfaction

Descriptor: Overall satisfaction of patient with the treatment, recovery process and outcome

Example codes: general impression of treatment; information from healthcare professionals; diagnostic process; accessibility to further information

Through this domain I drew together aspects of the patient journey that were perhaps quite disparate at first glance but appeared to naturally coalesce around the concept of participants' impressions of the overall treatment process and end result. While information from healthcare professionals may not itself appear to be an outcome, it can contribute towards patients' opinion on the overall experience and therefore their satisfaction with that experience.

4.4.4.4.4 Outcome domain: Appearance

Descriptor: How the hand/wrist looks after treatment is complete

Example codes: scarring, deformity, dry skin, swelling

Appearance was raised as a concern by only a few participants, generally from the younger age group: *"I think my only worry is when I do take the cast off, my finger's such a weird shape…I've been warned that it will probably never be straight"* (H119, female, ≤50 years, focus group).

At times the underlying concern was regarding the impact on the participant's attractiveness to others and how they might perceive the injured hand or wrist:

> *"I was thinking, 'oh, if it doesn't heal then that's…I'm going to have that stuck like that'. And it is a bit upsetting. It's not a very nice feature to have…it obviously wasn't attractive."*
>
> (H025, male, ≤50 years, interview).

4.4.4.5 *Outcome theme: Impact to personal and healthcare resources*

The various outcome domains in this theme were a result of impacts in other domains (e.g. being unable to work because of functional impairment, leading to a loss of income). However, they stood out in their own right as distinct ways in which participants could assess the impact of their injury.

Focus group discussion helped to clarify that participants did consider these outcome domains as belonging within an overall theme of economic impact:

> *"Whether you were able to work, what arrangement there was for getting paid or not paid if you weren't able to work, but also the economic impact of having to use taxis*

> *more if you can't drive, having to go to the hairdresser more because you can't wash your hair, having to buy ready-made food because you can't prepare it."*
>
> (H104, female, >50 years, focus group)

4.4.4.5.1 Outcome domain: Loss of income

Descriptor: Lost income from not being able to work

Example codes: reduced income; sick pay; working longer hours

In some participants' circumstances this clearly was of great significance: *"I only get statutory sick pay in my job….it'd impact me hugely in terms of my monthly salary. And then…obviously I've still got my mortgage to pay, bills to pay"* (H002, male, ≤50 years, interview).

4.4.4.5.2 Outcome domain: Added expenses due to injury

Descriptor: The increased costs to a patient while recovering from a hand/wrist injury

Example codes: reduced income; sick pay; working longer hours

Added costs secondary to injury could tally up quite quickly: *"I can't drive. It's 15 quid from my house in a taxi, so every time I've come [to fracture clinic] it's cost me £30"* (R029, female, ≤50 years, interview) and:

> *"I bought this protective sleeve and that was £10.80, and then I bought a sleeve for the actual cast just to keep it clean for when I go back to work…buying shoes that slip-on instead of tying, because I have trouble tying my laces…little things like that add to your shopping list."*
>
> (H135, female, >50 years, focus group)

While most considered this a minor inconvenience, for some it had an important effect on their cost of living:

> *"It's a lot of money. And that's on your mind. You're thinking about it all the time, aren't you? Because I have to say to myself, 'if I'm not financially able to sort myself out, what am I gonna do?' There's nobody else I can go to. I have to be financially…I have to budget."*

(R030, female, >50 years, interview)

4.4.4.5.3 Outcome domain: Healthcare resources used

Descriptor: Cost of treatment, follow-up and any associated healthcare provision costs (until end of recovery or discharge from follow-up)

Example codes: general clinician time use; operative treatment cost;

Although hinted at in some interviews, it was also through focus groups that the final resource-related outcome domain of use of healthcare resources was clarified as being something that some participants considered of relevance to their injury:

> *"[W]hen I actually got it operated on, I felt very humbled and I felt very stupid because…you know, the resources that that consumed on that day from me entering to exiting, it's just staggering…there must have been 12 people in the [operating] theatre and I'm lying there thinking 'what an idiot' and I felt really ashamed and embarrassed actually. That I was using those resources."*

(H130, male, ≤50 years, focus group)

4.4.4.6 Outcome theme: Treatment factors and complications

This was a miscellaneous theme which summarised outcome domains related to the treatment itself. These domains stood apart from those of the other themes in that I used them to capture outcomes that were either aspects of the interventions that participants experienced or ensuing adverse events in the course of treatment.

4.4.4.6.1 Outcome domain: Treatment convenience / comfort

Descriptor: The inconvenience or discomfort caused by a treatment (such as having a cast or splint on, dressings changes and stitch removal after surgery)

Example codes: avoiding water damaging splint or cast; hygiene or maintenance of splint or cast; having stitches removed; having to attend hospital; splint or cast hindering movement

While participants understood the importance of the treatment they received, several described inconveniences directly due to that treatment: *"I've got it strapped up…And it's quite annoying, having to redo it every day and every time it gets wet the bandage comes off"* (H119, female, ≤50 years, focus group).

4.4.4.6.2 Outcome domain: Treatment side effects or complications

Descriptor: Any side effects or complications that arise as a result of a specific treatment

Example codes: malunion of fracture; post-op infection; post-traumatic arthritis

Unfortunately, a few had complications or adverse events which clearly stood apart from more routine issues:

> *"I now don't have my splint on and I'm supposed to be doing physio but, unfortunately, there's tendon damage now. And I believe that's from where the wires*

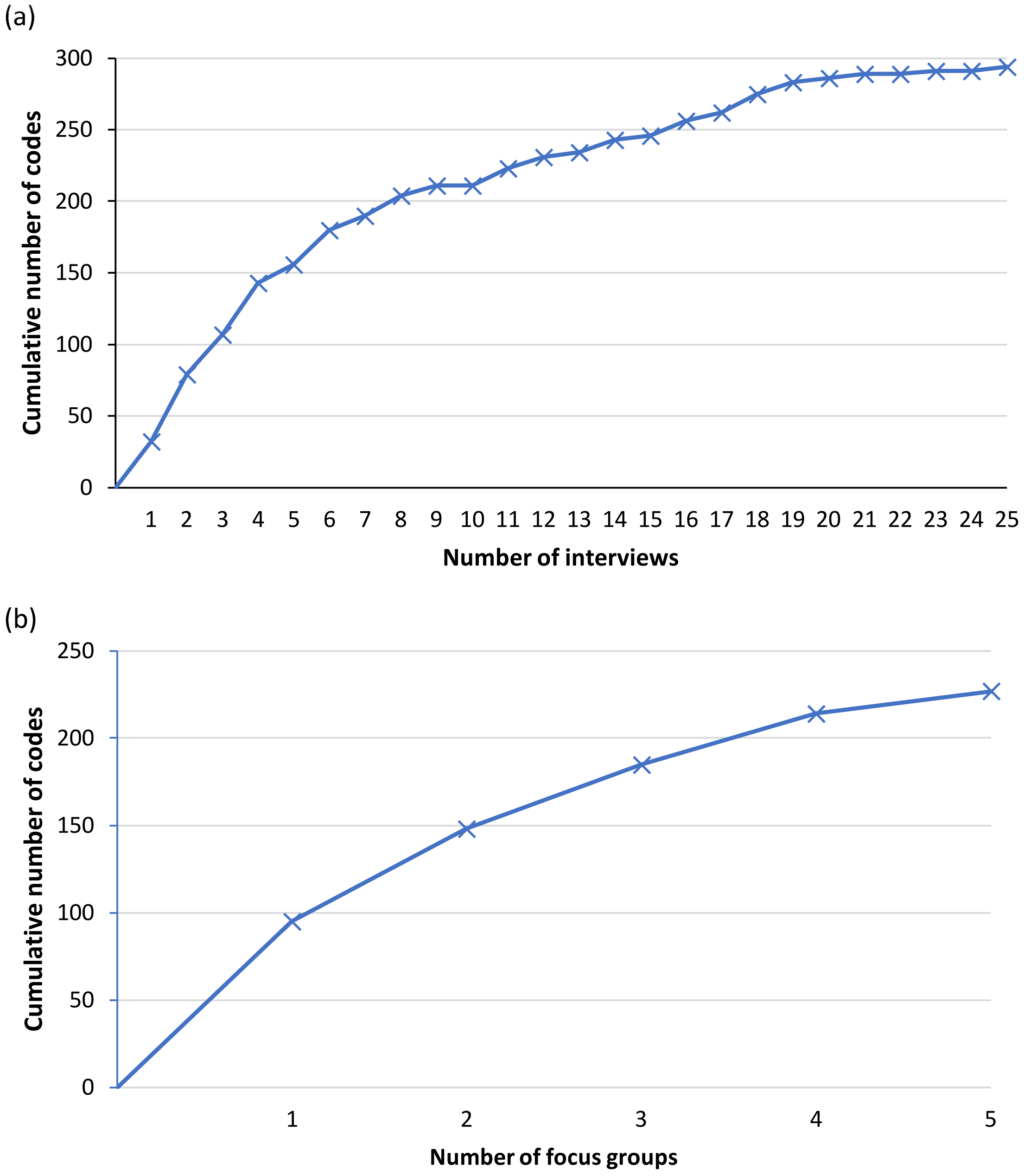
> *were put in. So this is my best thumbs up, it's not very good. So my thumb's not working properly now."*
>
> (R110, female, >50 years, focus group)

4.4.5 Coding and data saturation

In total >300 codes were generated over the course of analysis of the full data set (Appendix 6). The coding saturation curves for interviews and focus groups are shown in Figure 4.5 (a) and (b), respectively, and were a means by which I could prospectively try to assess whether our target of data saturation was achieved.

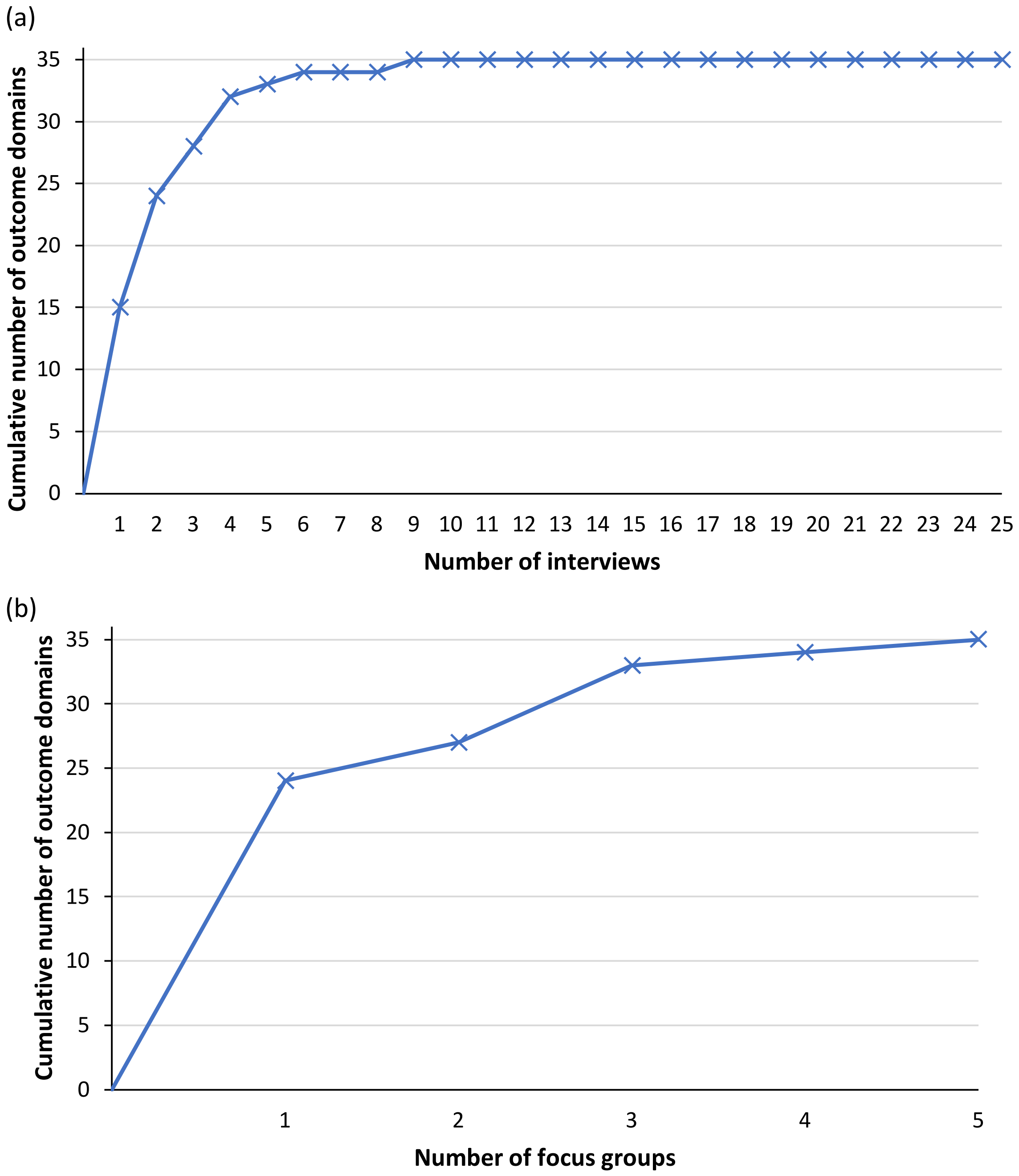
These curves plateau, such that by interview 19 I generated approximately 96% (283/294) of total codes developed by analysis of interview transcripts. Similarly, by focus group 4 there were already 94% (214/227) of the codes generated through analysis of all focus groups.

Figure 4.5 Saturation curves for coding of (a) interviews, and (b) focus groups



Many of the codes were quite granular, and I believed that perhaps the saturation of outcome domains achieved in interviews and focus groups was more relevant (Figure 4.6). I was able to perform this retrospective check after interpretive analysis and generation of the outcome domains was complete.

Figure 4.6 Saturation curves for outcome domains of (a) interviews, and (b) focus



Outcome domain saturation was reached by interview 9. For the focus groups, domain saturation was essentially reached by focus group 3, though technically a further outcome domain was generated through each of the remaining two focus groups.

4.4.6 Comparison of outcome domains for distal radius fracture and other hand fracture and joint injury patient interviews

I analysed for overlap in outcome domains between the two injury groups according to the chronology of interviews. Table 4-8 shows the overlap in outcome domains after two, five and all interviews for each of the injury groups.

Table 4-8 Summary of overlap in interview-based outcome domains between both injury groups at various checkpoints

	Number of interviews completed for each injury group		
	2	**5**	**All**
Overlap across both injury groups	22	31	33
Unique to DRF group	7	1	0
Unique to non-DRF group	2	2	2

As the interviews progressed there were very few outcome domains unique to one injury group or the other, and by the end there were just two; both ‘healthcare resources used’ and ‘sexual activity’ arose in the non-DRF interviews only. Each was only mentioned in a single interview though, and were the only outcome domains linked to a single interview). This reassured me that there was a high degree of overlap in outcome domains considered relevant by the two injury groups, as appeared likely to be the case based on the close overlap in patient journey themes found during the descriptive analysis phase.

4.5 Discussion

4.5.1 Key findings

Through this qualitative study, I explored the patient journey and thereby identified a comprehensive set of outcome domains of relevance to patients with hand fractures and joint injuries. The 35 outcome domains were summarised in six themes, ranging from fundamental concepts such as the sequelae to patient biology and physiology, through to downstream impacts on various basic functional tasks to maintain independence and associated functional impact on their wider roles in terms of family, work, etc. There were also themes of emotional and economic impacts precipitated by the injuries. Although attempting to encapsulate different concepts across the 35 outcome domains, in many cases there is interconnectedness between them.

It was not clear from this study alone which outcome domains patients with these injuries would prioritise. However, the general impression gained from the data suggested the potential for a hierarchy, at least in the outcome themes, as reflected in a shift in the balance of priorities for several participants as recovery progressed.

4.5.2 Comparing outcome domains of relevance for distal radius fracture versus other hand fractures and joint injuries

As a secondary aim, I wanted to clarify variation in outcome domains highlighted by patients who had distal radius fractures compared to those who had other forms of hand fractures and joint injuries. As shown in the results, there were occasionally contrasting features in the patient journey between these injury groups – such as no patients with distal radius fractures delaying seeking medical attention.

As presented in the results and evident in the full data set, virtually all outcome domains were considered of relevance to both injury groups. I suspect this is because all these injuries still have the consequence of impairing use of the upper limb, causing pain, etc. The treatments for these injuries also overlap in many ways – a cast for a distal radius fracture tries to leave the digits free to move, but functional ability is typically still impaired due to pain, inability to exert necessary force, inability to position the hand as one normally would, etc. Purposive sampling is likely to have contributed to this finding of outcome domain overlap as well – while the typical demographics for each injury group (older women with distal radius fractures, younger men with other hand fractures and joint injuries) made up the majority of the sample, I ensured that a broad range of demographics was covered overall. This actually better reflects the true demographics of those affected by these injuries, as all these injuries do occur in any age group or sex.

A few domains did seem to be cited by participants of one injury group noticeably more or less than the other, as noted in the results. ‘Appearance’ was infrequently a concern, but by fewer participants with distal radius fractures. ‘Sexual activity’ was mentioned very rarely, by only two participants from the ‘other hand fractures and joint injuries’ group. Given that most did not state these domains as a priority or concern, it may seem unlikely that these would be considered priority outcomes to the wider group. As stated previously though, an alternative explanation is that despite being a priority they are somewhat sensitive topics that participants simply did not feel comfortable discussing. Outcome domain priorities amongst patients is explored more definitively in the Delphi and consensus meeting chapters, as the goal of this step was to achieve a comprehensive set of outcome domains based on the patient perspective rather than establishing prioritisation.

4.5.3 Existing literature

There have been a few qualitative studies on patient experience of injuries to the hand or wrist, with the latter being more frequently covered. Despite a variety of approaches, aims and scopes, this literature indicates some common themes of potential relevance to patients with injuries within the scope of the SO-HANDI COS: impact to function, pain, impact on roles and relationships with others, and treatment consequences.

4.5.3.1 Impact to function

Impact to function was covered in all studies. Cheng (2008) focused on how this affected return to work, while in most others it was considered more broadly. They considered function in terms of activities of daily life, recreation, work, etc. ranging from fine motor tasks to those where crude grip strength was more of a concern (Gustafsson et al. 2000; Bialocerkowski 2002; Ammann et al. 2012; Kingston et al. 2015; Troianello et al. 2017; Watson et al. 2018; Andreasson et al. 2019). This broadly corresponded to the SO-HANDI themes of ‘impact on self-care and independence’ and ‘return to wider social roles, engagement and responsibilities’. In commenting on impact to function, several studies highlighted factors such as strength and range of movement.

Impairment to function, particularly with added stressors such as the financial implications when ability to work is affected, in some cases lead to emotional distress and anxiety

(Gustafsson et al. 2000; Bialocerkowski 2002; Chan et al. 2004; Coenen et al. 2013; Watson et al. 2018; Dias et al. 2020). These elements related to the SO-HANDI theme of ‘emotional impact and psychological wellbeing’ and the ‘loss of income’ outcome domain. However, I also had the domain ‘added expenses due to injury’ which was only raised in one study (Kingston et al. 2015) and ‘healthcare resources used’ which was not described by past qualitative studies.

4.5.3.2 Pain

Pain was addressed in the majority of studies and was the main focus of two of them (Chan and Spencer 2005; Sale et al. 2017). It has been highlighted as a stress factor (Gustafsson et al. 2000) which in turn can have numerous impacts not just on physical function (Bialocerkowski 2002; Kingston et al. 2014) but through disturbed sleep, anxiety and effects on relationships with others (Coenen et al. 2013; Andreasson et al. 2019).

Pain can also be seen by patients as something much more concerning – an indication of increasing damage to the injured region (O’Brien and Presnell 2010).

4.5.3.3 Impact on roles and relationships with others

Impact to function and pain resulted in some having to rely on others for support leading to a, usually temporary, shift in or loss of typical life roles such as ‘worker’ or ‘caregiver’ for many patients (Bialocerkowski 2002; Schier and Chan 2007; Bamford and Walker 2010; Kingston et al. 2014; Watson et al. 2018). Relationships with those outside the household, such as friends or neighbours, were potentially disrupted by inability to participate in socialising activities (Andreasson et al. 2019). All these impacts proved frustrating for many patients, with several stating a desire to return to independence (Gustafsson et al. 2000; Lai 2004; Troianello et al. 2017; Dias et al. 2020). However, some recognised that the change in dynamic in certain relationships highlighted a supportive side to people in their lives (Bamford and Walker 2010). This support was sometimes emotional in nature (Gustafsson et al. 2002). This also corresponded to the SO-HANDI themes of ‘return to wider social roles, engagement and responsibilities’ and ‘emotional impact and psychological wellbeing’.

A desire to return to the normal state of relationships was seen by some as a motivating factor in their own adaptation during recovery either to avoid having to rely on others (Lai 2004), or to be able to resume valued activities with others (Chan and Spencer 2004).

4.5.3.4 Treatment consequences

The aim of treatment is to enable appropriate healing from injury, and patients clearly recognise the protective value of a cast at an early stage after their injury (Watson et al. 2018). However, for many patients the treatment itself had various negative attributes. Removal of casts or splints was a source of relief for patients as they can be uncomfortable (O’Brien and Presnell 2010) and can hinder function (O’Brien and Presnell 2010; Watson et al. 2018; Dias et al. 2020). Surgical treatment has its own accompanying risks related to scarring, anaesthesia, etc that are not incurred by conservative treatment (Dias et al. 2020). This essentially matched the theme ‘treatment factors and complications’.

‘Treatment consequences’ also encompasses the concept of some treatment options facilitating an earlier return to function over others. For example, surgery is typically seen as being able to enable a quicker return to work (Dias et al. 2020) or simple activities of daily living (Nasser et al. 2018; Dias et al. 2020) over conservative treatment in a cast or splint.

4.5.3.5 Overall assessment of existing literature

The existing literature is quite limited and heterogeneous in its scope and reach. Some have involved detailed exploration of just three or four participants (Schier and Chan 2007; Ammann et al. 2012) while others have opted for data sets from a larger participant pool, such as the 49 participants interviewed as a nested study within a scaphoid fracture trial (Dias et al. 2020). While most of the studies involved single interviews, a few were longitudinal in nature with two or more interviews per participant. One was entirely based on focus groups (Coenen et al. 2013). The hand and wrist injuries being focused upon in each study have varied from specific injuries to a broad mix of injuries including those lying outside the scope of the SO-HANDI COS.

The variation makes synthesis of this body of work challenging and it would be difficult to state with conviction that I could establish the outcome domains of relevance to patients with injuries within the scope of the SO-HANDI COS based solely on the existing literature. However, the general themes summarised in the literature did overlap with those developed through thematic analysis of the SO-HANDI data set. These past studies did not generally aim to generate outcome domains of relevance to the patient participants, with the exception of the study by Coenen et al (2013). Even that mapped focus group data to an existing WHO ICF framework rather than generating their own outcome domains through an inductive approach. Therefore, although the various themes summarised above reflected a general overlap with the findings of the SO-HANDI qualitative study, on the whole there were not explicit outcome domains in the existing literature which could directly compare with the SO-HANDI ones. However, I attempted to determine which of the SO-HANDI outcome domains could be considered to have been raised in past publications based on the various authors’ comments and data extracts provided (albeit of varying depth and quantity). This was of course merely a crude comparison. As shown in Table 4-9, although several of the SO-HANDI themes corresponded to those reported by many of the studies, the specific outcome domains were infrequently found. Indeed a few, such as ‘confirmation of bone healing and alignment’, ‘walking / mobility’, ‘fatigue / tiredness’, ‘overall patient satisfaction’, ‘added expenses due to injury’ and ‘healthcare resources used’ were present in none or only a single study. It was unclear whether the differences between the SO-HANDI study and previous studies were due to varied aims, methodology or a limitation in terms of the published data from those studies.

Table 4-9 Summary of linkage of SO-HANDI outcome domains to published data and analysis in past studies

Outcome theme	Outcome domain	Amman et al (2012)	Andreasson et al (2019)	Bamford & Walker (2010)	Bialocerkowski (2002)	Chan & Spencer (2004)	Chan & Spencer (2005)	Cheng (2008)	Coenen et al (2013)	Dias et al (2020)	Gustafsson (2000, 2002)	Kingston (2014, 2015)	Lai (2004)	Nasser et al (2018)	O'Brien & Presnell (2010)	Sale et al (2017)	Schier & Chan (2007)	Troianello et al (2017)	Watson et al (2018)
DISRUPTION TO BIOLOGY AND PHYSIOLOGY	Change in sensation	🟥	🟩	🟥	🟩	🟥	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Pain / discomfort	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟥	🟩
	Strength or muscle power	🟥	🟥	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟥
	Range of movement	🟥	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥
	Confirmation of bone healing and alignment	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Confirmation of soft tissue healing	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩
IMPACT ON SELF-CARE AND INDEPENDENCE	Eating and drinking	🟩	🟩	🟥	🟩	🟥	🟥	🟥	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥
	Self-hygiene / personal care	🟥	🟩	🟥	🟩	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟩	🟩	🟩
	Cooking / preparing meals	🟩	🟩	🟥	🟩	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟩	🟥	🟩	🟩	🟥
	Grasping and moving light objects	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟥
	Walking / mobility	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Using hands to change body position	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	General household chores	🟩	🟩	🟩	🟩	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟩	🟥	🟩
	Shopping for groceries	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Sleeping	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥
	Dressing	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟩	🟩	🟥

Green boxes are used to indicate studies where we judged there was linkage to an outcome domain, and red where no linkage was apparent

Table 4-9 Summary of linkage of SO-HANDI outcome domains to published data and analysis in past studies (cont'd)

Outcome theme	**Outcome domain**	**Amman et al (2012)**	**Andreasson et al (2019)**	**Bamford & Walker (2010)**	**Bialocerkowski (2002)**	**Chan & Spencer (2004)**	**Chan & Spencer (2005)**	**Cheng (2008)**	**Coenen et al (2013)**	**Dias et al (2020)**	**Gustafsson (2000, 2002)**	**Kingston (2014, 2015)**	**Lai (2004)**	**Nasser et al (2018)**	**O'Brien & Presnell (2010)**	**Sale et al (2017)**	**Schier & Chan (2007)**	**Troianello et al (2017)**	**Watson et al (2018)**
RETURN TO WIDER SOCIAL ROLES, ENGAGEMENT AND RESPONSIBILITIES	Return to work / job	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟩	🟥	🟩	🟥	🟩
	Return to driving	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟩
	Fine hand use	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟩	🟥
	Hobbies and recreation	🟥	🟩	🟥	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟥
	Socialising / relationships with friends and family	🟩	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟩
	Carrying heavy objects	🟥	🟩	🟥	🟩	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟥	🟥
	Sexual activity	🟥	🟥	🟥	🟩	🟥	🟥	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Looking after dependents within household	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟩
	Keyboard and mouse / touchscreen use	🟩	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟥	🟥
	Fatigue / tiredness	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
EMOTIONAL IMPACT AND PSYCHOLOGICAL WELLBEING	Emotional / mood impact to self	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟩
	Emotional / mood impact to friends and family	🟥	🟩	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥
	Overall patient satisfaction	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥
	Appearance	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩	🟩	🟥	🟥	🟩	🟩	🟥	🟥	🟥	🟥

Green boxes are used to indicate studies where we judged there was linkage to an outcome domain, and red where no linkage was apparent

Table 4-9 Summary of linkage of SO-HANDI outcome domains to published data and analysis in past studies (cont'd)

Outcome theme	**Outcome domain**	**Amman et al (2012)**	**Andreasson et al (2019)**	**Bamford & Walker (2010)**	**Bialocerkowski (2002)**	**Chan & Spencer (2004)**	**Chan & Spencer (2005)**	**Cheng (2008)**	**Coenen et al (2013)**	**Dias et al (2020)**	**Gustafsson (2000, 2002)**	**Kingston (2014, 2015)**	**Lai (2004)**	**Nasser et al (2018)**	**O'Brien & Presnell (2010)**	**Sale et al (2017)**	**Schier & Chan (2007)**	**Troianello et al (2017)**	**Watson et al (2018)**
IMPACT TO PERSONAL AND HEALTHCARE RESOURCES	Loss of income	🟥	🟩	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟩
	Added expenses due to injury	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥
	Healthcare resources used	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
TREATMENT FACTORS AND COMPLICATIONS	Treatment convenience / comfort	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟥	🟥	🟩	🟥	🟥	🟥	🟩
	Treatment side effects or complications	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥

Green boxes are used to indicate studies where we judged there was linkage to an outcome domain, and red where no linkage was apparent

4.5.4 Strengths and limitations

Purposive sampling achieved participant numbers broadly in line with the planned sampling frames. Other factors that were not part of the sampling frames were also considered after an initial period of recruitment. Occupation (no occupation/retired/non-manual versus light/heavy manual occupations) was not a deciding factor in initial sampling (i.e. potential participants were initially approached based on the other parameters) but was considered when deciding whether a suitable mix had been achieved. There were further subtleties when sampling, such as an awareness that even within the ‘other hand fractures and joint injuries’ the goal was to include participants with a range of injuries rather than have over-representation of a few specific ones. Furthermore, I attempted to recruit participants from a range of time points in their care and recovery pathway. Not all permutations of the various demographic characteristics had high representation – for instance only two heavy manual workers were recruited, neither of whom had distal radius fractures. However, trying to consider these additional parameters during recruitment improved the exploration of diverse perspectives which enriched the iterative process of data collection and analysis.

The coding and outcome domain data saturation curves provide reassurance that data saturation was being reached, particularly for the latter in interviews. It could be argued that new codes were arising from data at the end of the run of interviews, but the codes in general were quite granular. Given the study’s goal it seems reasonable to prioritise the saturation of outcome domains over that of the codes. Although in retrospect it appears saturation of outcome domains occurred quite early across the run of interviews there was a benefit of recruitment having continued to achieve the purposive sampling frame as this helped to ensure that a broad range of perspectives was explored.

The rise of COVID-19 did impact on the study, with my preference for face-to-face data collection having to be sacrificed due to safety concerns. Most participants were able to manage virtual ones with both video and audio available. I believe that for the purposes of this study in developing a longlist of outcome domains, this method was sufficient. Nevertheless, there were a few patients who did not have access to a device to enable virtual interview/focus group participation. Although interviews could be facilitated for such patients via telephone, the focus group sessions required multiple participants connecting to the meeting which could not be done via landline. Therefore, input from such patients was restricted to the interviews.

There were issues such as last-minute cancellations and ‘no shows’ for the focus groups. Focus groups where fewer than the anticipated number of participants attended still proceeded out of respect for those who had set time aside to take part, as per the protocol. The reduced number of participants per group led to a decision to hold more than the three sessions originally planned.

Finally, while this step involved an inductive, thematic analysis, it is possible that some existing knowledge of frameworks of health outcome domains (as discussed in Section 1.4.2 and identified in the systematic review) could have influenced the development of the outcome domains and themes here. While attempting to keep these fully grounded in the data, a subconscious influence is impossible to entirely rule out. My medical background and clinical training likely shaped the choice of vocabulary in some of the coding, but I specifically worked to ensure lay-appropriate language was used for outcome domains and their descriptors and refined these during the focus groups.

4.6 Conclusion

This study developed a longlist of 35 outcome domains in six themes, based on the perspective of patients with hand fractures and joint injuries. A broad range of outcome domains are considered of relevance by such patients and there was great overlap in the experiences of those with distal radius fractures and other hand fractures and joint injuries.

The next step was to conduct a Delphi consensus study to determine what consensus could be found for the very important outcome domains amongst the key stakeholders – patients, hand surgeons and hand therapists. The foundation of the Delphi study consisted of a synthesis of the outputs of both the prior systematic review and this qualitative study, as described in the next chapter.

Chapter 5 Delphi consensus study

5.1 Introduction

The work described in Chapters Chapter 3 and Chapter 4 described the outcome domains used in the recently published clinical research on hand fractures and joint injuries and deemed of relevance by patients, respectively. In the study described in this chapter, these outputs were synthesised to generate a longlist of outcome domains that served as a starting point for consensus development. Consensus development took place over two phases, the first being an international Delphi study. The second phase was an international consensus meeting and is detailed in Chapter 6.

The COMET Initiative considers the Delphi technique useful in learning the opinion of a wide group of participants and report it as a common choice for COS development both past and present (Williamson et al. 2017). In essence, a Delphi study or survey, is a method of seeking the consensus opinion of numerous participants through sequential rounds of questionnaires answered anonymously – the participants need to have appropriate experience or expertise in the topic under consideration (Keeley et al. 2016). Delphi surveys are designed to, “obtain the most reliable opinion consensus” (Dalkey and Helmer 1963). In terms of COS development, the questionnaires usually involve some variation of rating of outcome domains such that a longlist derived from a combination of the systematic review and qualitative work can be refined to a shortlist or ranked in some manner.

By avoiding direct interaction between participants, the risk that the views of outspoken individuals will dominate are reduced compared to a face-to-face setting (whether in person or via teleconference) (Jones and Hunter 1995). Participants do receive some information on how others have rated particular outcome domains, and it is this information which can influence their opinion from one round of the Delphi process to the next. The Delphi study output is often discussed in a subsequent consensus meeting involving key stakeholders, to finalise the COS (Keeley et al. 2016).

5.2 Aims

The purpose of this study was to identify consensus of opinion on which outcomes are important for inclusion in a COS on hand fractures and joint injuries in adults, based on the views of key stakeholder groups: patients, hand therapists and hand surgeons.

Primary objective: To conduct a three-round Delphi study to determine consensus amongst key stakeholders in prioritisation of outcome domains from a comprehensive list of items relevant to the treatment of hand fractures and joint injuries.

Secondary objective: To consider aspects of Delphi methodology in general (consensus criteria based on single overall panel vs individual stakeholder group panels; setting removal of items per round vs only after final round; attrition) and more specifically for this COS (assessing difference in prioritisation between patient participants with distal radius fracture as opposed to other hand fractures and joint injuries).

5.3 Methods

5.3.1 Background

The Delphi technique originates from the RAND Corporation, a non-profit global policy think tank that developed this method to systematically determine consensus of opinion from a group of experts through a series of “questionnaires interspersed with controlled opinion feedback” (Dalkey and Helmer 1963). In situations where direct empirical evidence is lacking, this might be considered a reasonable substitute based on a “two heads are better than one” rationale (Dalkey 1969). While the original method involved an initial qualitative approach through questionnaire or interview to generate potential responses for which subsequent consensus was determined, Delphi surveys in COS development typically use a longlist of outcomes developed through one or more other means which are then presented to Delphi participants (Keeley et al. 2016).

The key principles of a Delphi study are (Hasson et al. 2000):

- Involvement of an ‘expert’ panel
- A series of structured questionnaires
- Controlled feedback to participants between each questionnaire round summarising the responses given
- Participant responses anonymised (to each other)

Since participants do not directly interact with one another, the Delphi technique avoids situations where a few outspoken individuals ‘steer’ the opinion of the overall group. Delphi studies have also benefitted from advancements in communications technology, which nowadays makes it quite practical to have a panel consisting of geographically distant participants in large numbers (Sinha et al. 2011).

5.3.2 Participants

5.3.2.1 Stakeholder groups

Delphi studies aim to learn the consensus opinion of a group of experts. It is therefore critically important to identify suitable 'experts' for the topic being studied. There is some controversy over what constitutes an 'expert' in this scenario (Hasson et al. 2000) but for COS development a range of stakeholders might be selected. I have included hand therapists, hand surgeons and patients as these are the three key stakeholder groups. Hand therapists and surgeons are directly involved in administering treatment to patients, with hand therapists typically involved in patient follow-up well beyond that by the surgeons. As modern COSs aim to represent the opinions of patients in terms of outcomes to be measured, I believed it vital to include patients with hand fractures and joint injuries in this consensus study.

Demographic details were recorded as part of the registration process for participants, and for all included month and year of birth, sex, country of residence and existing level of research experience (with dropdown options provided). Demographic details were collected as follows:

- Patients: injury sustained; whether dominant hand was injured; occupation
- Therapists: duration of experience managing patients with these injuries; whether they are an accredited/certified hand therapist (or whether no such option exists in their country); predominant training background (whether physiotherapy or occupational therapy)
- Surgeons: duration of experience managing patients with these injuries; predominant specialist training background (whether orthopaedic or plastic surgery)

The eligibility criteria outlined in Section 5.3.2.3 summarise the definitions for ‘expert’ in each of the three stakeholder groups.

5.3.2.2 Sample size

The number of participants for Delphi studies is not based on statistical power calculation and is instead a choice based on practical matters, which has led to wide variation in Delphi panel sizes (Williamson et al. 2017). I set a minimum recruitment number of 30 for each of the three stakeholder groups – this would enable input of at least 20 participants from each group even accounting for an attrition of up to one third. There was no maximum limit set, in accordance with the COMET Initiative guidance that the more participants there are to represent each of the stakeholder groups, the better for the purposes of generalisability to future patients and producing a COS which is more likely to be of convincing value to others (Williamson et al. 2017). A Health Technology Assessment (HTA) monograph on consensus development methods agreed with the overall principle that when combining judgments of individuals, more is better – however they did not specify numbers of participants for a Delphi but rather cited numbers used in studies on more directly interactive groups (Murphy et al. 1998).

5.3.2.3 Participant eligibility and recruitment

Eligible patient participants (using the same criteria as for the interviews and focus groups) were invited to participate via a range of pathways, including social media call to recruitment, posters displayed in fracture clinics at Queen’s Medical Centre, relevant patient advisory groups (such as that of the Centre for Evidence Based Hand Surgery, University of

Nottingham), patients who participated in the earlier interviews and focus groups and gave consent to be contacted about further research, and from the adult fracture clinic at Queen’s Medical Centre itself.

Health professional participants had to be a member of one of the following groups to be eligible:

- Hand surgeon (independent practitioner level such as Consultant or equivalent) – this could include plastic surgeons or orthopaedic surgeons if a majority of their workload involves hand and wrist surgery
- Hand therapist – once qualified, working independently (i.e. beyond training years) and subspecialised in hand therapy

These clinicians were invited to participate from a number of sources:

- Through contacting national and international orthopaedic, hand therapist and hand surgery societies to ask for information about the study to be disseminated to members. Anyone willing to participate would thereby be able to contact us for further details
- The Centre for Evidence Based Hand Surgery (CEBHS) Hands Surgery Evidence Updates mailing list – this was freely accessible and run by the CEBHS information specialist. Subscribers to the list included mostly clinical practitioners with an interest in clinical research updates relevant to hand surgery
- Invitations were sent to research groups with recent publications in the field of hand fractures and joint injuries by emailing corresponding authors of papers covered in the SO-HANDI systematic review

- An announcement about the study was included within one of the webinars run by the Pulvertaft Hand Centre, which has an international audience of hand therapists and surgeons

While international participation was welcomed, for pragmatic reasons participants were required to understand written/typed English. While translation of instructions for the Delphi questionnaires into most common languages might be straightforward, I was concerned that subtleties in translation could alter the meaning of outcome domains in a way which could detract from the eventual result obtained.

5.3.3 Questionnaire development

There were several important stages in developing the Delphi questionnaire, from initial generation of the outcome items, stakeholder refinement of the wording and descriptors for outcome items, deciding on a scoring system and setting consensus criteria.

5.3.3.1 Initial generation of Delphi outcome items

I synthesised the outputs of the systematic review and qualitative work done with patients in order to develop a unified longlist of outcome items for the Delphi. The outcome domains identified through the systematic review were based on the WHO ICF (Chapter 3), but the terminology of these domains was felt to be unclear at times, particularly for laypeople. In contrast, the phrasing of the outcome domains generated through interview and focus groups with patients (Chapter 4) was based on more accessible language as a result of discussion in focus group sessions. I therefore began with linking the review's WHO ICF domains to the existing qualitative study's domains, where appropriate. In any cases where the review and qualitative study outcome domains did not appear to link, Delphi outcome items were generated based on the single available source. Descriptors were produced which explained the meaning of each item, and where necessary illustrative examples were added to improve clarity.

5.3.3.2 Drafting of questionnaire instructions

Recommendations for design of questionnaires include clear instructions that enable participants to understand the purpose of the questionnaire as well as how they are meant to complete it (McColl et al. 2001). Context of a questionnaire is important because it can influence responses. I therefore drafted a set of instructions for the first round of the Delphi that clearly established the purpose of the questionnaire towards COS development. This highlighted the aim to learn was the opinions of participants on the level of importance of the various outcome items in terms of inclusion in a COS on hand fractures and joint injuries. It also clarified that it was extremely important that participants complete all three rounds, the focus was on 'what' to measure and not 'how' to measure, and that they would have opportunity to suggest additional outcomes at the end of the first round questionnaire.

For the second round the instructions made clear that some additional items were included after consideration of participant suggestions in Round 1. They also explained the feedback of results being provided in Rounds 2 and 3 (as detailed in Section 5.3.6).

On the question pages, there were instructions at the top detailing the scoring system we were using. These clarified that patients should rate the items based on their experience of their own injury, but professionals should do so based on their experience of managing the overall range of injuries.

5.3.3.3 Stakeholder refinement of wording

The Delphi outcome items and instructions were subsequently shown to representatives from each of the three stakeholder groups, to gain feedback on specific wording and determine whether the items, descriptors and instructions were clear or needed further change. This involved input from three patients, the British Association of Hand Therapists Research Committee and the study Steering Group which included mostly surgeons but also a hand therapist and two non-clinicians. I believed that this was an important step in trying to ensure that the language used was clear and accessible, and that the domains synthesised from the review and qualitative study outputs were deemed appropriate by stakeholder representatives. In Appendix 7, I indicate the few changes to the wording of the Delphi outcome domains resulting from this step. Probably the biggest change was the concept of ‘return to driving’ being amended to ‘being able to control a vehicle for transport’ –feedback raised the issue of some people using a motorbike or bicycle instead of a car, but the same underlying theme of hindered use of a vehicle for transport (rather than recreation) applied.

5.3.3.4 Generation of additional outcome items

At the end of the first round participants were encouraged to suggest any further outcome items which they felt were lacking amongst those presented in the questionnaire. I discussed these suggestions with the research team and made decisions on whether to add any as novel domains. The approach to analysis of this data is described in Section 5.3.9.

5.3.3.5 Scoring system

There are many potential variations on scoring including:

- Ranking of outcomes in terms of relative importance (Vargus-Adams and Martin 2009)
- Shortlisting a set number of outcomes as being the most important – done as the first round of a Delphi which subsequently used a ranking system in its second round (Ruperto et al. 2003)
- A binary yes/no choice for inclusion in a COS – this was done as the final Delphi but preceding work involved separate stakeholder group Delphi rounds with a 3-point scale (Bartlett et al. 2012)
- Points-based voting – with the allocation of a limited number of 'points' amongst Delphi outcome items based on relative importance (Kloppenburg et al. 2014)
- Rating of outcomes on Likert scales – with difference numbers of 'points' on the scale being a further factor for consideration

A binary choice does not afford participants the option of a neutral position and we felt it important that there be some granularity allowed in participants' responses. It also restricts the information participants gain through feedback of results, with magnitude of importance of each outcome to other participants effectively being concealed.

A points-based scoring system was felt to be quite demanding for patients and this, along with shortlisting a set number of outcomes or ranking of outcomes, encourage the comparing of each outcome to the others whereas we intended for outcomes to be rated for importance on an individual basis.

I therefore settled on the concept of rating of outcomes on a Likert scale, in keeping with the majority of COS development studies (Guyatt et al. 2011). Previous empirical research has had mixed findings. Some suggest that maximum reliability is found with 7- or 9-point scales (Alwin and Krosnick 1991), while others found increasing reliability from 2- to 3- to 5-point scales but no real benefit to a scales with 7 points or more (Lissitz and Green 1975). Some have even used theoretical models and concluded that there is no "magical number" which is correct for all applications (Givon and Shapira 1984).

Of course, the context of the question being posed might play a role in selecting an appropriate Likert scale. Considering COS development studies which have tackled this query, one directly compared a 3-point and 9-point Likert scales and found that nearly twice as many outcome were selected as 'critical' for inclusion in a COS when using the latter scale (De Meyer et al. 2019). The authors suggested that a 3-point scale might therefore be recommended if the Delphi aims to determine the final consensus for a COS, while a 9-point scale might be preferable if the Delphi is informing a further consensus process such as a consensus meeting. However, a subsequent study comparing 5-pont and 9-point Likert scales found that fewer outcomes reached a consensus for inclusion in their COS on the 9-point scale (Remus et al. 2021).

I selected a 9-point Likert scale, with each tertile of numerical scores given an 'importance label': 1-3 – 'less important/not important'; 4-6 – 'important but not very important'; and 7-9 – 'very important'. This is in accordance with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach that has been recommended for selecting outcomes and deciding on their relative importance in clinical decision-making (Guyatt et al. 2011) and is similar to the scale in the RAND/UCLA Appropriateness Method (Fitch et al. 2001). The approach has been used by several COS development studies and some have also included an option of ‘unable to score’ which we opted to include as well (Williamson et al. 2017).

5.3.3.6 Platform

The DelphiManager platform was used to host the Delphi study (COMET Initiative 2020). This is an online system developed by the COMET Initiative to aid in the creation and management of Delphi surveys and includes inbuilt features such as reminding participants of their own rating for each item from the previous round and facilitating email reminders to those who have not yet completed a given round.

DelphiManager was able to facilitate the selected scoring system, with a modification made to the default ‘importance labels’ used by the platform.

5.3.3.7 Ordering of outcome items

There has been research on questionnaire structuring which demonstrates that the order of presentation of questions can affect responses and response rates (McColl et al. 2001). I considered presenting outcome items in a random order that varied from one participant to the next. While possibly the most methodologically rigorous approach that could avoid any potential effects secondary to the order in which outcome items were presented, the opinion of the study Steering Group was that it would increase the risk of cognitive burden to participants and thereby reduce response rates.

I therefore grouped Delphi outcome items according to themes akin to the outcome themes generated in the qualitative study (Chapter 4). Grouping is accommodated in the

DelphiManager platform for presentation of survey items in the first round. I availed of the option in DelphiManager to randomise the order in which those groups of items were presented to each participant. This was done to mitigate a scenario where participants might tend to fatigue midway through the round and start giving ratings without due consideration to only the later outcome items. The same ordering was maintained in subsequent rounds for each individual participant.

Grouping by qualitative study outcome themes was not revealed to participants – they were presented with the various outcome items in groups but not informed of the group labels.

5.3.4 Defining consensus criteria

Consensus criteria have varied across COS development studies. I selected a combination of a percentage-based system using both the upper and lower tertiles of the 9-point Likert scale as cited in the COMET Handbook (Williamson et al. 2017), as well as specifying the condition that the percentage criteria must be met by all three of the stakeholder groups. This was stricter than having criteria set for a single overall panel, but helped to avoid the issue of prioritisation weighting being influenced by any inequality in proportions of stakeholder recruitment. The consensus criteria for any given item were:

- Consensus in – rated 7-9 by ≥70% and 1-3 by ≤15% by all three stakeholder groups
- Consensus out – rated 1-3 by ≥70% and 7-9 by ≤15% by all three stakeholder groups
- No consensus – all other item rating distributions

For the purposes of calculating these percentages, any response of ‘unable to score’ was removed from the denominator of respondents for that participant’s stakeholder group. The ‘70/15%’ consensus threshold was based on this demonstrating that the majority of participants agree that an item is either of critical or little/no importance (depending on whether the ‘consensus in’ or ‘consensus out’ thresholds are being considered) while only a small minority agree to the contrary (Williamson et al. 2012). The condition requiring stakeholder groups to individually reach the threshold was based on the approach of another COS development study where at least three of the four stakeholder groups had to reach the consensus threshold (Schmitt et al. 2011). In a more recent one published subsequent to the SO-HANDI Delphi protocol development, all stakeholder groups had to reach the threshold (Alkhaffaf et al. 2021).

Categorisation of outcome items using the above consensus criteria was noted at the end of each round, but the final status of outcome items was only decided after the third round.

5.3.5 Management of Delphi outcome items reaching consensus threshold in first or second round

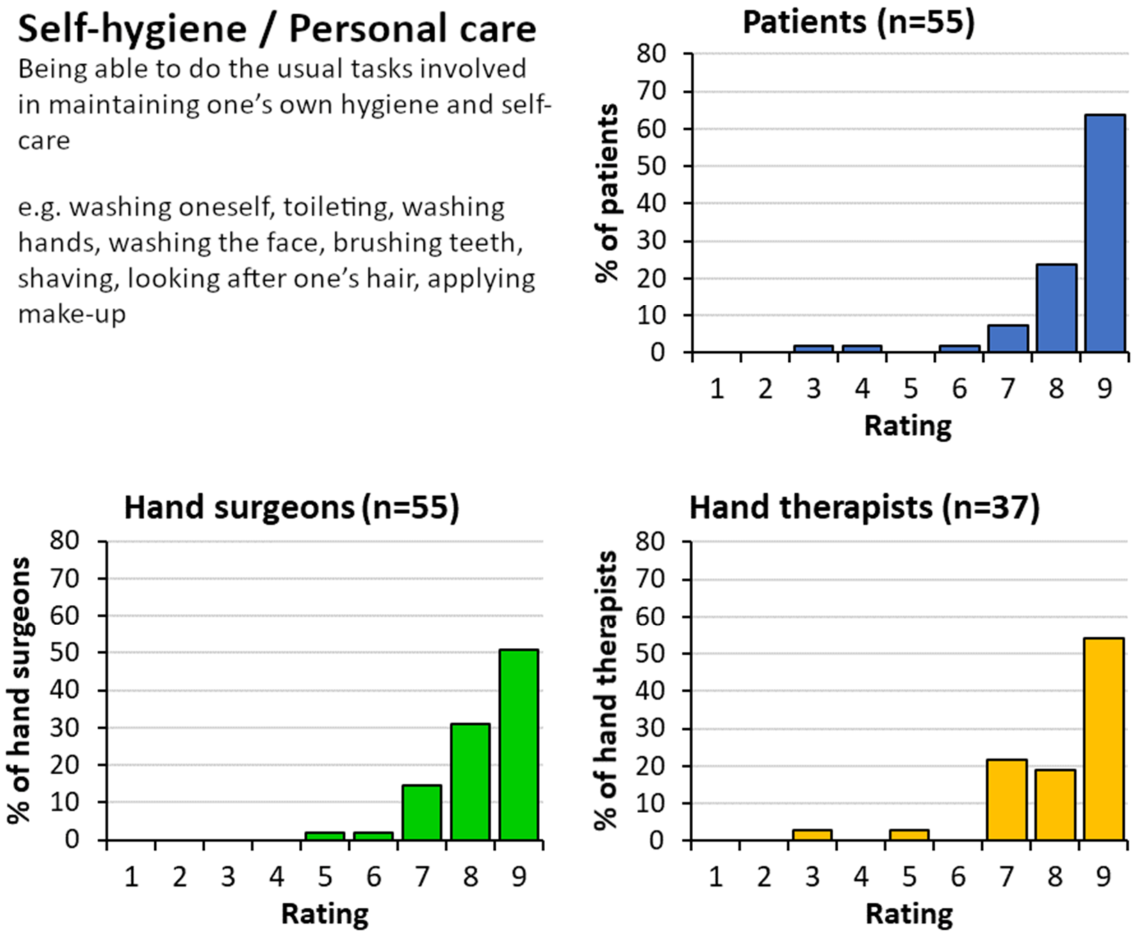
I decided to retain rather than drop outcome items reaching consensus in or out in the first or second rounds. Dropping items that reach consensus could be of benefit if the initial longlist is large, as this would reduce the burden of subsequent rounds on participants (Williamson et al. 2017). However, outcome items might reach a consensus for critical importance or little/no importance in the first round, and if dropped at this point there would have been no opportunity for feedback of results to participants. While there would likely only be niche scenarios where the feedback might lead to a shift away from the pre-defined consensus criteria, I felt that the starting point of 37 items was low enough that retaining all of these would not be too great a burden on participants. Furthermore, stability of response has also been regarded as an endpoint for Delphi surveys (Rowe and Wright

1999), and keeping all outcome domains in until the end would help to establish whether stability of responses was achieved.

5.3.6 Feedback of results in Rounds 2 and 3

For Rounds 2 and 3, each outcome item was presented on a separate page, with the descriptor and then three bar charts representing the distribution of ratings according to each stakeholder group. An example is presented in Figure 5.1.

Figure 5.1 Example of feedback presented in Rounds 2 and 3 for each outcome domain



Self-hygiene / Personal care

Being able to do the usual tasks involved in maintaining one's own hygiene and self-care

e.g. washing oneself, toileting, washing hands, washing the face, brushing teeth, shaving, looking after one's hair, applying make-up

I selected this over a 'whole panel' summary of results which has been used by some COS developers, as it allows differing views between stakeholder groups to be made apparent to participants. Some suggest that this approach might improve consensus between the groups

(Brookes et al. 2016). Although one study noted no difference in individual versus combined stakeholder group feedback, their Round 1 results began with a high level of agreement which might have masked any potential impact of the different feedback being presented to participants (MacLennan et al. 2018).

A reminder of each participant's previous rating for that item was provided at the bottom of the page, along with a rating scale to provide their rating for the current round.

This overall feedback layout and separation of each outcome to its own page was facilitated by DelphiManager and we felt participants were more likely to engage with the data than if they were presented with a separate document of the previous round's results.

5.3.7 Defining an endpoint

One of the key decisions when conducting a Delphi study is in prospectively defining an endpoint. Were the endpoint to require all or a pre-specified number or proportion of outcome items reaching a 'consensus in' or 'consensus out' there would be a theoretical concern that rather than finding a genuine consensus, the endpoint criteria are 'forcing' a consensus where none really exists. Delphi studies for COS development usually select a pre-determined number of rounds to run. A minimum of two rounds is needed in order to allow participants to have feedback on the responses of others. However, having too many rounds could lead to participant fatigue and burdening participants like this might result in higher attrition. Typically two or three rounds are selected (Williamson et al. 2017) and we opted for three rounds on the basis that if any additional outcome domains were added based on participant input in Round 1, then these would be rated by participants for the first time in

Round 2 and would only have feedback of the results and a further rating if running a third round.

5.3.8 Final study regimen

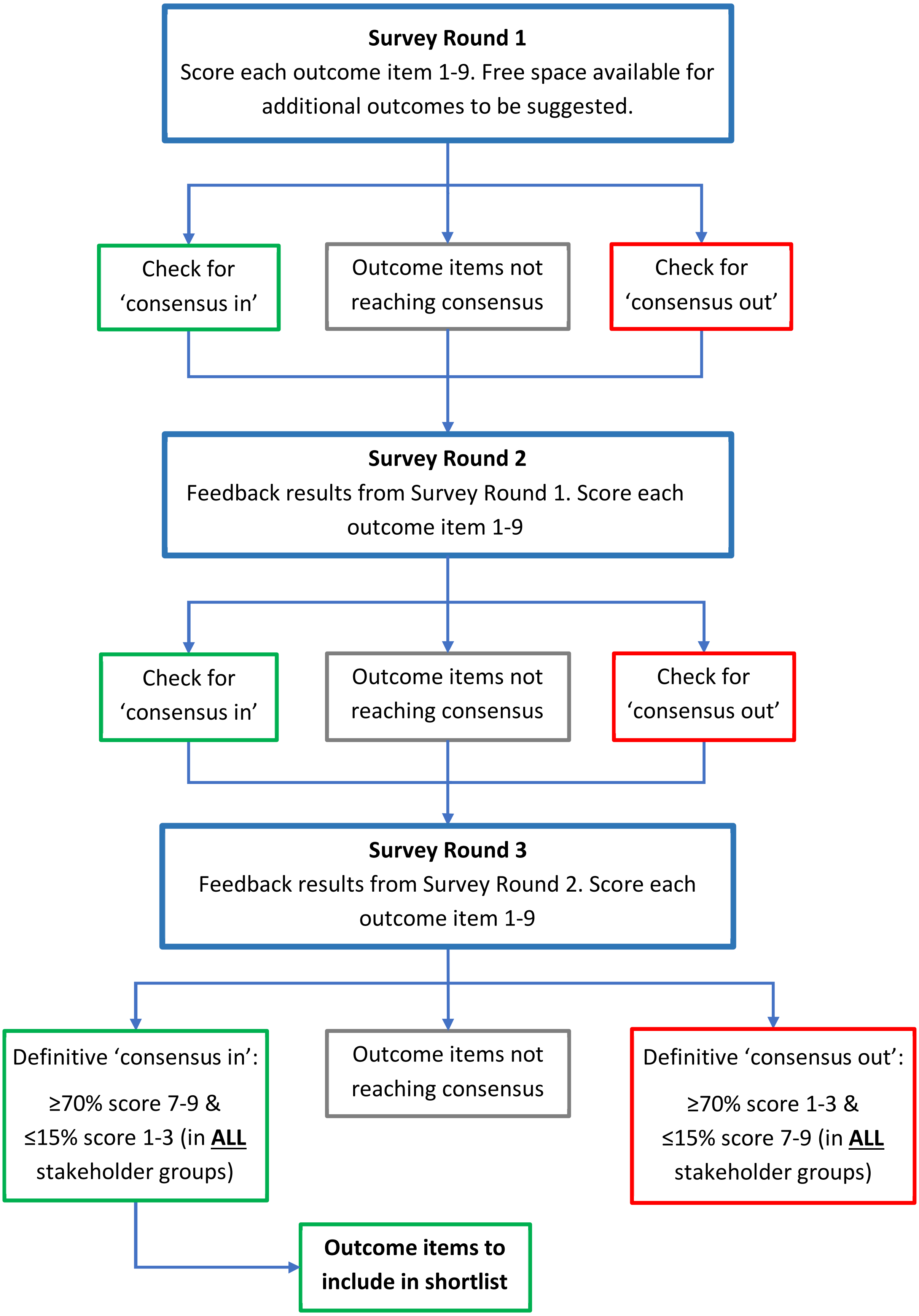
The Delphi study was conducted as per the flowchart in Figure 5.2. Each questionnaire round aimed to run for approximately 4 weeks, to allow as high a response rate as possible whilst balancing the risk of loss of interest by early responders if a round had too long a duration.

An email prompt was sent at the start of each round to all participants. Reminders were sent to those who had not yet submitted their response to a round, with the typical time points for such emails being two weeks into a round, then at three weeks, then 3-4 days before the closure of a round and finally in the last 24 hours before a round closed. Any non-responders to a round were not invited to provide ratings in remainder of the study.

An option was available for participants unable to access the online questionnaire to instead have a hard copy of the instructions and questions posted out. These were entered into the online DelphiManager system on their behalf for each round.

Up to two weeks were allocated between Rounds 1 and 2 in order to process any additional suggested outcomes and produce charts to feedback results. One week was allocated between Rounds 2 and 3 to produce the feedback charts based on participant ratings from the second round.

Figure 5.2 Flowchart illustrating the overall Delphi regimen



5.3.9 Data analysis

On completion of Round 1 all results were collated and analysed. Any outcome domains which met the consensus ‘in’ or ‘out’ threshold were noted, but all outcome items entered Round 2. In addition, suggestions from participants for additional outcomes were considered based on the following factors:

- The suggestion was an ‘outcome’
- The suggestion was a novel domain or helped to highlight a major aspect of an existing outcome item, rather than merely being a very specific example within/narrow aspect of an existing outcome item
- Number of participants who made the same or a similar suggestion

Suggestions that were deemed to be a very specific example within/narrow aspect of an existing outcome item were instead added in the descriptor of that item.

In Rounds 2 and 3 we provided feedback on the results from the previous round through bar charts created for each outcome item, showing the distribution of ratings from each stakeholder group.

The results from Round 3 were analysed to determine definitive consensus status for each outcome item. The trends of numerical changes in ratings were analysed. DelphiManager also has an option to prompt participants to provide their reasoning if a change in numerical rating led to a ‘boundary change’ (i.e. a change in the ‘importance label’ by crossing the 3/4-point or 6/7-point boundaries). I performed a rudimentary qualitative analysis of the reasons participants reported for such boundary changes, developing categories from the raw free text commentary based on an inductive, thematic approach. Finally, we also reviewed the results of those participants who were lost to follow-up over the course of the study, and performed analyses to estimate the potential impact of this attrition.

Demographic data were processed in order to convey a simple description of the overall composition of the panel.

As part of the secondary objective of the study, we compared the final results with that of a theoretical alternative Delphi method which would have removed ‘consensus in’ or ‘consensus out’ results after each round.

5.3.10 Ethics approval

The study protocol was reviewed and given favourable opinion by London (Harrow) Research Ethics Committee.

5.4 Results

The three-round Delphi which ran between December 2020 and March 2021. In total 144 participants completed all three rounds. Specific dates for each round were as follows:

- Round 1 was open from 02/12/20 to 06/01/21
- Round 2 was open from 18/01/21 to 16/02/21
- Round 3 was open from 19/02/21 to 24/03/21

5.4.1 Recruitment

While the true number of people invited to take part could not be clearly determined, we were contacted by 86 patients, 50 therapists and 60 surgeons who confirmed that they wished to participate. Of this set of potential participants, 70% (60/86) of patients, 86% (43/50) of therapists and 95% (57/60) of surgeons completed the registration form and consent statements. Of the 86 patients who expressed willingness to participate, five wished to do so by postal questionnaire – three of the five went on to complete the registration form and consent statements which were incorporated into a booklet along with the first round questionnaire. Unfortunately, one return suffered significant delay in the postal system and only reached us after the Round 2 window had opened – this response was therefore not included in the study nor in the response figures above. Response rates by stakeholder group across the rounds are shown in Table 5-1.

Table 5-1 Numbers registered and response rates per round and for overall study, by stakeholder group

Stakeholder groups	**Number registered for study**	**Response rates per round** % and number (completed/eligible to enter round)			**Completion rate of overall study for those who completed Round 1**
		1	**2**	**3**	
Patients	60	95% (57/60)	96% (55/57)	98% (54/55)	95% (54/57)
Hand therapists	43	93% (40/43)	93% (37/40)	95% (35/37)	85% (35/40)
Hand surgeons	57	96% (55/57)	100% (55/55)	100% (55/55)	100% (55/55)
Total	160	95% (152/160)	97% (147/152)	98% (144/147)	94.7% (144/152)

Demographic characteristics of the participants who completed the overall study are summarised in Table 5-2.

Table 5-2 Demographic characteristics of Delphi participants who completed all three rounds

Characteristic		**Patients (n=54)**	**Therapists (n=35)**	**Surgeons (n=55)**
Median age (interquartile range)		49 (38-62)	49 (40-53)	52 (43-59)
Sex	Male	23 (47%)	7 (20%)	44 (80%)
	Female	31 (57%)	28 (80%)	11 (20%)
Injury group (DRF – distal radius fracture non-DRF – all other hand fractures and joint injuries)	DRF	20 (37%)		
	Non-DRF	33 (61%)		
	Both	1 (2%)		
Injury management	Surgical	17 (31%)		
	Conservative	36 (67%)		
	Both	1 (2%)		
Patient occupation type	None/retired	10 (19%)		
	Office/non-manual	26 (48%)		
	Light manual	14 (26%)		
	Heavy manual	4 (7%)		
Clinician experience (years)	<5		2 (6%)	5 (9%)
	5-15		8 (23%)	13 (24%)
	15-25		12 (34%)	15 (27%)
	25+		13 (37%)	22 (40%)
Surgeon training background	Orthopaedic			48 (87%)
	Plastic			7 (13%)
Therapist training background	Occupational therapy		23 (66%)	
	Physiotherapy		12 (34%)	
Therapist accredited / certified	Yes		19 (54%)	
	No		9 (26%)	
	N/A		7 (20%)	

Presented as: number (% of n)

Table 5-2 Demographic characteristics of Delphi participants who completed all three rounds (cont'd)

Characteristic		**Patients (n=54)**	**Therapists (n=35)**	**Surgeons (n=55)**
Country of residence	Australia	0	3 (9%)	2 (4%)
	Brazil	0	0	2 (4%)
	Canada	0	0	1 (2%)
	Chile	0	1 (3%)	0
	France	0	0	1 (2%)
	Guatemala	0	0	1 (2%)
	India	0	0	1 (2%)
	Ireland	0	0	1 (2%)
	Israel	0	1 (3%)	0
	Japan	0	1 (3%)	0
	Netherlands	0	0	2 (4%)
	Norway	0	0	1 (2%)
	Philippines	0	0	1 (2%)
	Poland	0	0	1 (2%)
	Qatar	0	1 (3%)	0
	South Africa	0	5 (14%)	0
	Spain	0	0	2 (4%)
	Sweden	0	2 (6%)	2 (4%)
	Switzerland	0	2 (6%)	0
	Ukraine	0	0	1 (2%)
	United Kingdom	54 (100%)	16 (46%)	33 (60%)
	United States of America	0	3 (9%)	3 (5%)

Presented as: number (% of n)

5.4.2 Finalised outcome domains in Delphi questionnaire

Table 5-3 shows the final longlist of Delphi outcome items used for Round 1, along with descriptors developed to clarify their meaning for Delphi participants. These items were based on a synthesis of the systematic review and qualitative study outcome domains. Groups are also shown, which reflects how items were clustered together when presented in Round 1 of the questionnaire.

There were two cases where the WHO ICF domains did not appear to link to any of the outcome domains derived from the patient perspective leading to the generation of the Delphi outcome items ‘stability of joints’ and ‘thinking and memory’.

There were three cases of outcome domains from the qualitative study not having any corresponding domains from the systematic review, resulting in the creation of Delphi outcome items ‘added expenses due to injury’, ‘looking after dependents within household’ and ‘emotional/mood impact to friends and family’.

A small number of WHO ICF domains from the systematic review were not included within the Delphi outcome items (with ‘nc’ denoting ‘not covered within WHO ICF’):

- nc-Technical (related to intervention) – these were outcomes focused on some technical aspect of an intervention itself (e.g. as part of feasibility studies)
- nc-Blood tests – a very specific outcome of a single study which included these tests to assess effects of anaesthetic medication rather than the surgical intervention
- nc-Individualised rating scale – this was also used in a single study, and essentially allows patients to determine what outcome to assess on an individual basis. There is thus no classification into any one outcome domain, as the patient is at liberty to select literally any outcome they believe to be important.

A complete table showing how the review WHO ICF domains linked to qualitative study outcome domains, and the corresponding Delphi outcome items that were generated, is provided in Appendix 7.

Table 5-3 Delphi outcome items along with descriptors and groupings

Group	Delphi Outcome Item	Descriptor
Biology and physiology	Change in sensation	Change from normal feeling in the hand / wrist (either sensations in the hand or in how things feel to touch) e.g. numbness, pins and needles, tingling, being able to sense hot / cold through touch, dysesthesia – in which the things that one touches feel different to normal
	Pain / discomfort	Discomfort or pain in the hand or wrist, whether at rest or with activity e.g. ache, shooting pain, sharp pain, throbbing
	Strength or muscle power	Being able to use strength or exert power through the hand or wrist e.g. pinch between thumb and fingers, gripping with hand, moving the wrist with force
	Range of movement	How much movement one has through the joints of the hand or wrist, whether active (i.e. moving it with the muscles of the injured side) or passive (e.g. if someone else were to try to move it for the patient) Includes stiffness in the joints or how much one can bend or straighten the thumb, fingers or wrist
	Confirmation of bone healing and alignment (fractures only)	Having confirmation of healing of broken bone(s) including how well they are lined up. e.g. x-rays to show how the bones are healing (NOT cases where the bones are obviously not aligned well and from the outside appearance one can already see the change from normal shape)
	Confirmation of soft tissue healing	Having confirmation of healing of soft tissues (e.g. tendons, ligaments, surgical wound) e.g. clinical assessment or tests to confirm the soft tissues are healing.
	Stability of joints	Having the joints in the hand / wrist feel stable during use e.g. no feeling that the bones will 'come out of joint' when being used

Table 5-3 Delphi outcome items along with descriptors and groupings (cont'd)

Self-care and independence	Grasping and moving light objects	Being able to hold onto and move / lift objects weighing less than ~1kg / 2lb e.g. lifting a cup of tea, opening doors with a handle / knob, opening small containers
	Self-hygiene / personal care	Being able to do the usual tasks involved in maintaining one's own hygiene and self-care e.g. washing oneself, toileting, washing hands, washing the face, brushing teeth, shaving, looking after one's hair, applying make-up
	Cooking / preparing meals	Being able to complete the tasks involved in preparing meals for oneself to ensure proper nutrition
	Dressing	Being able to put on and take off one's own clothing and footwear
	Eating and drinking	Being able to carry out the tasks involved in eating and drinking
	Shopping for groceries	Being able to carry out the tasks involved in shopping for groceries, putting shopping away, etc (NOT the travel to and from shops)
	Walking / mobility	Being able to walk / mobilise as one usually would after an injury, and have the confidence to do this e.g. general walking, stairs, using walking aids such as stick or frame (if normally needed)
	Using hands to change body position	Being able to use one's hands to push up and stand from a chair or turn in bed
	General household chores	Being able to do general tasks around the house e.g. household cleaning, doing dishes, laundry, making the bed, putting out bins, general maintenance, gardenwork
	Sleeping	Being able to sleep and rest overnight, and the impact on activities if having trouble sleeping
	Thinking and memory	Being able to think about and remember things clearly

Table 5-3 Delphi outcome items along with descriptors and groupings (cont'd)

Wider social roles, engagement and responsibilities	Fine hand use / dexterity	Being able to do fine motor tasks or precise activities with the hand/wrist e.g. writing, drawing, picking up coins from a table, using a key
	Hobbies and recreation	Being able to do usual hobbies or pastimes e.g. sports, exercise, reading, playing an instrument, playing video games
	Return to usual work / job	Being able to return to the work or job that one was doing prior to their hand/wrist injury (NOT including the financial impact of any lost income)
	Carrying or lifting heavy objects	Being able to lift or carry objects weighing more than ~5kg / 10lb (i.e. more than a shopping bag with groceries)
	Being able to control a vehicle for transport	Being able to drive a motor vehicle or ride a bicycle, whether for work, social or recreation reasons
	Keyboard and mouse / touchscreen use	Being able to make use of keyboard and mouse or touchscreens e.g. being able to perform the hand / wrist actions needed to use devices such as a computer, laptop, tablet computer or smartphone
	Socialising / relationships with friends and family	Being able to socialise and maintain relationships with family and friends
	Fatigue / tiredness	General feeling of low energy, tiredness or fatigue which restricts how much one can do (NOT due to issues with sleeping)
	Looking after dependents within household	Being able to do the tasks involved in looking after dependents within the household e.g. caring for children, pets, elderly relative within household)
	Intimate / sexual activity	Being able to engage in one's usual intimate/sexual activity

Table 5-3 Delphi outcome items along with descriptors and groupings (cont'd)

Miscellaneous – emotional & psychological impact, personal/healthcare resources, treatment factors/complications	Appearance	How the hand / wrist looks after treatment is complete e.g. change in shape of injured hand / wrist, scars, swelling
	Emotional / mood impact to self	Effect on mood / emotions to oneself e.g. anxiety, depression / low mood, frustration, anger
	Emotional / mood impact to friends and family	Effect of one's injury on mood / emotions of their friends and family e.g. anxiety, depression / low mood, frustration, anger
	Overall patient satisfaction	Overall satisfaction with the treatment, recovery process and result from the patient's perspective
	Added expenses due to injury	The increased costs to oneself while recovering from a hand / wrist injury (this is separate to the lost income from not being able to work). e.g. using public transport / taxi, buying ready meals, buying different clothing / footwear to accommodate cast/splint
	Loss of income	Any loss of income due to the hand / wrist injury from not being able to do usual work
	Healthcare resources used	Medical resources used in providing healthcare for the hand / wrist injury. Includes cost of treatment, follow-up, etc until end of recovery
	Treatment complications	The risks of unwanted and unexpected events due to a treatment e.g. infection after surgery, anaesthetic risks, pressure sore / blister due to cast
	Treatment inconvenience / discomfort	The inconvenience or discomfort caused directly by a treatment e.g. inconvenience / discomfort with having a cast or splint on, dressings changes and stitch removal after surgery

Suggestions for additional outcome domains from participants were considered by the research group on the basis of the factors specified earlier, resulting in the addition of four outcome domains in Round 2 as per Table 5-4. A detailed explanation of the consideration of all participant suggestions is provided in Appendix 8.

Table 5-4 Additional Delphi outcome items (added after Round 1) with descriptors and groupings

Group	**Delphi Outcome Item**	**Descriptor**
Biology and physiology	Pain/discomfort during activity	Discomfort or pain in the hand or wrist specifically during activities (NOT at rest) e.g. ache, shooting pain, sharp pain, throbbing, discomfort/pain due to not being able to tolerate hot or cold sensation
	Pain/discomfort during rest	Discomfort or pain in the hand or wrist specifically during rest (i.e. with the hand/wrist not moving, so NOT during activities) e.g. ache, shooting pain, sharp pain, throbbing, discomfort/pain due to not being able to tolerate hot or cold sensation
	Speed of movement	How quick movements through the joints of the hand or wrist can be, including any issue of delay in being able to start the movement
Miscellaneous – emotional & psychological impact, personal/healthcare resources, treatment factors/complications	Patient satisfaction with outcome/result	Satisfaction with the overall result from the patient's perspective (NOT with treatment or recovery process, but the end result only)

5.4.3 Final consensus status of outcome domains

Given that the chosen approach involved retaining all outcome domains through the rounds of the study, and that each domain's consensus status was set by whether or not consensus criteria were reached by the end of the third round, a summary of the results from simply the final round is informative. This is presented in Tables 5-5 and 5-6, showing the results for outcome domains that reached 'consensus in' and those that did not, respectively.

Table 5-5 Rating threshold results for outcome domains reaching 'consensus in' status by end of Delphi study

Outcome domain	**Patients (Total = 54)**			**Surgeons (Total = 55)**			**Therapists (Total = 35)**			**Combined (Total = 144)**		
	% rating 1-3	**% rating 7-9**	**Number rating**	**% rating 1-3**	**% rating 7-9**	**Number rating**	**% rating 1-3**	**% rating 7-9**	**Number rating**	**% rating 1-3**	**% rating 7-9**	**Number rating**
Fine hand use / dexterity	0	98.1	54	0	100	55	0	97.1	35	0	98.6	144
Return to usual work / job	3.8	94.3	53	0	98.2	55	0	100	35	1.4	97.2	143
Grasping and moving light objects	0	98.1	54	0	96.4	55	0	94.3	35	0	96.5	144
Pain / discomfort	0	94.4	54	0	100	55	0	94.3	35	0	96.5	144
Self-hygiene / personal care	1.9	98.1	54	0	94.5	55	0	94.3	35	0.7	95.83	144
Patient satisfaction with outcome / result	0	96.3	54	1.8	89.1	55	0	100	35	0.7	94.4	144
Pain / discomfort during activity	0	94.4	54	0	90.9	55	0	97.1	35	0	93.8	144
Stability of joints	0	98.1	54	0	83.3	54	0	100	35	0	93	143
Overall patient satisfaction	0	92.6	54	0	87.3	55	0	100	35	0	92.4	144
Pain / discomfort during rest	0	96.3	54	1.8	89.1	55	0	91.4	35	0.7	92.4	144
Treatment complications	1.9	84.9	53	0	92.7	55	0	97.1	35	0.7	90.9	143
Strength or muscle power	0	98.1	54	0	83.6	55	0	91.4	35	0	91	144
Eating and drinking	0	94.4	54	0	87.3	55	0	88.6	35	0	90.3	144
Range of movement	0	88.9	54	0	72.7	55	0	100	35	0	85.4	144
Change in sensation	0	85.2	54	0	87.3	55	0	88.6	35	0	86.8	144
Dressing	0	85.2	54	0	80	55	2.9	88.6	35	0.7	84	144
Sleeping	5.6	83.3	54	3.6	87.3	55	0	74.3	35	3.5	82.6	144
Being able to control a vehicle for transport	3.7	88.9	54	1.8	72.7	55	0	80	35	2.1	80.6	144
Cooking / preparing meals	0	81.5	54	0	70.9	55	0	82.9	35	0	77.8	144
Keyboard and mouse / touchscreen use	5.8	76.9	52	0	78.2	55	0	80	35	2.1	78.2	142

Table 5-6 Rating threshold results for outcome domains that did not reach 'consensus in' status by end of Delphi study

Outcome domain	**Patients (Total = 54)**			**Surgeons (Total = 55)**			**Therapists (Total = 35)**			**Combined (Total = 144)**		
	% rating 1-3	**% rating 7-9**	**Number rating**	**% rating 1-3**	**% rating 7-9**	**Number rating**	**% rating 1-3**	**% rating 7-9**	**Number rating**	**% rating 1-3**	**% rating 7-9**	**Number rating**
Hobbies and recreation	0	**81.5**	54	1.8	**65.5**	55	0	**82.9**	35	0.7	75.7	144
Emotional / mood impact to self	3.7	**77.8**	54	0	**63.6**	55	2.9	**80**	35	2.1	72.9	144
Loss of income	15.7	**66.7**	51	5.5	**67.3**	55	0	**71.4**	35	7.8	68.1	141
Confirmation of soft tissue healing	1.9	**86.6**	52	1.9	**44.4**	54	2.9	**68.6**	35	2.1	66	141
Carrying or lifting heavy objects	3.7	**72.2**	54	1.8	**54.5**	55	0	**65.7**	35	2.1	63.9	144
Looking after dependents within household	12.2	**69.4**	49	1.8	**41.8**	55	5.7	**80**	35	6.5	61.2	139
Treatment inconvenience / discomfort	5.6	**72.2**	54	10.9	**50.9**	55	5.7	**62.9**	35	7.6	61.8	144
Confirmation of bone healing and alignment (fractures only)	0	**88.2**	51	7.4	**35.2**	54	0	**60**	35	2.9	60.7	140
Using hands to change body position	1.9	67.9	53	3.6	58.2	55	5.7	45.7	35	3.5	58.7	143
General household chores	5.6	61.1	54	1.8	49.1	55	2.9	57.1	35	3.5	55. 6	144
Healthcare resources used	5.6	68.5	54	10.9	27.3	55	0	65.7	35	6.3	52.1	144
Speed of movement	0	**72.2**	54	14.5	**25.5**	55	8.6	**48.6**	35	7.6	48.6	144
Socialising / relationships with friends and family	7.7	50	52	10.9	43.6	55	8.6	48.6	35	9.2	47.2	142
Shopping for groceries	5.6	59.3	54	5.5	29.1	55	5.7	28.6	35	5.6	40.3	144
Walking / mobility	12	58	50	21.8	32.7	55	54.3	20	35	26.4	38.6	140
Appearance	13.0	42.6	54	16.4	38.2	55	5.7	25.7	35	12.5	36.8	144
Thinking and memory	20.8	50	48	35.2	31.5	54	48.6	22.9	35	33.6	35.8	137
Added expenses due to injury	20.4	38.9	54	21.8	23.6	55	14.3	40	35	19.4	33.3	144
Fatigue / tiredness	7.8	54.9	51	23.6	23.6	55	17.1	20	35	16.3	34.0	141
Intimate / sexual activity	21.2	34.6	52	13.0	38.9	54	20	20	35	17.7	32.6	141
Emotional / mood impact to friends and family	13.5	40.4	52	21.8	16.4	55	8.6	22.9	35	15.5	26.8	142

Some outcome domains did not meet the percentage-based criteria in all three of the stakeholder groups to reach ‘consensus in’ status, but did meet the criteria for in one or two groups. The domains that reached the criteria in only two stakeholder groups were:

- Hobbies and recreation
- Emotional / mood impact to self

The domains that reached the criteria in only one stakeholder group were:

• Confirmation of bone healing and alignment (fractures only) • Confirmation of soft tissue healing • Treatment inconvenience / discomfort • Carrying or lifting heavy objects • Speed of movement	Patients only
• Looking after dependents within household • Loss of income	Therapists only

Consensus status if determined using overall (combined) ratings with consensus thresholds

The results based on a pooled panel of all three stakeholder groups are shown in Tables 5-5 and 5-6. Had the consensus criteria been set so that the thresholds were applied to the combined ratings instead, then two further outcome domains would have reached ‘consensus in’ status: ‘hobbies and recreation’ and ‘emotional / mood impact to self’.

5.4.4 Change of ratings between rounds

Appendix 9 shows a crude summary of the ratings in the form of means and standard deviations for each outcome domain per study round. These figures convey a sense of the overall changes in ratings taking place within each stakeholder group across rounds and changes in the variance in ratings. However, reciprocal changes in rating within a stakeholder group are masked (i.e. if some participants increase their ratings but a similar number decrease their rating, then there will be minimal net change). Since overall stability of responses across rounds can be an alternative endpoint in a Delphi (Rowe and Wright 1999), I believed it worthwhile to explore the changes of ratings in more detail.

Table 5-7 shows a summary of the changes in ratings across all outcome domains per stakeholder group across the Delphi rounds. This reflected the overall stability of scoring in a very broad sense and did not distinguish between outcome domains nor on how many participants changed their ratings. Appendix 10 provides per domain data on the number of individual participants who changed their score for each domain between rounds, the magnitude and direction of their changes in score, and the number of participants who did not change their score.

Table 5-7 Summary of changes in ratings across all outcome domains between rounds, per stakeholder group

		Round 1 to 2	**Round 2 to 3**
Patients' ratings	Number that decreased	279/1969 (14%)	94/2173 (4.3%)
	Mean decrease	-1.9	-1.3
	Number with no change	1216/1969 (62%)	1788/2173 (82%)
	Number that increased	474/1969 (24%)	291/2173 (13%)
	Mean increase	+1.9	+1.8
Surgeons' ratings	Number that decreased	188/2030 (9.3%)	157/2246 (7.0%)
	Mean decrease	-1.3	-1.2
	Number with no change	1512/2030 (74%)	1904/2246 (85%)
	Number that increased	330/2030 (16%)	185/2246 (8.2%)
	Mean increase	+1.4	+1.3
Therapists' ratings	Number that decreased	197/1364 (14%)	70/1426 (4.9%)
	Mean decrease	-1.9	-0.9
	Number with no change	884/1364 (65%)	1207/1426 (85%)
	Number that increased	283/1364 (21%)	149/1426 (10%)
	Mean increase	+1.4	+1.4

Note: Numbers of ratings above are the product of number of participants and number of outcome domains

5.4.4.1 'Significant' changes in ratings

Given that the consensus criteria were based upon percentages rating within the top and bottom importance categories, I proceeded with a closer assessment of the changes of ratings resulting in such a 'boundary change'. There were also instances when participants would transition from being 'unable of score' a domain to then giving it a rating on the 1-9 scale, or vice versa – while not technically 'boundary changes' these too were considered 'significant' and part of the overall stability of responses.

Table 5-8 summarises the numbers of participants in each stakeholder group who made such changes in their ratings along with an indication of the number of domains for which changes were made per participant. This table reflects the overall stability in terms of the significant changes only, and suggested that for all three stakeholder groups there was increasing stability in ratings – there were generally more participants tending to make no, or fewer, domain ratings changes between Rounds 2 and 3 compared to between Rounds 1 and 2. Again, a detailed breakdown per outcome domain (Appendix 11) was needed to identify whether any specific domains were more or less stable than others. The vast majority of domains (31/37, 84% – out of 37 as the four domains added after Round 1 only had one point of reference for a boundary change) had fewer boundary changes as the Delphi survey progressed.

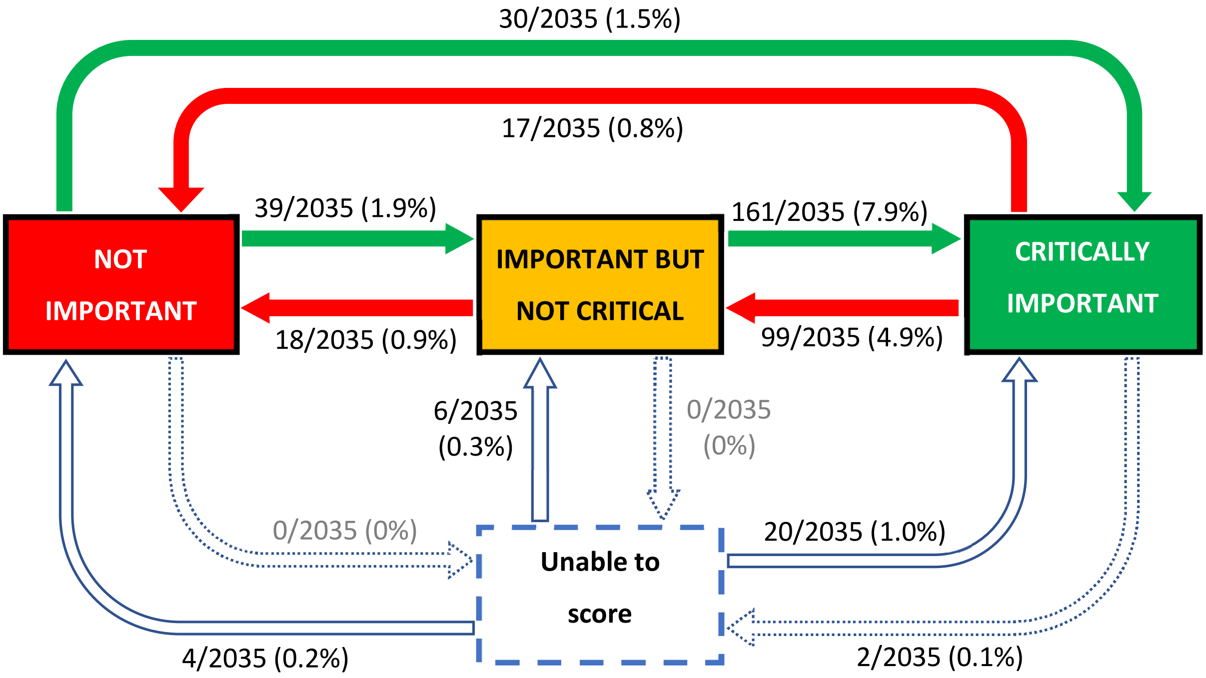
Table 5-8 Number of participants per stakeholder group making significant changes in their ratings between rounds and corresponding number of domains affected

Stakeholder group			**Round 1 to 2**	**Round 2 to 3**
Patients	**Number of domains with boundary changes**	0	5/55 (9.1%)	24/54 (44%)
		1-2	11/55 (20%)	15/54 (28%)
		3-5	16/55 (29%)	6/54 (11%)
		6-10	12/55 (22%)	7/54 (13%)
		11-15	4/55 (7.3%)	1/54 (1.9%)
		>15	7/55 (13%)	2/54 (3.7%)
	Number of domains changing to or from 'unable to score'	0	43/55 (78%)	49/54 (91%)
		1	7/55 (13%)	4/54 (7.4%)
		2-3	1/55 (1.8%)	1/54 (1.9%)
		4-5	2/55 (2.2%)	0/54 (0%)
		6-10	2/55 (2.2%)	0/54 (0%)
Surgeons	**Number of domains with boundary changes**	0	9/55 (16%)	24/55 (44%)
		1-2	15/55 (27%)	19/55 (35%)
		3-5	14/55 (25%)	8/55 (15%)
		6-10	16/55 (29%)	3/55 (5.5%)
		11-15	1/55 (1.8%)	1/55 (1.8%)
	Number of domains changing to or from 'unable to score'	0	55/55 (100%)	54/55 (98%)
		1	0/55 (0%)	0/55 (0%)
		2-3	0/55 (0%)	0/55 (0%)
		4-5	0/55 (0%)	1/55 (1.8%)
Therapists	**Number of domains with boundary changes**	0	0/37 (0%)	9/35 (26%)
		1-2	10/37 (27%)	18/35 (51%)
		3-5	17/37 (46%)	5/35 (14%)
		6-10	9/37 (24%)	2/35 (5.7%)
		11-15	0/37 (0%)	1/35 (2.9%)
		>15	1/37 (2.7%)	0/35 (0%)
	Number of domains changing to or from 'unable to score'	0	36/37 (97%)	31/35 (89%)
		1	0/37 (0%)	2/35 (5.7%)
		2-3	0/37 (0%)	1/35 (2.9%)
		4-5	1/37 (2.7%)	1/35 (2.9%)

Figures 5.3, 5.4 and 5.5 summarise the overall flow of changes in these ratings between rounds for patients, surgeons and therapists, respectively. The denominators represent the theoretical maximum number of ratings changes that could have taken place.

Figure 5.3 Flow of significant rating changes between rounds for patients

a) Patients Round 1 to Round 2 changes



b) Patients Round 2 to Round 3 changes

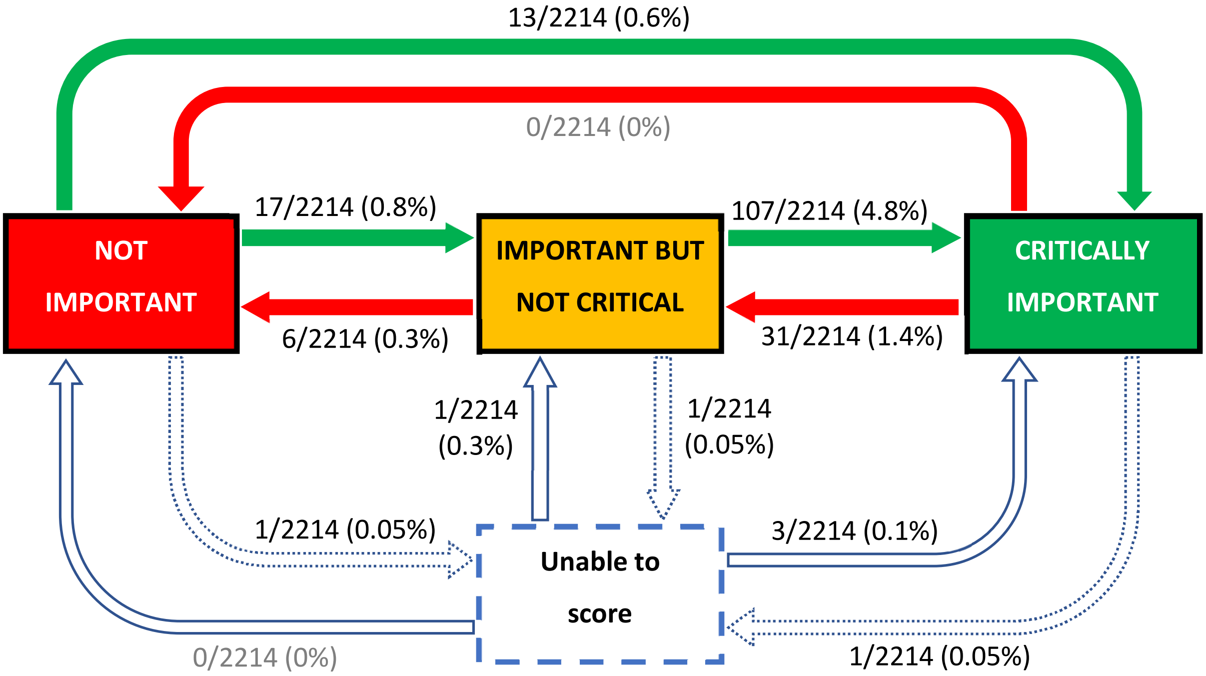
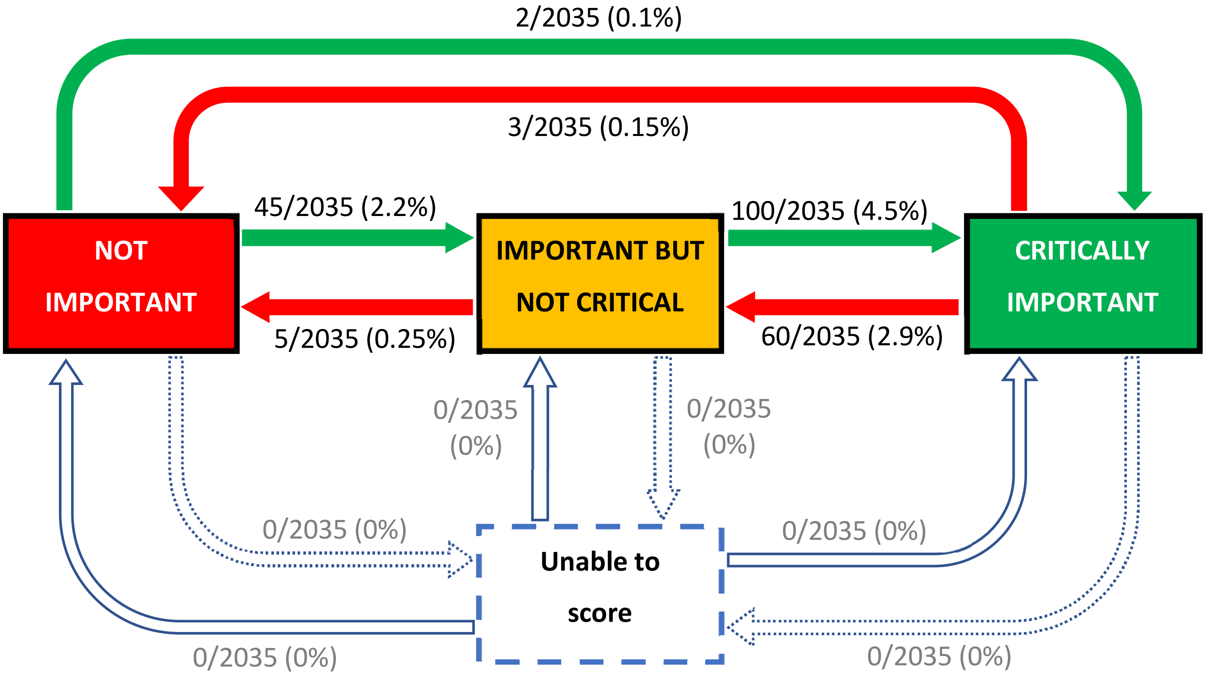


Figure 5.4 Flow of significant rating changes between rounds for surgeons

a) Surgeons Round 1 to Round 2 changes



b) Surgeons Round 2 to Round 3 changes

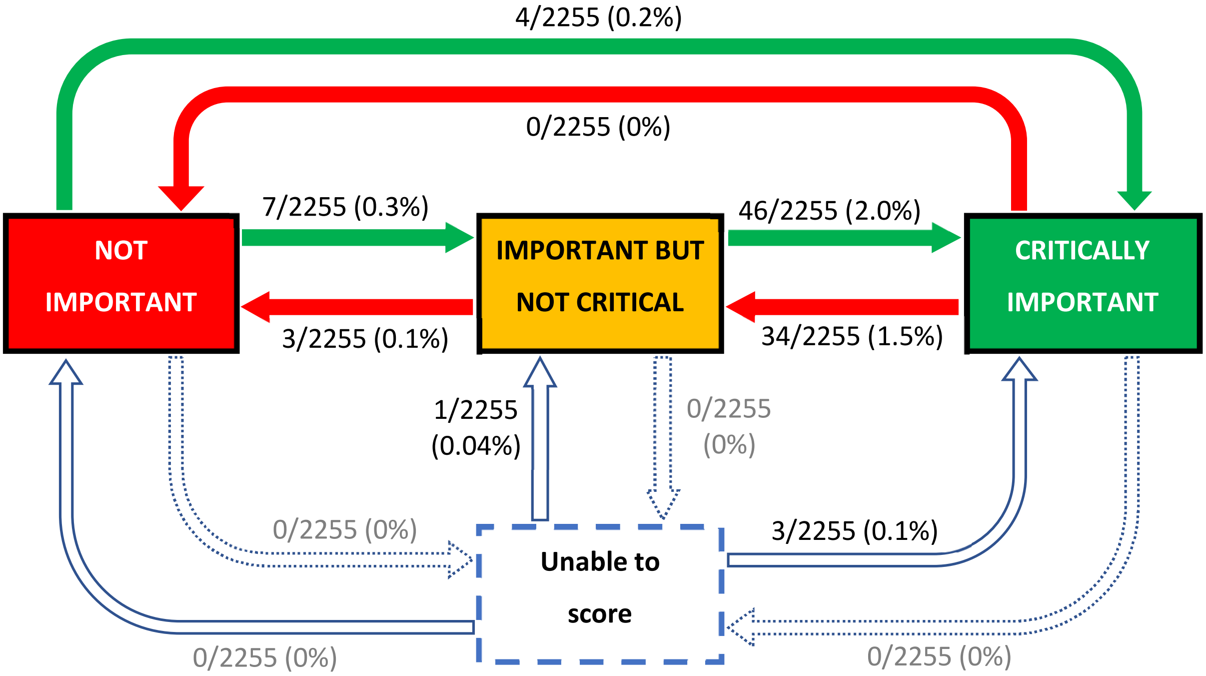
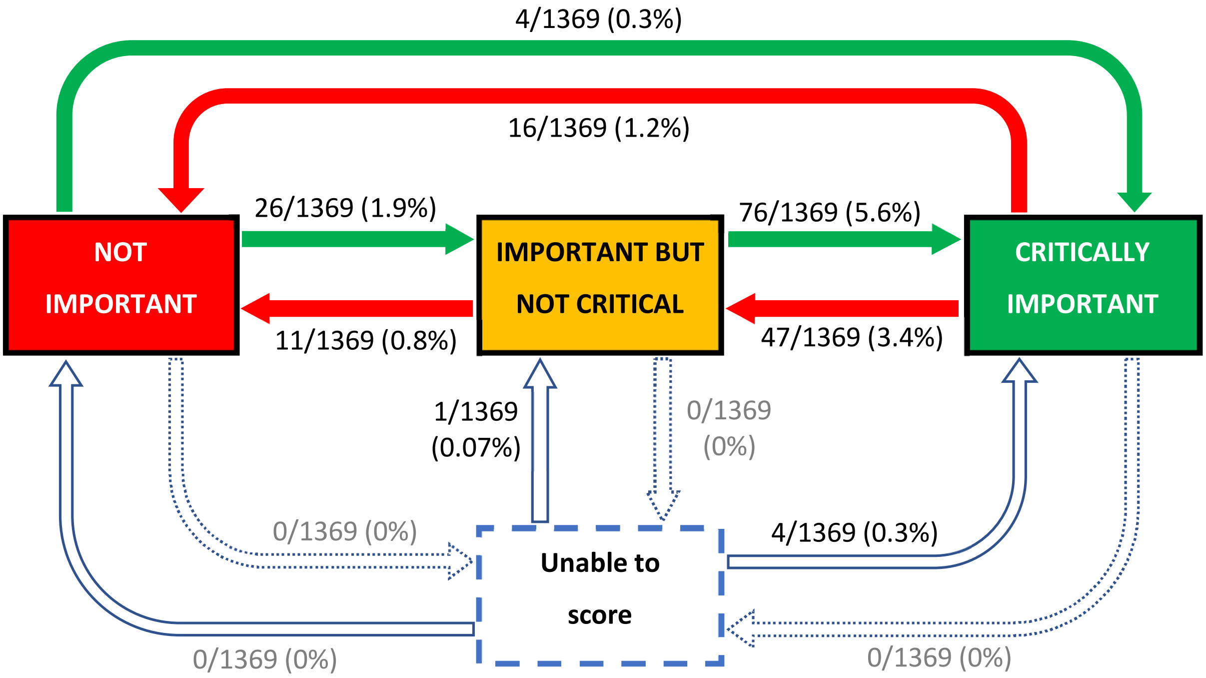
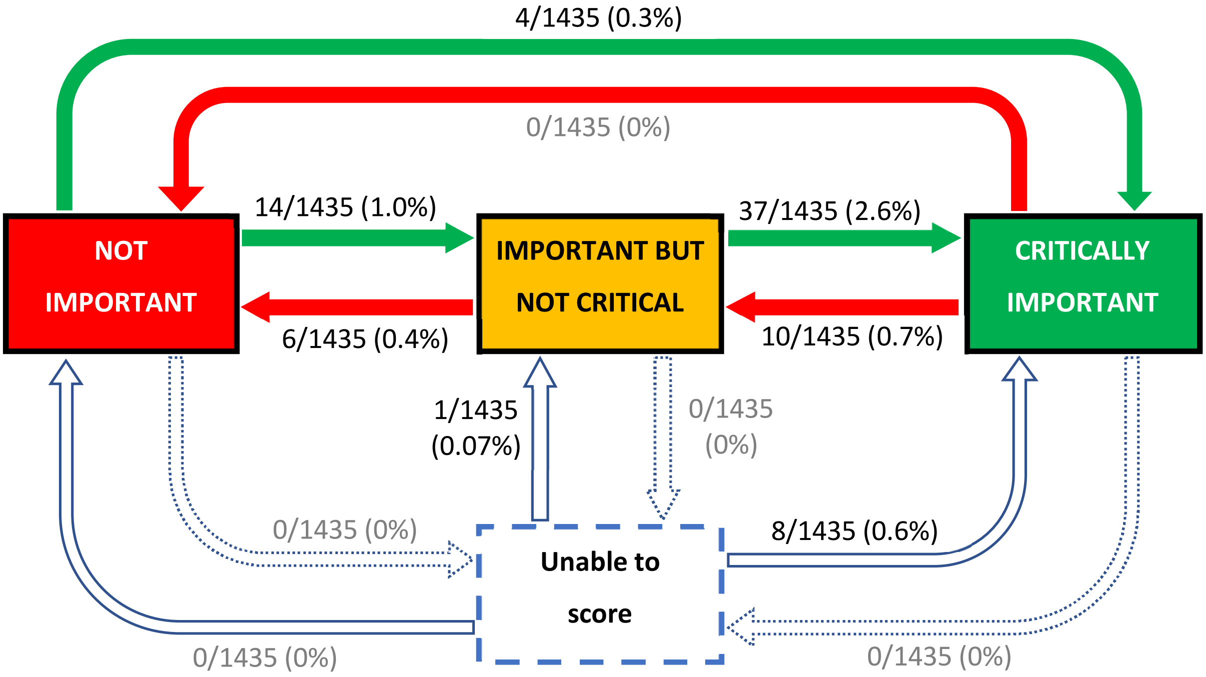


Figure 5.5 Flow of significant rating changes between rounds for therapists

a) Therapists Round 1 to Round 2 changes



b) Therapists Round 2 to Round 3 changes



In each stakeholder group the total number of significant ratings changes was higher between Rounds 1 and 2 than between Rounds 2 and 3, despite there being more potential for change in the latter stage due to the four additional outcome domains. There was an overall flow of increasing the ratings of importance, particularly in the patient group.

Through DelphiManager, participants who had a ‘boundary change’ or a transition to or from a ‘unable to score’ rating were prompted to provide free text commentary on the reason motivating the change. As Table 5-8 shows, several made such changes to more than one domain. Correspondingly, when analysing the reasons provided by participants some appeared more common simply because a participant stated them numerous times. Appendix 12 provides data on the detailed breakdown of categories of reasons along with frequency and the number of participants who stated each reason as a cause for their rating change. The description below focuses on the latter figure though. These are based on a thematic analysis type of approach, but the baseline data was often very brief and certainly not of the quality that might be obtained through formal interview.

5.4.4.1.1 Rounds 1 to 2

Between Rounds 1 and 2 patients made significant rating changes in 19% (396/2035) of possible cases, while surgeons did so for 11% (215/2035) and therapists for 14% (185/1369).

Patients provided commentary for 171 of the 396 changes. The most common reason related to the patients’ experience of symptoms or problems as a result of their injury, as time progressed – it appeared that they had either under- or overestimated the importance of the domain. Some patients reported more specifically that a change in rating was prompted by increasing activity as recovery progressed, bringing to light certain domains that were not considered a priority at the time of the previous round. Only one patient made a significant change to their rating in an attempt to conform to the general Delphi panel's scores, and this was for only one domain.

Surgeons gave feedback on 141 of the 215 changes. Most surgeon participants making such changes declared their reasoning as trying to align better with the ratings of the patient stakeholder group, while some instead mentioned an effort to conform better with the general Delphi panel's scores. Several mentioned that they had not fully understood a domain previously or had a subsequent change in their perception of the domain.

Therapists commented on 106 of the 185 changes and their top three reasons were the same as the surgeons', but the commonest was to do with not fully grasping the meaning of a domain or a change in perception which lead to changing their rating significantly.

5.4.4.1.2 Rounds 2 to 3

Between Rounds 2 and 3 patients made significant changes in just 8.2% (181/2214) of ratings, surgeons for 4.3% (98/2255) and therapists for 5.6% (80/1435).

Patients commented on 68 of 181 changes, with the same top three reasons as between Rounds 1 and 2. This time two patients made a change to conform to the general Delphi panel's ratings.

Surgeons provided reasons for 69 of the 98 changes and again most commonly changed scores to align with the patient group or the Delphi panel as a whole. As might be expected, a change in how the domain was perceived became much less of an issue between these later rounds of the Delphi.

Finally, therapists provided feedback on 55 of the 80 significant changes. As per the surgeon group, therapists continued to make changes to conform to either patients or the overall Delphi panel. Some other reasons given by a few included: misunderstanding the true nature of the domain; believing that there were workarounds for problems in certain domains (bypassing any issues); and that on reconsideration a domain may be important to a subset of patients.

5.4.5 Change of consensus between rounds

Despite the previous consideration of 'significant' changes in scoring, consensus status would only change if meeting the percentage-based pre-defined criteria. Table 5-9 provides a summary of whether or not each domain met 'consensus in' criteria per stakeholder group. The last four outcome domains in the table were introduced in Round 2, and three of these were 'consensus in' by the end of Round 2 for all three stakeholder groups. Considering just the original 37 outcome domains present from the start of the study, then:

- Patients rated 18 'consensus in' in Round 1, which rose to 22 'consensus in' in Round 2. For Round 3, only one additional domain of the 37 joined the 'consensus in' group.
- Surgeons had the fewest domains reaching 'consensus in' at each point in the Delphi, with just 14 in Round 1. This increased by just one of the original 37 in Round 2, and a further two in Round 3.
- Therapists rated 18 as 'consensus in' after Round 1 and added another three of the original 37 in Round 2. There were no changes in domain consensus status in Round 3.

Most changes in consensus status occurred between Rounds 1 and 2 and almost always gaining rather than losing 'consensus in' status.

Table 5-9 Summary of outcome domains reaching 'consensus in' status per stakeholder group across all three rounds

Outcome domain	Patients			Surgeons			Therapists		
	1	2	3	1	2	3	1	2	3
Change in sensation	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Pain / discomfort	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Strength or muscle power	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Range of movement	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟩
Confirmation of bone healing and alignment (fractures only)	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥
Confirmation of soft tissue healing	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥
Stability of joints	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Grasping and moving light objects	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Self-hygiene / personal care	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Cooking / preparing meals	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩
Dressing	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Eating and drinking	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Shopping for groceries	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Walking / mobility	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Using hands to change body position	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
General household chores	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Sleeping	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟩	🟩
Thinking and memory	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Fine hand use / dexterity	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Hobbies and recreation	🟥	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟩
Return to usual work / job	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Carrying or lifting heavy objects	🟥	🟩	🟩	🟥	🟥	🟥	🟩	🟥	🟥
Being able to control a vehicle for transport	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩
Keyboard and mouse / touchscreen use	🟥	🟥	🟩	🟩	🟩	🟩	🟥	🟩	🟩
Socialising / relationships with friends and family	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Fatigue / tiredness	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Looking after dependents within household	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟩
Intimate / sexual activity	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Appearance	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Emotional / mood impact to self	🟥	🟩	🟩	🟥	🟥	🟥	🟩	🟩	🟩
Emotional / mood impact to friends and family	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Overall patient satisfaction	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Added expenses due to injury	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Loss of income	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩
Healthcare resources used	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
Treatment complications	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
Treatment inconvenience / discomfort	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥
Speed of movement	⬛	🟥	🟩	⬛	🟥	🟥	⬛	🟥	🟥
Pain / discomfort during activity	⬛	🟩	🟩	⬛	🟩	🟩	⬛	🟩	🟩
Pain / discomfort during rest	⬛	🟩	🟩	⬛	🟩	🟩	⬛	🟩	🟩
Patient satisfaction with outcome / result	⬛	🟩	🟩	⬛	🟩	🟩	⬛	🟩	🟩
Total	**18**	**25**	**27**	**14**	**18**	**20**	**18**	**24**	**24**

Green denotes that the 'consensus in' criteria were met, while red indicates that they were not.

5.4.6 Attrition

Overall response rate in terms of participants who completed Round 1 going on to complete the study overall (i.e. both Rounds 2 and 3) was quite high across all stakeholder groups, particularly surgeons (with a 100% completion rate for Round 1 completers). I therefore decided to analyse the final round data in light of three potential scenarios:

1. A ‘last known rating’ scenario

   The last known ratings for participants who had completed only Round 1 or Rounds 1 and 2 were combined with those of the participants who had completed Round 3. This scenario assumes participants would not shift from their last known rating.

2. A ‘contrary to stakeholder group’ scenario

   For participants who did not complete Round 3, I assumed that their Round 3 rating for each domain would be opposite to the consensus status reached by that domain <u>within</u> their stakeholder group (e.g. ‘pain / discomfort’ was ‘consensus in’ by the end of the study in the patient group, so I assumed all three patients who didn’t provide ratings in the final round would score it 1-3 in importance). This scenario offers the most extreme difference in final consensus status per stakeholder group that could have been achieved had all participants completed the full study.

3. A ‘contrary to overall Delphi panel’ scenario

   For participants who did not complete Round 3, I assumed that their Round 3 rating would be contrary to the final consensus status of each domain for the overall Delphi panel (i.e. across all three stakeholder groups). This scenario is perhaps the most relevant test of the final consensus status of each domain.

Given the 100% completion rate of surgeons, consensus status remains unaltered for that stakeholder group regardless of the scenario. Detailed breakdowns of projected figures for each of the three scenarios is provided in Appendix 13.

5.4.6.1 'Last known rating' scenario

In this scenario, consensus status for the outcome domains is not altered in the surgeon and therapist stakeholder groups. For the patient group though, 'looking after dependents within household' would have reached the threshold for 'consensus in' – this domain had 69.4% rating 7-9 in importance, and in the scenario would reach 71.2% rating 7-9. However, this domain failed to reach the 'consensus in' status in the surgeon group and therefore would not have achieved 'consensus in' status overall.

5.4.6.2 'Contrary to stakeholder group' scenario

As expected, given the scenario assumed that the 'missing' ratings would be contrary to the final consensus status in each stakeholder group, there were several differences found.

For the therapist stakeholder group, five Round 1 participants did not complete the final round. The domains which had a change in consensus status within the therapist group were:

- 'Confirmation of soft tissue healing' – 'consensus in' (69% to 73% rating 7-9)
- 'Carrying or lifting heavy objects' – 'consensus in' (66% to 70% rating 7-9)
- 'Healthcare resources used' – 'consensus in' (66% to 70% rating 7-9)
- 'Loss of income' – no consensus (71% to 63% rating 7-9)

- ‘Sleeping’ – no consensus (74% to 65% rating 7-9)
- ‘Looking after dependents within household’ – no consensus. This domain was unique, in that its status in the scenario changed not because of an insufficient percentage rating 7-9 (in the scenario 70% still rate it 7-9) but because 18% would have rated it 1-3

The patient stakeholder group had three Round 1 participants who did not complete the final round, and showed a change in patient consensus status for the following outcome domains:

- ‘Looking after dependents within household’ – ‘consensus in’ (69% to 71% rating 7-9)
- ‘Healthcare resources used’ – ‘consensus in’ (69% to 70% rating 7-9)
- ‘Carrying or lifting heavy objects’ – no consensus (72% to 68% rating 7-9)
- ‘Treatment inconvenience / discomfort’ – no consensus (72% to 68% rating 7-9)
- ‘Speed of movement’ – no consensus (72% to 68% rating 7-9)

5.4.6.3 ‘Contrary to overall Delphi panel’ scenario

In this scenario I assumed the ‘missing’ ratings would be contrary to the overall consensus status of each domain. This resulted in a consensus status change within the therapist group for four domains:

- ‘Confirmation of soft tissue healing’ – ‘consensus in’ (69% to 73% rating 7-9)
- ‘Carrying or lifting heavy objects’ – ‘consensus in’ (68% to 70% rating 7-9)
- ‘Healthcare resources used’ – ‘consensus in’ (68% to 70% rating 7-9)
- ‘Sleeping’ – no longer ‘consensus in’ (74% to 65% rating 7-9)

The patient stakeholder group had just two domains where the consensus status changed:

- ‘Looking after dependents within household’ – ‘consensus in’ (69% to 71% rating 7-9)

- ‘Healthcare resources used’ – ‘consensus in’ (69% to 70% rating 7-9)

5.4.6.4 Overall impact of attrition

While the scenarios outline several domains where individual stakeholder group consensus status could have been different had all participants completed the study, the true impact would have been minimal. With the pre-defined criteria of ‘consensus in’ needing to meet the percentage thresholds in all three stakeholder groups, in the ‘last known rating’ scenario there would be no change to the overall consensus status of any outcome domain. For both the ‘contrary to stakeholder group’ and ‘contrary to overall Delphi panel’ scenarios, only one domain had a different status: ‘sleeping’ would not have been ‘consensus in’ under either of those scenarios but was deemed ‘consensus in’ in the actual Delphi result.

5.4.7 Round 1 responses of patient participants per injury group

I recognised an opportunity to explore whether any major differences existed in prioritisation of outcomes for a COS for patients with distal radius fractures compared to those with other hand fractures and joint injuries. This was limited to the first round in which any given domain was presented – beyond that point, feedback would be seen from the patient stakeholder group as a whole and could therefore distort the judgment of patient participants from each injury group when re-rating domains. I therefore analysed the ratings and consensus status for the original 37 domains after the end of Round 1 for the two injury groups, and that of the four added domains after the end of Round 2.

There was concordance in terms of rating status for 34/41 (83%) of outcome domains across the two injury groups. However, ≥70% of the distal radius fracture patient group rated the following domains as ‘very important’ while the other group did not reach this threshold:

- Hobbies and recreation
- Carrying or lifting heavy objects
- Looking after dependents within household
- Speed of movement

Conversely, the following domains were rated as ‘very important’ by ≥70% of the patient group with injuries other than distal radius fractures only:

- Change in sensation
- Cooking / preparing meals
- Keyboard and mouse / touchscreen use

Despite these differences, six of the seven reached consensus as ‘very important’ in the overall patient group by the end of Round 3. Only ‘looking after dependents within household’ not reaching this threshold by the end of the Delphi.

It is not possible to be certain how patients in the two injury groups might have altered their ratings across rounds. There were three domains where neither of the patient injury groups reached consensus of ‘very important’ on the initial scores given, but this status was reached in the final result at the end of Round 3 for patients overall: ‘sleeping’, ‘emotional / mood impact to self’ and ‘treatment inconvenience / discomfort’.

A further factor is the rating by the other stakeholder groups. Of these domains, the four reaching threshold of ‘very important’ by the distal radius fracture patient group did not reach such a status amongst surgeons and therapists. Conversely, the three domains which were ‘very important’ amongst ≥70% of patients with hand fractures or joint injuries other than distal radius fractures all also reached ‘very important’ status amongst surgeons and therapists.

5.4.8 Alternative Delphi method with ‘per round removal’ of domains reaching consensus

Removal of domains per round is employed in several COS development Delphi studies and has the benefit of reducing burden on participants as they don’t have to re-score these domains. I explored this alternative method by seeing whether eliminating outcome domains that reached ‘consensus in’ or ‘consensus out’ status after each round of the Delphi might have led to a different shortlist of ‘consensus in’ domains by the end. The final outcome domains selected would likely have been the same whether domains reaching consensus status per round were removed from further rounds or not, albeit this prediction was limited by the unknown impact to re-rating of domains when some have been removed.

5.5 Discussion

5.5.1 Key findings

The Delphi study provided data on the priorities placed by key stakeholders on the range of outcome domains synthesised from Chapters Chapter 3 and Chapter 4. Twenty domains reached the pre-determined consensus criteria to be considered ‘very important’ to measure in all future research.

Although the study design afforded a predominantly quantitative approach to analysis, there was a simple qualitative component exploring participants’ reasons for significant rating changes across rounds. For those patients who provided commentary on this, alteration in scores was based on their evolving experience of symptoms and problem. This may explain the additional ‘very important’ domains for this stakeholder group across successive rounds. Clinicians often changed their ratings to conform to those of the patient stakeholder group or the overall Delphi panel. The former reason likely drove the addition of ‘consensus in’ domains across rounds as both clinician groups began with fewer domains reaching consensus as ‘very important’ in Round 1 as compared to the patient group.

Consensus criteria for domain shortlisting required all three stakeholder groups to achieve ≥70% rating a given domain 7-9. Several domains did not reach an overall ‘consensus in’ status. Two domains reached the threshold of ≥70% rating 7-9 for therapists and patients, but not surgeons (‘hobbies and recreation’ and ‘emotional / mood impact to self’). A further five domains reached consensus threshold for ‘very important’ in the patient group alone; another two domains reaching the threshold only amongst therapists. Theoretically, if given the opportunity to explain the reason for prioritisation by certain stakeholder group(s) then the holdout group(s) might also have been convinced of their importance. This factored in the implementation of a ‘salvage’ stage in the consensus meeting, as detailed further in Chapter 6.

5.5.2 Findings in context of existing literature

Chapter 1 summarised relevant work on attempts to standardise outcome domains for distal radius fractures or hand and wrist conditions more broadly. None of these involved a Delphi process to determine outcome domain priorities of all three of the key stakeholder groups involved in the SO-HANDI COS. Each presented a non-prioritised longlist of outcome domains and then the final prioritised shortlist. The post-Delphi shortlist produced by the work described in this chapter is therefore not directly comparable with existing literature, as it is an intermediate stage to either of those types of lists. Nevertheless, the WHO ICF Core Set for hand conditions comprised 23 outcome domains (Rudolf et al. 2012) – much longer than the final shortlists of other existing literature and perhaps more reasonable to consider as analogous to this Delphi shortlist of 20 outcome domains.

There appears to be overlap between both for several outcome domains related to function (both in physiological/abstract and task-oriented/real-world examples). Pain features in both, as does emotional impact and some sense of social/relationship and work roles. Being restricted to the WHO ICF framework though, the Brief ICF Core Set could not specify domains of patient satisfaction nor consider granularity in the pain domain (such as distinguishing between pain at rest and during activity). It also does not include certain domains that the SO-HANDI participants prioritised, despite those being available in the pool of the Comprehensive ICF Core Set from which the developers were drawing the Brief ICF Core Set: examples include ‘eating’, ‘drinking’ and sleep functions’. This may be due to their health condition scope being much broader and/or lack of patient involvement in developing the set.

The patient group had more domains reaching the consensus threshold of ‘very important’ (27/41, 66%) than either of the other two groups (surgeons 20/41, 49%; therapists 24/41, 59%). While the general pattern of patients tending to prioritise more outcome domains than clinicians is evident in other recent COS development studies (Smith et al. 2019; Alkhaffaf et al. 2021), it does not hold universally (Al-Jabri et al. 2021). Perhaps by largely shaping the outcome domains based on the qualitative work with patients I produced a longlist so resonant with patients’ own experiences that they found the domains more challenging to prioritise. Data collection in the Delphi was limited to participants’ reasons for significant changes in ratings; I did not gather data on the reasoning behind their scores overall. This meant that I was not able to gain insight into why domains might have been ‘very important’ for one or two stakeholder groups but not all three. Through the subsequent consensus meeting (Chapter 6) I was able to gain some understanding of the motivations for prioritisation of these domains.

There was a general pattern of the patients changing ratings based on experience, while the clinicians did so based on conforming to the patients’ ratings. Few COS publications report participants’ reasons for ratings changes, and even fewer do so while discerning between patient and clinician stakeholder groups. Fish et al (2020) analysed three COS projects and labelled the concept of participants trying to consider the priority of an outcome from the perspective of others as “vicarious thinking”. This was more commonly cited as the reason for rating change amongst healthcare professionals than patients across all three COS studies, while the latter group more commonly cited ‘personal experience’. A similar trend of considering the scores of others was found in a post-Delphi survey for a COS on acute respiratory failure; 67% of patients and 95% of clinicians answered that they considered voting results from the other stakeholder groups when re-rating domains (Turnbull et al. 2018). This pattern of impact on re-rating appears to work in the favour of prioritising the patient voice when shortlisting outcome domains. This is desirable in a COS and would have been lost with a combined panel format of feedback.

5.5.3 Limitations

To promote inclusivity in the recruitment of clinician participants, I contacted hand surgery and therapy societies worldwide to ask for the study details to be communicated to members, inviting those interested to contact us. While this approach increased numbers and the internationality of the SO-HANDI participant pool, a consequence was that accurately establishing what proportion of those invited went on to participate was not possible.

Although clinicians from several countries completed the study, almost half the therapists and 60% of the surgeons were from the UK. All patients were from the UK, a pragmatic limitation due to ethics approval required for clinicians in other countries to assist with recruiting patients to a UK-based study. A COS attempts to be international in scope because it aims to compare and contrast all research on a given health condition (Williamson et al. 2017). The systematic review confirmed that recently published studies on hand fractures and joint injuries were performed in several countries. The potential benefits for generalisability and credibility of a COS involving relevant stakeholders from a wide range of countries have been highlighted alongside some of the challenges in including patients from multiple countries (Williamson et al. 2020). A recent survey of COS developers reported that approximately half of the COS projects for which a response was provided limited patient participation to a single country (Biggane et al. 2018).

The Delphi survey was available only in English and therefore had a language barrier to non-English readers. The challenges go beyond merely translating the words in English to a different language though. For the results to be useful, the meaning of the outcome domains, descriptors and examples would need to be preserved and validated across language and cultural divides. Resource limitations made such translation of the questionnaire into multiple languages untenable.

On a rudimentary level, I did not find any major differences in the initial ratings of patients with distal radius fractures as opposed to those with other injuries in the scope of 'hand fractures and joint injuries'. Of course, the study design was not developed with this comparison in mind and therefore I cannot be certain of quite how the ratings of each injury group would have altered across the rounds. In hindsight, a proper exploration of this might have been achieved by running the Delphi with four stakeholder groups, having two of these as the patient participants divided by injury group. Nevertheless, with consensus requiring a high percentage of therapists and surgeons to also rate a given outcome domain as 7-9, I expect that the overall result would not have significantly differed to what was found using the single patient group.

Finally, while I gained a semblance of the reasons for why participants who altered their scores in a significant way did so, unfortunately many of these participants chose not to supply a comment. Therefore the picture is incomplete, which must be borne in mind when drawing conclusions. This is related to another aspect of the Delphi methodology I used, wherein feedback was limited to the distribution of scores per stakeholder group – while this offered participants some information on which to base their re-rating of domains, it did not convey the reasons for changes in scores and could therefore be considered a “normative rather than informational influence” (Rowe et al. 1991; Murphy et al. 1998).

5.5.4 Strengths

A key strength of this study was the preceding work which informed development of the outcome domains used in the Delphi. The extensive systematic review and detailed exploration of the patient perspective through interviews and focus groups enabled development of a comprehensive set of domains. I believe this is reflected in there being few added domains based on participant suggestions after Round 1. Of the four added, three were actually more specific facets of existing domains.

Another key strength was the completion rate, with over 94% of participants who completed the first round going on to complete the whole study. It is recommended that attrition be minimised so as to avoid overestimating consensus in the process (Sinha et al. 2011). An analysis of COS Delphi surveys has emphasised the need to carefully consider the number of survey items (outcome domains) and overall panel size, as these appear to influence attrition (Gargon et al. 2019a). Although I had a fairly large panel size compared to most in that study (where ~83% had ≤100 participants) the SO-HANDI Delphi began with only 37 outcome domains and added merely four.

By having a relatively low number of outcome domains, I was able to retain all outcomes throughout the three rounds. This theoretically allowed a domain to shift in consensus status across rounds, which would not be possible domains that reached consensus status in

Rounds 1 or 2 were removed at those points. Despite this theoretical advantage to retaining all outcome domains until the end, in this Delphi I found that a ‘per round’ or ‘final round’ deciding stage would have led to no difference in consensus status.

Although the range of countries represented by the various participants could be improved, conducting an online Delphi improved the inclusivity for geographically distant participants. There is a subset of the population that do not find an online survey to be an accessible method, and they were accommodated by the option of postal questionnaires.

On considering the participant demographics it was clear that there were a reasonably broad range of perspectives even within stakeholder groups. Patients with various injury types took part, along with a mix of age, sex and treatment (surgical vs conservative). The surgeons were of a predominantly orthopaedic background and most were male, while the therapists tended to have an occupational therapy background and most were female. Importantly, the majority of clinician participants had ≥15 years of managing patient with hand fractures and joint injuries and many had been involved in conducting research projects themselves bringing both a clinician and researcher perspective to the process.

5.6 Conclusion

This step achieved initial consensus prioritisation of the longlist of outcome domains synthesised from the findings detailed in Chapters Chapter 3 and Chapter 4. Twenty outcome domains reached the consensus threshold of 'very important' to measure in future clinical research on hand fractures and joint injuries.

The next step was to conduct a consensus meeting to decide a final COS based on consensus of the core stakeholder groups of patients, hand surgeons and hand therapists, as well as having some input from other stakeholder perspectives. The output from this Delphi survey served to shape the consensus meeting and provide background information and data for the consideration of the meeting participants as detailed in Chapter 6.

Chapter 6 Consensus meeting

6.1 Introduction

The Delphi survey was an important step towards reaching a consensus on the prioritisation of outcome domains to be included in the COS by the key stakeholder groups of patients, surgeons and therapists. This chapter describes an online international consensus meeting to select the final COS domains.

The COMET Initiative recommends that representatives of key stakeholder groups have an opportunity to discuss the results of a Delphi survey, and have additional voting to select a final COS as needed (Williamson et al. 2017). A consensus meeting enables detailed discussion of domains prioritised through the Delphi process: while the latter revealed the distribution of ratings for each domain across each stakeholder group to survey participants, the consensus meeting participants would be able to explore the reasons behind prioritisation of domains.

There are many possible formats for consensus meetings (Murphy et al. 1998; Humphrey-Murto et al. 2017; Williamson et al. 2017), and in this chapter I describe the key choices made as well as some of the alternative options considered. I then present the main discussion points and voting results of the meeting.

6.2 Aims

The purpose of this study was to identify consensus of opinion on the most important outcomes to include in a COS for hand fractures and joint injuries in adults. This was to be based mainly on the views of patients, hand therapists and hand surgeons, but also considering input from representatives of other relevant stakeholder groups such as those involved in clinical research but not the direct clinical management of patients.

Primary objective: To select a final COS of domains for hand fractures and joint injuries in adults.

Secondary objective: To develop further insight about the prioritisation of outcome domains by reviewing group discussion.

6.3 Methods

6.3.1 Background

While the Delphi survey had a core principle of anonymity between participants which is central to the methodology, consensus meetings typically involve direct interaction and the identity of participants is not hidden. The format, setting and organisation of such meetings affect how the participants interact and can influence the judgments they reach (Murphy et al. 1998).

The nominal group technique (NGT) was developed by Van de Ven and Delbecq in the 1960s and presented as an instrument for exploratory health studies in 1972. It is a structured meeting involving a sample of people with “experience, expertise or perceptions” directly related to the area being explored (Van de Ven and Delbecq 1972). In the original description, groups of five to eight were deemed an acceptable size and took part in the following stages:

- Introduction – background information and the key question/aim of the meeting are conveyed to participants
- Silent generation of ideas – individual participants would spend time working independently to develop a list relevant to the question that the meeting is attempting to address
- Round-robin feedback – each participant in turn is asked to share one item from their list which is then recorded by a facilitator, proceeding until all items from the idea generation stage have been exhausted
- Discussion of items – open discussion to clarify, dispute and defend items, or add new ones which emerge through discussion
- Prioritising items – the group is asked to rank the priority of items (originally specified as the top ten most critical items), with each participant adding their selection to a tally
- Further discussion of initial tally – open discussion of the initial vote/tally results
- Re-ranking and rating of priorities – each participant is permitted to review and change their top ten priority items as they wish, and then asked to rate the relative importance of items within their top 10

This process involves key stakeholders generating items related to the underlying problem being considered and proceeds with individuals making judgments, discussing their reasoning and then reviewing their initial judgments over the course of the meeting.

The SO-HANDI consensus meeting had some key differences borne out of being part of a wider COS development process rather than a standalone NGT meeting. For example, the ‘generation of ideas’ (outcome domains in this case) by key stakeholders had already been accomplished through the various steps described in previous chapters. I therefore considered the approach used by the James Lind Alliance (JLA) for their Priority Setting Partnerships (PSPs). The JLA PSP Guidebook details the workshop methodology used to select the top ten research priorities for a given health condition and their starting point is also a list of research questions generated by prior steps (James Lind Alliance 2021). Their workshops involve small and whole group steps with the aim of encouraging participation by all, minimising the discussion being dominated by a single person and taking everyone's opinions into consideration. There are five phases:

1. Small group discussion – each participant in turn contributes their views of the top and bottom research questions
2. Small group ranking – discussion and then ranking of priority of research questions by the groups
3. Whole group review – the various small groups' rankings are aggregated and the aggregated list is discussed by the whole group
4. 2nd round of small group ranking – the aggregated list is discussed further and re-ranked within the small groups
5. Final whole group review – a final aggregated ranking is generated based on the output of the small groups. The final list is then discussed by the whole group, with an aim of reaching a consensus on the top ten research questions

The JLA PSP approach is effectively an extension of the original NGT approach, with both sharing the key concepts of discussion and iterative scoring/ranking. I adopted a broadly similar approach to that of the JLA PSP workshop for the COS consensus meeting, with the same principle of inclusivity prompting the use of a combination of small and whole group discussion and ranking/voting. Due to the ongoing COVID-19 pandemic, congregating for a face-to-face meeting was not possible and therefore the online medium was selected.

6.3.2 Participants

6.3.2.1 Stakeholder groups

The key stakeholder groups of patients, surgeons and hand therapists were involved in the Delphi survey and were again the priority when deciding on recruitment for the consensus meeting. However, I believed that some representation from additional stakeholder groups involved in the design of clinical research trials would be of benefit. There are several potential trial team members including medical statistician, data manager, trial manager and health economist who might have interest or insight into outcome selection and measurement, and COSs (Farrell et al. 2010). However, with the meeting focusing on ‘what to measure’ I decided that insight from two of the roles would be particularly relevant – a trial manager and a health economist. In my view, others such as medical statistician are likely to be of benefit at the ‘how to measure’ stage of COS development.

6.3.2.2 Sample size

There is no set rule on the total number of participants in a consensus meeting for COS development. A recent update to a systematic review of COSs for research identified 25 of 33 studies which provided details on participants numbers – of these, 15 had some form of consensus meeting as part of the process and the participant numbers ranged from 12 to 55, with a median of 19 (Gargon et al. 2021). I decided that the patient group should have an approximately equal number to the non-patient stakeholders in order to maintain my commitment to the role of the patient voice in determining the final COS. Although patients formed one of three key stakeholder groups, I wished to avoid having too low a patient to professional ratio which would effectively weaken patients' voting power. I also aimed for a patient participant mix covering a range of injuries within the COS scope as well as various ages. To cover this mix, I planned to recruit approximately 10-14 patients, 5-7 surgeons and 5-7 hand therapists, with one each of the trial manager and health economist groups. The total number would therefore be 22-30 participants. Given the planned approach described earlier, the meeting would have an element of small group work; a general suggestion is that small group size should range from 5 to 12 (Humphrey-Murto et al. 2017) which would correspond to three small groups.

6.3.2.3 Participant eligibility and recruitment

Invitations were sent to selected participants from the pool that completed the Delphi survey. In addition, I invited some patients who participated in the earlier interviews and focus groups and gave consent to be contacted about further research. Several of the surgeons and therapists were invited on the basis of being known to members of the research Steering Group as having significant clinical and/or research experience.

The eligibility criteria were as set for the qualitative study for patients (see Section 4.3.3 for details), and the Delphi study for surgeons and therapists (see Section 5.3.2.3 for details). For the health economist and trial manager being invited to the meeting, the requirement was prior involvement in orthopaedic trauma clinical trials.

Any member of the research group involved with facilitating the meeting would not be eligible to be a participant, nor to vote.

Given that this was an online meeting, international participation was welcome. However, unlike the Delphi where asynchronous engagement was accommodated by default, this live meeting had to work within the limits of time zones. The meeting was scheduled to enable participation from North America to Europe feasible. For pragmatic reasons, participants were required to be fluent in English.

6.3.3 Consensus meeting development

The structure of the consensus meeting underwent a development process beginning with forming an initial outline of the essential steps to define a COS of domains and approaching a suitable Chair for the meeting. I further refined the format in discussion with the appointed Chair and the research Steering Group, and then appointed facilitators who were briefed on the meeting objective and structure.

6.3.3.1 Appointing a Chairperson and facilitators

I aimed to appoint a Chair who had prior experience of chairing consensus meetings and had no pre-existing link to research in the field of hand fractures and joint injuries. Professor Jamie Kirkham was the invited Chair, fulfilling these criteria and having significant experience in COS development methodology through the COMET Initiative. His background in the latter included creating guidelines for developing a COS (Kirkham et al. 2017), writing a protocol (Kirkham et al. 2019) and reporting COSs (Kirkham et al. 2016). There were five meetings in the three months preceding the meeting between myself and the Chair and members of the research group across which I:

- Conveyed background information about the steps undertaken to date
- Ensured clarity about the meeting objective
- Developed and refined the meeting format

Facilitators were primarily selected from the study research group. All were briefed on the specific meeting objectives and format and provided with the pre-meeting information pack that had been sent to participants. Multiple facilitators were required as I felt that the online format would mean that in each small group one facilitator would take on the key 'facilitation' role while the other could contribute as needed but also manage the 'discussion board/shared screen' and monitor to ensure all group members have opportunity to participate and any 'raised hands' are noticed. Although a single person can often do all these tasks in a face-to-face setting, I believed that this would be much more challenging in an online meeting.

6.3.3.2 Defining consensus criteria

Several options for consensus criteria were described in Chapter 5, with the Delphi requiring clear criteria to establish whether a given domain had reached consensus for prioritisation. For the consensus meeting, I anticipated a binary voting process at key stages and set a threshold of ≥80% 'yes'. This was a reasonably high level of agreement, supporting the principle that only the outcome domains deemed critically important by key stakeholders should be included in a COS. It also helped to maintain the strength of the patient vote to some degree – if ~50% of patients did not feel a domain was a priority for inclusion in the COS, then they would not be overruled by all the other stakeholder groups.

Small group discussions aimed for agreement across the whole small group where possible, but settled upon a within-group majority vote as the minimum threshold for consensus. Although a lower threshold than for the whole group voting, the small group work served to elicit discussion points for whole group discussion and voting.

6.3.3.3 Pre-meeting information pack development and preparation

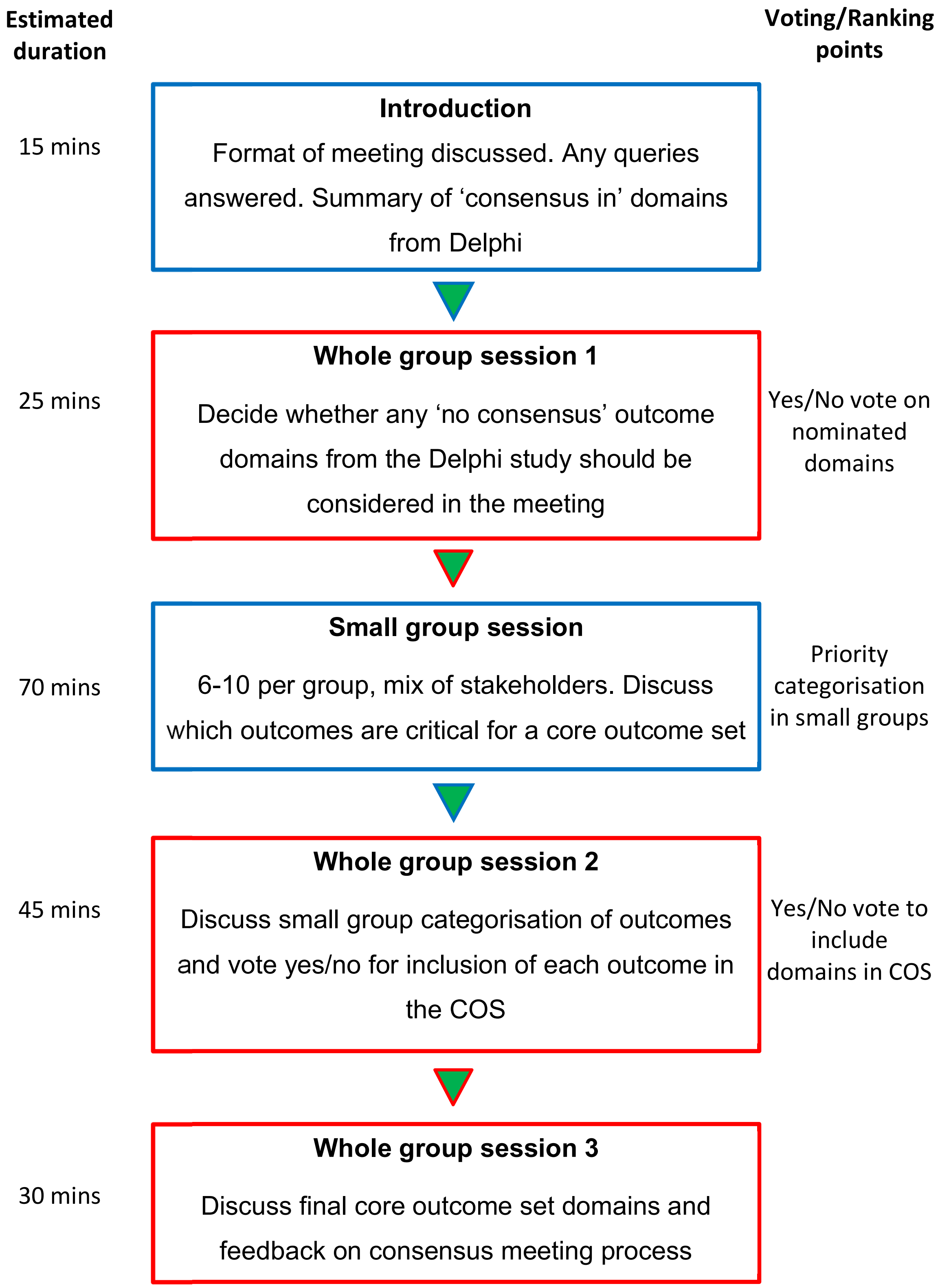
I produced an information pack summarising background information on COSs, the previous steps taken in developing this specific COS and then outlining the meeting agenda, as well as more detail on the areas of discussion planned for the meeting. Participants were presented with summary data from the Delphi and were essentially led through the key steps of the meeting itself, with encouragement to consider whether they would salvage any outcome domains, how they would categorise the domains in terms of importance to include in the COS, and how they might group the outcome domains. The information pack main text and an example page showing summary Delphi data for some domains is provided in Appendix 14. All participants were encouraged to raise any uncertainties with the research team. The pack was developed by me and then feedback sought from the Chair and research group. I then showed the final draft to some patient representatives for feedback to help ensure that the language used was clear and appropriate for laypeople. I emailed the pack to participants two weeks prior to the meeting, with a follow-up reminder a few days before the meeting strongly encouraging that participants read the pack carefully in order to help the meeting run efficiently and ease full involvement in the meeting processes.

As a preparatory step for the online meeting, all participants were offered a chance to test the online meeting system by joining a trial meeting with one of the researchers at a time convenient to them. At the trial meeting, participants could seek clarification on the use of Microsoft Teams or the polling platform that would be used in the actual consensus meeting.

6.3.4 Study regimen

The overall plan for the consensus meeting is shown in Figure 6.1.

Figure 6.1 Flowchart of final consensus meeting regimen



Estimated duration

Voting/Ranking points

15 mins

Introduction

Format of meeting discussed. Any queries answered. Summary of 'consensus in' domains from Delphi

25 mins

Whole group session 1

Decide whether any 'no consensus' outcome domains from the Delphi study should be considered in the meeting

Yes/No vote on nominated domains

70 mins

Small group session

6-10 per group, mix of stakeholders. Discuss which outcomes are critical for a core outcome set

Priority categorisation in small groups

45 mins

Whole group session 2

Discuss small group categorisation of outcomes and vote yes/no for inclusion of each outcome in the COS

Yes/No vote to include domains in COS

30 mins

Whole group session 3

Discuss final core outcome set domains and feedback on consensus meeting process

6.3.4.1 Salvage of outcome domains reaching no consensus in Delphi survey

After an introduction to the meeting objectives and agenda, as well as a reminder about core outcomes sets and the scope of injuries covered in this COS, participants were given an opportunity to test use of the online polling platform.

The initial step after introduction was a whole group one, in which participants were presented the various outcome domains which had reached no consensus through the Delphi survey. Further consideration of such domains is fairly standard in COS development studies (Smith et al. 2019; Al-Jabri et al. 2021; Alkhaffaf et al. 2021). Feedback in the Delphi was in the form of stakeholder group distribution of scores only, with no qualitative feedback of participants' reasoning for ratings. Therefore, allowing some discussion time for domains that reached no consensus meant that any of these domains felt to be particularly important could be advocated for. This "informational influence" could help to counter the potential concern of a "normative influence" at the Delphi survey stage (Rowe et al. 1991; Murphy et al. 1998). Participants were asked to comment on any domains that they wished to nominate to bring to a group vote to preserve for consideration in the rest of the meeting and had to provide their justification for the nomination. Open discussion was encouraged to assist participants in making their voting decision. The vote was conducted via Poll everywhere, with each domain presented in turn and participants completing a 'yes/no' vote on whether the domain should be kept for the rest of the meeting. The requirement for the vote of ≥80% 'yes' being needed felt justified given that including a domain at this stage would be going against the result of the Delphi which benefitted from a larger participant cohort.

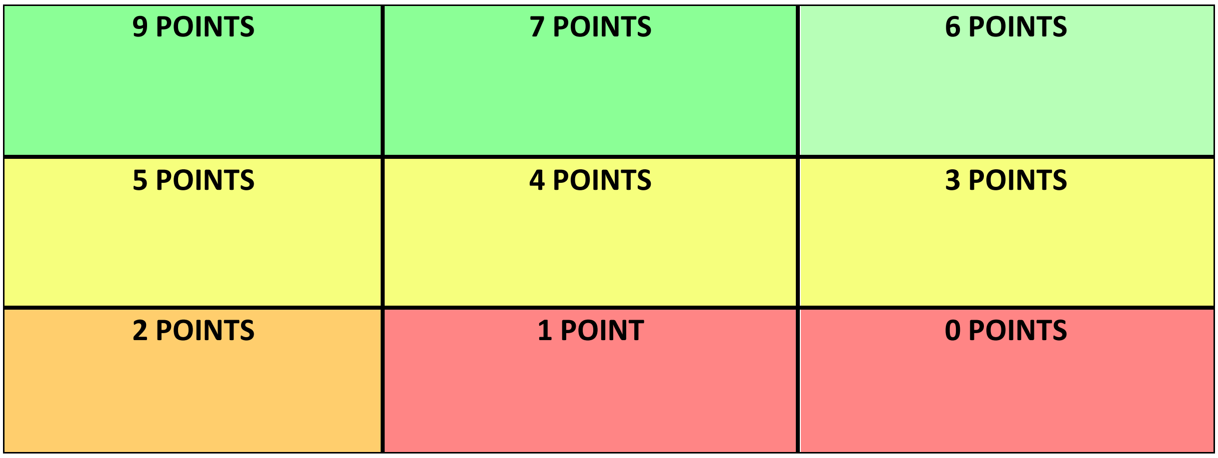
Domains were presented in the following order: initially those which had no stakeholder group with ≥70% rating 7-9 in the Delphi; then those with one stakeholder group reaching that level of prioritisation; and finally those with two stakeholder groups deeming the domain 'very important'.

6.3.4.2 Prioritisation of outcome domains for inclusion in COS

Participants were allocated to one of three small groups (specific allocations determined prior to the meeting, with stratified representation of each of the three key stakeholder groups). Each small group had two facilitators to guide the participants through the task of categorising the outcome domains deemed 'very important' through the Delphi (as well as any salvaged in the prior stage of the meeting) into one of three groups: 'critical for COS', 'important but not critical' and 'definitely not needed in COS'. Participants discussed the domains within their small groups and had to reach majority agreement when assigning all the domains into these categories.

Each of the categories was assigned differing points to support the ensuing aggregation step: the 'critical' category was assigned 3 points, the intermediate priority category 1 point, and the 'definitely not needed' category 0 points. The purpose of this scoring structure was to highlight the importance of a domain being rated as critical: even one small group categorising a domain as 'critical' would result in it outweighing a domain rated as 'important but not critical' by two small groups. I also felt it might be beneficial to have a more granular aggregate score than would be achieved by a 0-1-2 scoring system. The three small groups' categorisations were aggregated and a figure created to present to the whole group to show points scored by each domain (template shown in Figure 6.2).

Figure 6.2 Template for aggregate scoring of small group categorisations



9 POINTS	**7 POINTS**	**6 POINTS**
5 POINTS	**4 POINTS**	**3 POINTS**
2 POINTS	**1 POINT**	**0 POINTS**

Domains with an aggregate of 0 or 1 point only were eliminated from further consideration, but all the rest underwent whole group discussion, facilitated by the Chair. Once participants exhausted discussion points, the meeting proceeded to an individual ‘yes/no’ vote on whether each outcome domain should be included in the COS. Participants were provided a link to the Poll Everywhere survey for this voting process, and results were only revealed once all domains were voted on.

6.3.4.3 Post-meeting participant feedback

Despite various research group members and the Chair having experience of conducting face-to-face consensus meetings, this was the first one being run purely online. I therefore decided it would be useful to collect some participant feedback on the methodology being used as well as on the final COS produced.

I produced an online survey prior to the meeting using the JISC Online Surveys platform, and a link was provided to participants at the end of the meeting. The survey was anonymised in order to encourage honest comments and criticism, with the only identifier collected being whether the respondent was a ‘patient’ or ‘non-patient’.

6.3.5 Data collection and analysis

Participant demographic details were already available for the majority who had completed the Delphi survey. For those who did not, demographic details as specified in Section 5.3.2.1 were collected at time of online registration and consent, using JISC Online Surveys.

Voting results were extracted from Poll Everywhere and analysed contemporaneously to inform subsequent stages of the meeting as necessary. The data were retained for more detailed analysis (e.g. per stakeholder group) after the meeting.

Audio recordings were extracted from the Microsoft Teams recording files for the main meeting rooms and all breakout rooms. Transcription was done by a transcription service and reviewed together with the audio files for accuracy by the research team. NVivo v12 software was utilised in analysis of transcripts. These were coded to summarise the main areas of discussion by participants in the various steps of the meeting, along with identifying the stakeholder group from which comments arose. The overall methodological approach was similar to the inductive, thematic approach used for analysis of focus groups in the study described in Chapter 4.

6.3.6 Ethics approval

The consensus meeting protocol was added as a substantial amendment to the existing Delphi study protocol, which was accordingly reviewed and approved by London (Harrow) Research Ethics Committee (20/PR/0178).

6.4 Results

6.4.1 Recruitment

A total of 27 participants were recruited to the consensus meeting: 12 patients, seven surgeons, six hand therapists, one health economist and one trial manager. All patients were from the UK, while the non-patients were from five countries (UK for most, and one from each of South Africa, USA, Canada and Sweden). The majority of participants (23/27, 85%) had participated in, and completed all three rounds, of the Delphi survey.

Demographic characteristics of the participants from the three core stakeholder groups are provided in Table 6-1.

Table 6-1 Demographic characteristics of meeting participants from the three core stakeholder groups

Characteristic		Patients (n=12)	Therapists (n=6)	Surgeons (n=7)
Median age (interquartile range)		54 (32-64)	47 (43-50)	55 (52-58)
Sex	Male	4	1	5
	Female	8	5	2
Injury group	DRF	4		
	Non-DRF	8		
Injury management	Surgical	6		
	Conservative	6		
Patient occupation type	None/retired	3		
	Office/non-manual	6		
	Light manual	3		
Clinician experience (years)	10-15		1	0
	15-25		3	3
	25+		2	4
Surgeon training background	Orthopaedic			5
	Plastic			2
Therapist training background	Occupational therapy		4	
	Physiotherapy		5	
Therapist accredited / certified	Yes		3	
	No		2	
	N/A		1	

This broadly reflected a stratified sampling of the larger Delphi survey participant pool. I purposely recruited fewer patient participants with distal radius fractures than the other hand fractures and joint injuries given the range of injuries to be covered within the latter injury group. I also deliberately selected hand therapists and surgeons with greater experience in management of these injuries: clinician participants had between 13 and 36 years of experience.

Demographic data for the single trial manager and health economist are not presented as they were not part of a larger group of Delphi participants and therefore the details are unlikely to be of relevance here. However, both were experienced and established in their fields, and had participated in the design and running of orthopaedic trials in the past.

6.4.2 Salvage of domains failing to reach consensus in Delphi survey

6.4.2.1 Nomination and voting results of domains to potentially salvage

From the whole group discussion, a total of 10 outcome domains that had not reached consensus as 'very important' through the Delphi were nominated to be 'salvaged'. These are presented along with a breakdown of the voting result in Table 6-2. None of the 10 nominated domains reached the required threshold to take forward into the rest of the meeting.

Table 6-2 Voting results for domains reaching no consensus in Delphi survey that were nominated for potential consideration for COS

Outcome domain	**Patients (n=12)**		**Surgeons (n=7)**		**Therapists (n=6)**		**Trial manager / Health economist (n=2)**		**Total (n=27)**	
	Yes	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
Hobbies and recreation	9 (75%)	3 (25%)	4 (57%)	3 (43%)	4 (67%)	2 (33%)	1 (50%)	1 (50%)	18 (67%)	9 (33%)
Emotional / mood impact to self	11 (92%)	1 (8%)	2 (29%)	5 (71%)	5 (83%)	1 (17%)	0 (0%)	2 (100%)	18 (67%)	9 (33%)
Loss of income	4 (33%)	8 (66%)	2 (29%)	5 (71%)	1 (17%)	5 (83%)	0 (0%)	2 (100%)	7 (26%)	20 (74%)
Carrying or lifting heavy objects	5 (52%)	7 (42%)	1 (14%)	6 (86%)	2 (33%)	4 (67%)	2 (100%)	0 (0%)	10 (37%)	17 (63%)
Confirmation of bone healing and alignment (fractures only)	10 (83%)	2 (17%)	1 (14%)	6 (86%)	2 (33%)	4 (67%)	2 (100%)	0 (0%)	15 (56%)	12 (44%)
Confirmation of soft tissue healing	7 (58%)	5 (42%)	0 (0%)	7 (100%)	0 (0%)	6 (100%)	1 (50%)	1 (50%)	8 (30%)	19 (70%)
Healthcare resources used	1 (8%)	11 (92%)	2 (29%)	5 (71%)	0 (0%)	6 (100%)	0 (0%)	2 (100%)	3 (11%)	24 (89%)
Using hands to change body position	2 (17%)	10 (83%)	0 (0%)	7 (100%)	2 (33%)	4 (67%)	0 (0%)	2 (100%)	4 (15%)	23 (85%)
Shopping for groceries	3 (25%)	9 (75%)	0 (0%)	7 (100%)	1 (17%)	5 (83%)	1 (50%)	1 (50%)	5 (19%)	22 (81%)
Appearance	7 (58%)	5 (42%)	5 (71%)	2 (29%)	4 (67%)	2 (33%)	0 (0%)	2 (100%)	16 (59%)	11 (41%)

6.4.2.2 *Reasons for nomination of domains and surrounding discussion points*

Each nominated domain required an accompanying reason from the person proposing that the domain be put to a vote for salvage at this stage. All participants were then allowed to contribute their thoughts on the suggestion, to help inform the eventual vote. Since none were voted through to retain for further consideration in the meeting, instead of a detailed summary of the discussion points I have opted to present the majority with merely indicative quotes to represent the reason for nomination or participants' reasoning for either prioritising the domain or not regarding it as important (Table 6-3). The stakeholder group origin of the quote is stated in each case.

Table 6-3 Table of indicative quotes for nomination/discussion of outcome domains for the early meeting salvage vote

Outcome domain	**Indicative quotes**
Shopping for groceries	*"Can I just say that shopping for groceries to me is essential. If you're a single person and you've got a broken wrist, it's pretty bloody hard to shop for groceries and that's fairly important. That's your food"* (Patient)
Carrying and lifting heavy objects	*"Lots of things are heavy that you do need to carry, i.e. a shopping bag or a pile of books or, you know, just even making the bed when you've got to move the mattress or something"* (Patient)
Using hands to change body position	*"I'd like to mention using hands to change body position. I think this can depend on the individual of course but quite often people do use their hands and arms quite a lot if their legs are not as active"* (Patient)

Table 6-3 Table of indicative quotes for nomination/discussion of outcome domains for the early meeting salvage vote (cont'd)

Outcome domain	**Indicative quotes**
Confirmation of soft tissue healing	*"I think soft tissue can be really important and we can get an awful lot of pain and stuff from that. So to me it's quite important"* (Patient)
Looking after dependents within household	*"Looking after dependents is quite specific to people that have dependents that they need to look after and perhaps the activities required in that are also captured in some of the other functional measures"* (Therapist)
Confirmation of bone healing	*"If I have confirmation I can see an x-ray and I can have a chat about the healing of the bone, then it enables me to move on with my life and not having this hanging over me, and also have…increased confidence in the use of the joint, knowing that it's fully healed"* (Patient 1) *"In hand fractures I teach our residents…that x-ray healing of hand fractures is useless; that clinical healing is much more important because frequently radiologists tell us 'oh, this metacarpal's not healed but it's maluniting' and we go in…but I need dynamite to get the pieces of bone apart. So clinical healing is way more important"* (Surgeon) *"We don't need x-rays but it is a communication thing and not all surgeons and doctors actually know that people are waiting to hear that from them"* (Patient 2)

Table 6-3 Table of indicative quotes for nomination/discussion of outcome domains for the early meeting salvage vote (cont'd)

Outcome domain	**Indicative quotes**
Appearance	*"I think appearance should be in the core set. If you look at the outcomes that we collect…appearance has got a massive impact on satisfaction and normal hand scores"* (Surgeon 1) *"The vast majority of people with hand fractures have appearances at the end of the treatment that are relatively similar to the beginning. So it may not be quite as emotive in this group of patients as in elective patients"* (Surgeon 2)
Healthcare resources used	*"I think probably that's going to get measured in any full trial but it will get measured in what we would say…the cost side of the equation, and as a health economist I tend to err people away from putting that into an outcome set because it will very much be collected and it is really important, but usually addressed in a separate form"* (Health economist) *"Health economics is important for many trials but not all studies have a health economics component"* (Surgeon)
Loss of income	*"I think personally 'loss of income' links very closely to work…being able to 'return to usual work' has made it through and perhaps ['loss of income' is] captured almost within the work outcome that we've got"* (Therapist) *"I think 'return to work' is a much better capture of that measure than income, which just is really not a good thing to put in a survey. It annoys people, it annoys patients…the exact word it's considered too "intrusive"…a bad thing on the whole"* (Health economist)

‘Emotional / mood impact to self’ and ‘hobbies and recreation’

Two nominated domains are not referred to in Table 6-3 and instead are elaborated on in more detail as they were eventually raised again in the final discussion towards the end of the meeting after the final vote on the domains to include in the COS.

‘Emotional / mood impact to self’ was nominated by a therapist: *“’Emotional mood to oneself’ like pain and anxiety and fear of movement and some things like that, I think that’s important…because it’s a big impact on the patients…I think it’s very important to not just look at the physical parts, also the psychological”* (Therapist). Some patients agreed with this in the meeting chat text, with one patient stating that the underlying reason for its importance in their opinion was the same as that of the domain of ‘hobbies and recreation’:

> *“The actual underlying reason is the same because both of these are very much about self-esteem and…part of our identity and who we are and what makes life liveable is to do with our hobbies and our recreation and of course if the injury or a healing outcome from the injury impacts severely negatively on the hobbies and recreational side of life then it’s also going to affect…’emotional mood impact’. It’s a self-esteem thing…they kind of are a bit intertwined there.”* (Patient)

The importance placed on these domains by patients (particularly ‘emotional / mood impact to self’ with 11 of the 12 patient panellists voting ‘yes’ for the domain to be taken forward to the rest of the meeting processes) and therapists was clear in the vote result that followed this discussion. Despite the patient group having the highest weighting amongst stakeholder groups at the meeting and having broad support from therapists, both domains had only a 67% ‘yes’ vote and failed to meet the 80% threshold across the overall consensus meeting panel (Table 6-2).

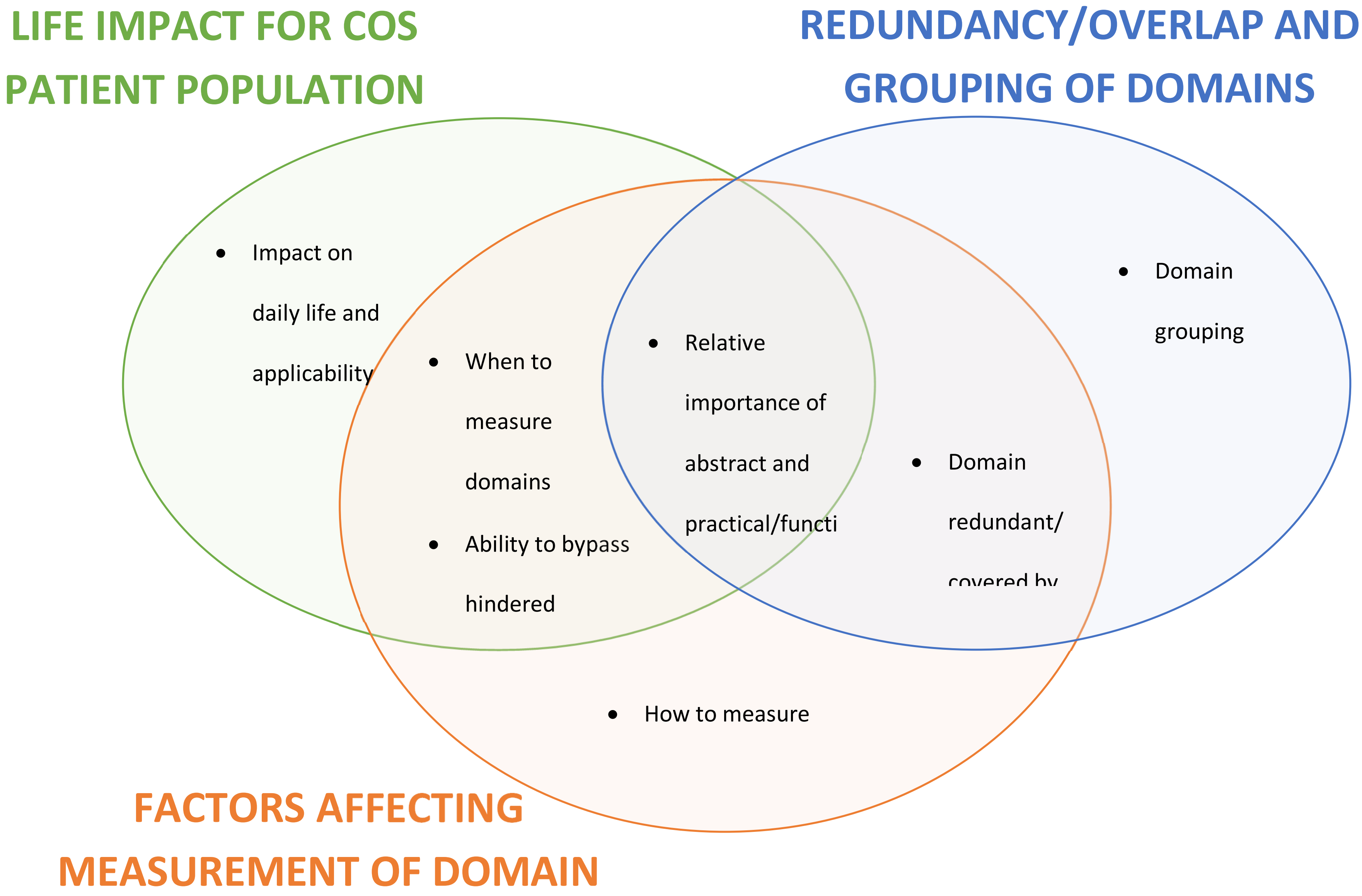
6.4.3 Prioritisation of outcome domains for inclusion in COS: small group discussions

Since no domains were salvaged, the three small groups proceeded with prioritisation of only the 20 outcome domains that had reached consensus of ‘very important’ in the Delphi survey. In this section I present a summary of the key themes arising from my analysis of the small group discussions, which provide some insight into the decision-making process that culminated in the final aggregated score:

- Impact on daily life and applicability to broad set of patients
- How to measure domains
- When to measure domains
- Domain grouping with one or more others
- Domain redundant/covered by information from other domains
- Domain influenced by pre-injury factors
- Ability to bypass hindered function
- Relative importance of abstract and practical/functional outcome domains

As part of the post-meeting analysis, I organised the themes into three broad areas as displayed in Figure 6.3. Some themes were associated with more than one broad area. The prioritisation of domains clearly involved consideration of a given domain’s impact to patients with hand fractures and joint injuries. There were subtleties to this though, with some discussion that certain impacts (e.g. on function) could be bypassed – patients could find ways around limited use of the functionally impaired hand. There were logistical considerations, with participants discussing whether some domains were essentially covered by another. In essence, very specific domains (e.g. ‘keyboard and mouse / touchscreen use’) may have lower priority as they could conceivably be covered within a broader one (e.g. ‘return to usual work / job’). Finally, some considered the practicality of measuring domains, and during the meeting the Chair did need to reinforce that such considerations should not weigh into the prioritisation of domains at this stage of COS development.

Figure 6.3 Areas and themes generated from analysis of small group discussions



In the rest of this section, I present each theme in greater detail, along with some accompanying quotes from participants that contributed to development of the theme. Along with stakeholder group, the particular small group from which the quotes arose is indicated by ‘SGx’ (where x is replaced with the corresponding small group number).

6.4.3.1 Impact on daily life and applicability to broad set of patients

Much of the discussion on selection of prioritisation category revolved around perceived impact on daily life: *“OK put in ‘important’ the things that everybody does in their day-to - day life” and then things that are more specific just put in ‘not needed’”* (Patient, SG1). However, participants also considered how broadly the domains applied across the COS target patient population.

Impact on daily life could be in the form of basic essential functions:

> *“I put ‘self-hygiene and personal’ care in the critical column [because] everyone does that”* (Surgeon, SG1)
>
> *“I think for just a basic human need, ‘eating and drinking’ is a critical point, you know, being able to do that yourself”* (Patient, SG2)

In other cases, participants considered the impact to everyday tasks, both work- and non-work-related:

> *“As life is now everything you do is online – at home everything, you know, all your bills, all your banking, all your ordering, so for not necessarily work, you need your ‘keyboard and mouse’ to live your life”* (Patient, SG1)
>
> *“I think [‘return to usual work / job’] is critical…if you're the main breadwinner…it’s not just a self-esteem thing, it’s actually my ability to do my usual job…this is life*

> *changing stuff – and I think it's a really critical outcome that 'return to work' is something that's measured"* (Patient, SG3)

Some participants felt that as important as some domains might be in cases where they do impact on patients, they were felt to not have broad enough applicability to deem for inclusion in the COS:

> *"'Change in sensation' – I have a hard time understanding…if you don't have numbness this doesn't count at all – but if you have numbness this is totally critical…Not that many people end up with numbness after a fracture"* (Surgeon, SG3)
>
> *"I think the 'keyboard and the mouse and touch screen use' [is definitely not needed]…just because it just doesn't apply to everybody"* (Therapist, SG1)

Several clinicians argued against a strict consideration of the applicability to a broad set of patients for a few domains, on the basis that something which might only be an issue in a small proportion of patients could nevertheless have such a big impact that it warranted a higher priority:

> *"'Change of sensation' is not a common problem with…hand and wrist fractures, but when there is a sensation loss it can cause…major long-term problems if the person has permanent loss, so I would contest that it is possibly more in the 'critical' side"* (Therapist, SG1)
>
> *"If you then want to for example compare treatments or compare different interventions, then it becomes quite important to know the complications…you can get to the same outcome but you've been through hell to get there sometimes"* (Surgeon, SG3)

The trial manager representative raised the point that complications which might matter to researchers may not necessarily translate into a significant impact and something which patients believed to be of critical importance: *"From a research perspective, treatment complications like "are you getting infections"…or "have you got implant problems from your surgery", etc. Normally we would think that was pretty critical and something to record. But it's not obvious how that would translate to what it would mean to somebody"* (Trial manager, SG2).

6.4.3.2 How to measure domains

Participants were informed that the meeting focus was on 'what to measure' rather than 'how to measure'. Nevertheless, there were a few times when they expressed concern about the difficulty in measuring a particular domain as a motivating factor when prioritising the domain.

One example was the domain 'stability of joints':

> *"Part of the problem with that…is that mostly we can't measure it…we're not sure quite what we mean by stability a lot of the time"* (Surgeon, SG1)

There was also debate on having a COS for which measures required patient attendance or review in-person. For example, several domains that traditionally involve assessment by clinicians (including 'change in sensation', 'range of movement' and 'strength or muscle power'):

> ***"Surgeon 1:*** *Anything that you have to see a patient in-person for is very difficult to collect within a core outcome set. So that's just worth bearing in mind when thinking of the critical things about what can actually be done in every single study because increasingly…follow-up is remote...*

> ***Surgeon 2:*** *Yes, except…as technology changes we can now collect range of motion through a phone, I suspect we'll collect more of this stuff, not least grip strength in future. I don't think we should restrain ourselves just based on current technology.*
>
> ***Surgeon 1:*** *Oh no I hear you; I'm not saying these couldn't be done, it's just that that consideration is important if you're going to include them in all trials."*
>
> (SG1)

While not pre-specifying a total number of outcomes for the COS, the importance of selecting only those which were critically important given the burden imposed on researchers and participants was emphasised. The trial manager representative mentioned the challenge of prioritising what to measure without also considering how the domains would be measured:

> *"Obviously as a researcher you're always thinking about what's the feasibility of getting this data, so what's the volume of data that you're collecting…how are you going to measure these things? Because if each of those questions is just like one thing, like "how is your sleeping" or "how is your dressing"…there might just be a single question. So therefore that's very, very, very simple, it can be done in a couple of seconds. Whereas with things like range of movement and dynamometers you need specific equipment and stuff, you need people to come into hospital."* (Trial manager, SG2)

There was a concern that the domains related to patient satisfaction were not necessarily 'standalone' outcome domains as they brought together several others as well as concepts of patient experience and expectation which would prove very difficult to measure:

> *"This may be a bit controversial but satisfaction isn't actually an outcome, it's a composite of experience, expectation and outcome, so I'd vote that actually satisfaction isn't part of a core outcome. It's obviously vitally important, but it's not something you'd measure as part of a core outcome set."* (Surgeon 1, SG1)

Despite this, there was agreement within the same small group that 'patient satisfaction with outcome / result' was of value to patients despite any difficulties in the clinimetrics or how it might be assessed:

> *"When you're looking at outcome measures and what we use them for to inform patients about various treatments, one of the strongest things that you can tell them is about what patients think of the result a year down the line. I know it's a complex matrix and depends on all sorts of things, so from that perspective it doesn't look at patient function, but if you say to patients "well this is the score number that you'll achieve", that's not something that people can interpret usually. But if you say "there's a 90% satisfaction rate with this procedure or this injury at one year", then patients…find that quite useful information."* (Surgeon 2, SG1)

6.4.3.3 When to measure domains

Although also considered an aspect of 'how' to measure outcomes, some participants raised the issue that the specific timing of measurement could impact on which domains should be prioritised for the COS. Their priorities could be influenced by time since the injury:

> *"It's the initial stages when you've got the break and you can't actually do anything and that's when you struggle with the 'self-hygiene and personal care'…It's as important I think as eating, drinking and…but if you've recovered from your injury,*

> *albeit you've got limited movement…If you're measuring it while they've still got the injury then that would be critical, but once they've got over the injury…"* (Patient, SG1)
>
> *"I suppose with a measure like ['being able to control a vehicle for transport'] it depends when it's being taken. If it's taken sort of 12 weeks after your injury, maybe you were back driving OK by then. If it's taken at the 6 week mark, maybe not"* (Therapist, SG2)

6.4.3.4 Domain grouping with one or more others

Despite the pre-meeting information pack informing participants of a dedicated session later in the meeting to develop grouping of domains, participants were inclined to group together some domains at this prioritisation stage. At times this led to the associated domains being placed within the same prioritisation category.

The informal grouping of domains proposed by some was not necessarily agreed upon by all participants within each small group. There was discussion around the three 'pain / discomfort' domains:

> *"The 'pain and discomfort' linked with the 'pain and discomfort during activity' and the 'pain and discomfort during rest', I've actually got those three things as all being encompassed in one outcome"* (Patient, SG1)
>
> *"I think pain during activity and pain at rest are different things…and pain at rest is a much worse situation…And I think it's critical that we capture that. My view would be pain during activity and pain during rest should be critical and pain in general…pain/discomfort in general is not"* (Surgeon, SG3)

There was also consideration of what grouped with ‘personal care’ as compared to a broader concept of ‘activities of daily living’:

> *“’Dressing’ I think is covered by ‘personal care’ so probably doesn’t need to be in”* (Surgeon, SG1)
>
> *“’Fine use and dexterity’ could also cover things like…‘eating and drinking’, ‘self-hygiene’, ‘grasping and moving of light objects’ and ‘cooking and preparing meals’…I group all of those in ‘activities of daily living’”* (Therapist, SG1)
>
> *“I'm not sure ‘cooking and preparing meals’ really comes under ‘self-care’. It’s more an ‘activity of daily life’…Because you can eat, if somebody else makes your meal for you, it’s whether you can then eat it, isn't it, that’s the important factor in your ‘personal self-care’”* (Health economist, SG3)

6.4.3.5 Domain redundant/covered by information from other domains

In parallel to the ‘grouping of domains’ theme, there were some outcome domains which participants felt overlapped with one or more other domains to the point that information gained by assessing the former would make the latter (unmeasured) domain redundant. This was distinct to the previous theme in which domains were grouped, because in this case participants were judging whether any meaningful information was gained by including all the related domains in a COS.

At times this was used to justify a lower prioritisation for a domain that participants felt was accounted for by another domain that had already been highly prioritised. Some examples were domains that appeared relatively granular or specific which were effectively redundant in the face of domains with a seemingly broader scope:

> ***"Patient:*** *'Keyboard and mouse/touchscreen use' should go because if you've got 'fine dexterity' then you're going to be able to do that, aren't you?*
>
> ***Therapist:*** *I agree. And if you've got work and that's your work, you're also measuring it in that one as well."*
>
> (SG2)

> ***"Surgeon:*** *I think if we have 'fine hand use and dexterity', we don't need 'grasping and moving light objects'…I'd take it all the way to [definitely not needed]…*
>
> ***Therapist:*** *Agree with that.*
>
> ***Patient:*** *Yes."*
>
> (SG3)

> *"I think 'cooking and preparing meals', 'grasping and moving light objects', should be actually in the critical but they're covered by other things"* (Patient, SG2)

'Eating and drinking' was also a typical example of a domain felt to be 'important but not critical' owing to the overlap with other 'broader' domains.

> *"'Eating and drinking', a lot of the activities which are global arm movements along with 'fine dexterity', will be included with 'self-hygiene'. So brushing your teeth, you have to do the same actions as that for 'eating', for instance, although it's a finer thing. So there's a lot of stuff that's crossing over"* (Surgeon, SG2)

This clearly presented a challenge to some of the participants when deciding on the priority of the various domains: *"If you have 'fine hand use and dexterity', doesn't that cover 'eating and drinking'? It's difficult to [prioritise] these because they all seem really important but then I think some of them are covered by other things"* (Patient, SG2).

At times, this revealed disagreement when trying to decide which of the overlapping domains should take priority:

> ***"Patient:*** *I think if you can cook and prepare meals you'd be able to eat them. So I think you need to leave cooking and preparing [in 'important but not critical'] but you could take away the eating and drinking."*
>
> ***Surgeon:*** *"Well I'd spin that around really. I'd say that everyone needs to eat and drink but not everyone needs to prepare food, so if you're living in a nursing home for instance you might not actually be preparing your food, but you're always going to need to eat and drink."*
>
> (SG1)

The domain with broader scope was not always the one given higher priority though, as exemplified by the two 'patient satisfaction' domains. While 'overall patient satisfaction' technically encompasses 'patient satisfaction with outcome / result', participants prioritised the latter on the basis that it actually addressed the main concern of patients: *"It's more the outcome…that's what's going to affect you for the rest of your life"* (Patient, SG3). The concern was that broader 'overall patient satisfaction' would actually dilute the key aspect of patient satisfaction.

This theme also linked to another – the concept of abstract versus practical domains. Participants mentioned the apparent overlap between abstract and practical domains:

> *"So you need to have enough 'sensation', 'muscle power' and 'range of movement' to be able to look after yourself and do your job and so on, but actually those things are already covered by the other critical outcomes"* (Surgeon, SG1)

> *"I think…the 'stability of a joint' would translate into 'fine hand use' or 'return to work', all those things. I think that's how it's going to manifest really"* (Therapist, SG2)

6.4.3.6 Relative importance of abstract and practical/functional outcome domains

This theme arose through participants working beyond the idea of overlap of abstract and practical domains and proceeding to discuss their relative priority. Prioritisation was challenging because of the interplay between the two categories of domains: *"I think the problem is that we are comparing…very different things – one thing is activity and different parts in activity and the other things are function. So that's hard to compare because there's two different things and to do activity you need the function"* (Therapist, SG3).

Several clinicians advocated for practical/functional domains being a priority over the more abstract ones:

> *"'Fine hand use and dexterity', you need to have 'range of movement'. And then being able to 'do your job' and being able to 'cook and prepare meals', you need some 'strength or muscle power' but it's just whether we decide to measure these in a purely sort of abstract, clinical way or whether we measure them in a sort of functional patient way"* (Therapist 1, SG2)
>
> *"I do think as therapists we can get a bit obsessed by actual physical measurements of things, which don't always translate to how the patient's managing"* (Therapist 2, SG2)
>
> *"Comfort and capability are more important than measures of impairment or objective impairment or pathophysiology"* (Surgeon 1, SG2)

The trial manager representative similarly prioritised functional outcome domains over abstract ones: *"I think how it manifests itself for the patient is what's important. So you know, the consequences of not being able to move your wrist or to be able to grip something properly…if you're talking about what's absolutely critical and being totally ruthless, maybe you don't really need to have ['range of movement' and 'strength or muscle power']"* (Trial manager, SG2).

Some clinicians who placed high value on the importance of certain abstract domains though, believing that measuring a fundamental physiological parameter had broader applicability because by extension it could imply greater functional capability: *"I think we will start showing that strength is going to be a key parameter of daily tasks. There are some tasks you require certain adequate strength to do and whilst I accept they may just be covered in the task, as it is very measurable we can see…what tasks are then achievable"* (Surgeon 2, SG1)

For others, there was a concern that assessing practical domains might mean that important information is missed in cases where patients have adapted to bypass hindered function: *"I'm kind of arguing to keep [abstract domains] because it makes us measure something specific to the injury pattern or the treatment pattern, whereas you may be able to modify 'dressing', you may 'go back to your work' but your hand may still be stiff, so it doesn't necessarily – they're kind of general outcomes but we also want something specific to the area of injury"* (Surgeon 2, SG2). This appears to be a concern more related to appropriate measurement of the domain though.

Interestingly, patients tended to advocate for inclusion of the abstract domains (though not necessarily at the expense of the practical ones):

> *"I think the 'self-hygiene/personal care'…I don't see how that can possibly apply to 'strength or muscle power'…I mean I'm a gardener and I need my 'strength and muscle power' so I think that is important…I don't see how 'self-hygiene/personal care' and that stuff covers the 'strength or muscle power'. It might cover 'range of movement'"* (Patient, SG2)
>
> *"[I] had a wrist fracture, it was my right wrist and I'm right-handed, so I was very, very debilitated. Luckily for me, I've got almost the same range of movement back…it's not quite back, but enough for me to do all that I did before. But I think it's critical for that 'range of movement' to be measured and the 'strength'"* (Patient, SG1)

By the end of the small group discussions it appeared that participants generally prioritised practical/functional domains but believed the abstract ones were also important and had their place: *"There are times when the objective outcome is the outcome you're interested in, something like healing, motion, something you can objectively measure. Then there are things – most often you're mostly interested in how your hand works for you – "are you comfortable, are you capable?" – and then there's also the experience of care…satisfaction type aspects"* (Surgeon 1, SG2).

6.4.3.7 Domain influenced by pre-injury factors

The domain that prompted this discussion theme was 'sleeping'. Patients expressed a concern that although the concept of a sound sleep was important, it could be impacted by many other factors separate from the injury and therefore may not be a useful to include in the COS:

> *"Can I just say sleeping for me is intermittent and…before I had my wrist fracture I'd got a problem. So I couldn't decide whether it was the wrist fracture that caused sleeping problems or not. So it is important but…I don't see how sleeping can be in the core outcomes because other things affect sleeping."* (Patient, SG1)

A surgeon agreed that sleep might not be of good quality as a default prior to injury: *"of course many patients are like me in their 50s and menopausal and actually sleep is a luxury"* (Surgeon, SG1).

6.4.3.8 Ability to bypass hindered function

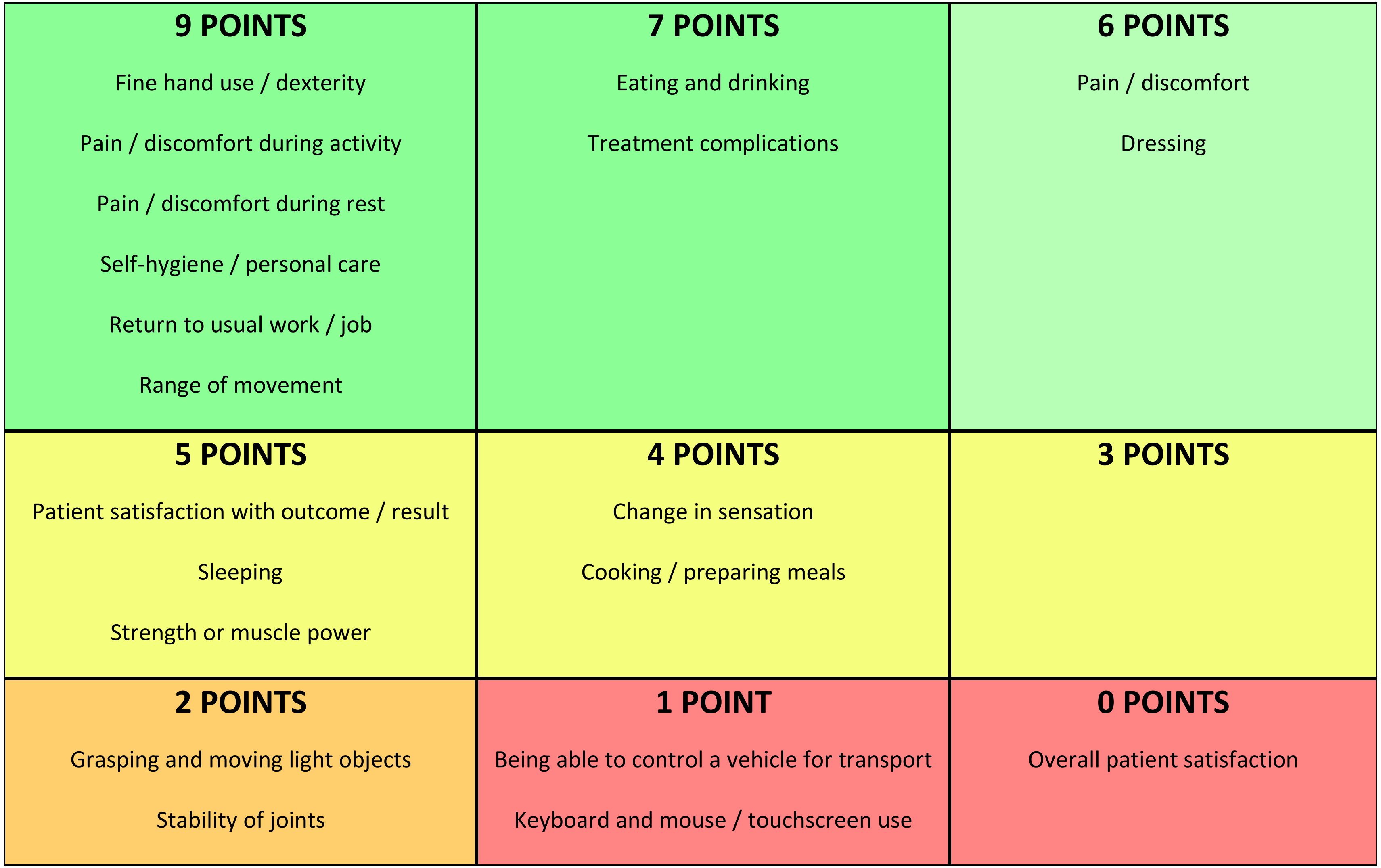
A few participants highlighted that it is possible for some functional impairments to be overcome, such that measuring these sorts of domains may not provide information that is necessarily of importance and perhaps justifying a lower prioritisation category than other domains where the impact could not be similarly lessened:

> *"If somebody's job is a keyboard and mouse occupation and their hand fractures are to such an extent that they're actually a little bit incapacitated in that, then it is something that one needs to look at, but there's lots of assistive devices, you can use audio recognition, voice recognition, computer technology and so on"* (Therapist, SG1)
>
> *"Some patients…may adapt. They may be able to do those activities but they may not do them in the way they did them before their injury or their treatment…So you may be able to still do personal care or self-hygiene but you may not use the hand that's got the hand fracture"* (Surgeon, SG2)

6.4.4 Prioritisation of outcome domains for inclusion in COS: whole group pre-vote discussion

The three prioritisation categories from each group were aggregated to show an overall prioritisation of the domains (Figure 6.4). This was revealed to all participants, following which there was whole group discussion to allow panellists to share their views on the outcome domains to be voted on. Participants were reminded that domains with an aggregate score of zero or one points would not be eligible for the vote and were removed from consideration for the COS.

Figure 6.4 Prioritisation of outcome domains based on aggregation of small group priority categorisation task



9 POINTS Fine hand use / dexterity Pain / discomfort during activity Pain / discomfort during rest Self-hygiene / personal care Return to usual work / job Range of movement	**7 POINTS** Eating and drinking Treatment complications	**6 POINTS** Pain / discomfort Dressing
5 POINTS Patient satisfaction with outcome / result Sleeping Strength or muscle power	**4 POINTS** Change in sensation Cooking / preparing meals	**3 POINTS**
2 POINTS Grasping and moving light objects Stability of joints	**1 POINT** Being able to control a vehicle for transport Keyboard and mouse / touchscreen use	**0 POINTS** Overall patient satisfaction

In the rest of this section, I summarise the main pre-vote discussion points, organised according to the specific domains being considered. While the Chair steered discussion towards those domains scoring four to seven points in an effort to help guide decision-making for the final vote, participants were free to raise other domains for discussion if they felt this was important. Nevertheless, several domains did not have discussion points raised and are therefore not covered within the following summary.

‘Pain / discomfort’ (6 points), ‘pain/discomfort with activity’ (9 points) and ‘pain/discomfort with rest’ (9 points)

Participants wanted to discuss this trio of domains due to concern that at the voting stage the underlying concept of pain/discomfort might be lost unless some level of agreement was reached at this discussion stage: *“The risk is that although everybody’s identified that some form of ‘pain and discomfort’…as a gross concept is important…I just wonder if having three things which are very similar and then asking us to vote with quite a high threshold of 80%, you get a split vote…we need to discuss and address how we might approach that when it comes to the poll”* (Therapist).

The Chair reiterated that all domains thought to be critical should be voted ‘yes’. Although the goal was to prioritise the core domains, this should not be at the expense of any believed to truly be essential for the COS. It became clear that participants felt it was reasonable to not prioritise all three domains, especially considering the disparity in aggregate score (the broader ‘pain / discomfort’ domain scoring just six points): *“I’m thinking that if we have ‘pain and discomfort during activity’ for nine points and we have* *'pain and discomfort during rest' for nine points, that we don't need 'pain and discomfort'…because we've already got it twice in our critical outcome set"* (Surgeon).

The conversation then focused on reasoning behind this disparity:

> *"'Pain and discomfort' is an umbrella term and it can be used, but if we're looking at…hand and wrist fractures, then perhaps it's better to do it as during activity and during rest because there are very broad sweeping pain assessments, like centralised pain for neuropathic pain and all of that kind of stuff that wouldn't necessarily pinpoint the point that one has post-fracture and during that rehab phase"* (Therapist)

The small group facilitator summarised the reasoning from their small group session (which had categorised the broader 'pain / discomfort' domain as 'definitely not needed in the COS'):

> *"Discussion from our group was…obviously they're all important because they all encompass 'pain and discomfort', but 'pain and discomfort at rest' is…clinically very relevant because it might indicate that something more significant is going on, i.e. deep infection, and if we're going to have 'during rest' then you should probably have 'during activity' as well. That was the reasoning in our group…for 'pain and discomfort' being given a zero"* (Facilitator – SG3)

<u>'Strength or muscle power' (5 points)</u>

Some participants advocated for including this domain in the COS:

> ***"Surgeon:*** *I think that 'strength' is important…I mean you could just do that in terms of function, you know, the ability to do daily tasks and so on, but I think it is quite an important independent variable.*

> ***Patient:*** *There's a lot of problem with this bunching things together I think, you know, but I agree, I think 'strength and muscle power' is separate from 'fine hand use / dexterity' or 'range of movement', so I think it should have higher points."*

Others felt that despite its importance, it was not a far-reaching domain critical for the COS: *"I personally don't think 'strength or muscle power' is – I mean to a certain extent it's hugely important but beyond that it's not"* (Patient).

<u>'Stability of joints' (2 points)</u>

While scoring only two points, this domain clearly resonated with a few patient participants: *"I think stability of joints is really important, depending on your injury of course, but it's all about the confidence and the ability to use your hand subsequent to the healing process, in a manner which is normal"* (Patient).

The clinicians' perspective was that this domain was of consequence for only a small subset of the injuries with the COS scope: *"One of the problems with stability is that even patients with some instability are often not functionally troubled by it…everyone with a wrist fracture has distal radioulnar joint instability and they rarely report it. So it's probably not a very material outcome for most hand injuries…there are a few specific sites…such as thumb [metacarpophalangeal] joint injuries, then you would add that in separately"* (Surgeon)

<u>'Cooking and preparing meals' (4 points) versus 'eating and drinking' (7 points)</u>

The explanation for a low prioritisation of 'cooking and preparing meals' as opposed to 'eating and drinking' by one of the small groups was that the former applied to a narrower range of patients: *"The logic was that everyone needs to eat and drink; not everyone actually* *prepares their own food and so on…therefore it wasn't a critical outcome for the purposes of this exercise"* (Surgeon).

'Patient satisfaction with outcome / result' (5 points)

Discussion on this domain was prompted by a patient participant raising concern that with 'overall patient satisfaction' already eliminated by virtue of scoring zero points then without retaining this domain, any direct assessment of patient satisfaction would be lost from the COS: *"'The patient satisfaction [with outcome / result]' has only got five points…I mean we decided that the 'general patient satisfaction' should be out but I'm a bit concerned about 'patient satisfaction with outcome/result' not getting very high points"* (Patient).

A therapist from one of the two of the small groups that had rated the domain as only 'important but not needed' rather than 'critical' for the COS explained their reasoning:

> *"We decided on 'patient satisfaction of outcome result' as opposed to 'overall patient satisfaction' because all the other aspects of the patient's rehab and therapy and all of that might not be specifically pertaining to the actual fracture process…And then we didn't rate it very highly because it's difficult to objectively assess it and so outcome measures don't often include that in their questionnaires…We do regard it very highly"* (Therapist)

A surgeon explained their perspective on the value offered by assessing the domain:

> *"It's such an important part of the outcomes that I collect, satisfaction, 'friends and family test' and 'normal hand score', these are the three pillars that we base the recommendations that we make to our patients on, rather than the 20-point change in DASH or the 0.3 change in EQ5D. So I feel that it's a very important measure to*

> *inform patients…it's like a patient-to-patient recommendation for what we're doing"* (Surgeon)

The Chair provided guidance to participants that when it came to the vote, all domains should only be considered in terms of their importance for the COS. Concerns about 'how' to measure the selected domains should not influence this decision and would be resolved in future work.

6.4.5 Prioritisation of outcome domains for inclusion in COS: vote for COS and final discussion

6.4.5.1 Voting results and final selected COS

Results of the vote are presented in Table 6-4, in descending order of percentage in the 'Total – Yes' column (as this was the one which mattered for the pre-determined consensus threshold of ≥80%). A breakdown of the voting across stakeholder groups is also provided.

Table 6-4 Voting results for domains to be included in the final COS

Outcome domain	**Patients (n=12)**		**Surgeons (n=7)**		**Therapists (n=6)**		**Trial manager / Health economist (n=2)**		**Total (n=27)**	
	Yes	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**
Self-hygiene / personal care	12 (100%)	0 (0%)	7 (100%)	0 (0%)	6 (100%)	0 (0%)	2 (100%)	0 (0%)	**27 (100%)**	0 (0%)
Pain / discomfort during activity	12 (100%)	0 (0%)	7 (100%)	0 (0%)	6 (100%)	0 (0%)	2 (100%)	0 (0%)	**27 (100%)**	0 (0%)
Fine hand use / dexterity	12 (100%)	0 (0%)	6 (86%)	1 (14%)	6 (100%)	0 (0%)	2 (100%)	0 (0%)	**26 (96%)**	1 (4%)
Return to usual work / job	12 (100%)	0 (0%)	7 (100%)	0 (0%)	5 (83%)	1 (17%)	2 (100%)	0 (0%)	**26 (96%)**	1 (4%)
Range of movement	11 (92%)	1 (8%)	5 (71%)	2 (29%)	6 (100%)	0 (0%)	2 (100%)	0 (0%)	**24 (89%)**	3 (11%)
Pain / discomfort during rest	10 (83%)	2 (17%)	6 (86%)	1 (14%)	6 (100%)	0 (0%)	2 (100%)	0 (0%)	**24 (89%)**	3 (11%)
Patient satisfaction with outcome / result	11 (92%)	1 (8%)	5 (71%)	2 (29%)	6 (100%)	0 (0%)	1 (50%)	1 (50%)	**23 (85%)**	4 (15%)
Eating and drinking	8 (67%)	4 (33%)	6 (86%)	1 (14%)	5 (83%)	1 (17%)	2 (100%)	0 (0%)	21 (78%)	6 (22%)
Treatment complications	9 (75%)	3 (25%)	4 (57%)	3 (43%)	4 (67%)	2 (33%)	2 (100%)	0 (0%)	19 (70%)	8 (30%)
Strength or muscle power	9 (75%)	3 (25%)	4 (57%)	3 (43%)	2 (33%)	4 (67%)	2 (100%)	0 (0%)	17 (63%)	10 (37%)
Dressing	8 (67%)	4 (33%)	1 (14%)	6 (86%)	4 (67%)	2 (33%)	2 (100%)	0 (0%)	15 (56%)	12 (44%)
Sleeping	8 (67%)	4 (33%)	0 (0%)	7 (100%)	2 (33%)	4 (67%)	2 (100%)	0 (0%)	12 (44%)	15 (56%)
Cooking / preparing meals	6 (50%)	6 (50%)	2 (29%)	5 (71%)	2 (33%)	4 (67%)	1 (50%)	1 (50%)	11 (41%)	16 (59%)
Change in sensation	6 (50%)	6 (50%)	3 (43%)	4 (57%)	2 (33%)	4 (67%)	0 (0%)	2 (100%)	11 (41%)	16 (59%)
Grasping and moving light objects	7 (58%)	5 (42%)	0 (0%)	7 (100%)	2 (33%)	4 (67%)	0 (0%)	2 (100%)	9 (33%)	18 (66%)
Pain / discomfort	5 (42%)	7 (58%)	2 (29%)	5 (71%)	1 (17%)	5 (83%)	0 (0%)	2 (100%)	8 (30%)	19 (70%)
Stability of joints	7 (58%)	5 (42%)	0 (0%)	7 (100%)	1 (17%)	5 (83%)	0 (0%)	2 (100%)	8 (30%)	19 (70%)

The final COS, along with the descriptors that were used to define the selected domains throughout the Delphi and consensus meeting processes, is presented in Table 6-5.

Table 6-5 Finale core outcome domain set and accompanying domain descriptors

Outcome domain	Descriptor
Self-hygiene / personal care	Being able to do the usual tasks involved in maintaining one's own hygiene and self-care e.g. washing oneself, toileting, washing hands, washing the face, brushing teeth, shaving, looking after one's hair, applying make-up
Pain / discomfort during activity	Discomfort or pain in the hand or wrist specifically during activities (NOT at rest) e.g. ache, shooting pain, sharp pain, throbbing, discomfort / pain due to not being able to tolerate hot or cold sensation
Pain / discomfort during rest	Discomfort or pain in the hand or wrist specifically during rest (i.e. with the hand / wrist not moving, so NOT during activities) e.g. ache, shooting pain, sharp pain, throbbing, discomfort / pain due to not being able to tolerate hot or cold sensation
Fine hand use / dexterity	Being able to do fine motor tasks or precise activities with the hand/wrist e.g. writing, drawing, picking up coins from a table, using a key
Return to usual work / job	Being able to return to the work or job that one was doing prior to their hand/wrist injury (NOT including the financial impact of any lost income)
Range of movement	How much movement one has through the joints of the hand or wrist, whether active (i.e. moving it with the muscles of the injured side) or passive (e.g. if someone else were to try to move it for the patient)
Patient satisfaction with outcome / result	Satisfaction with the overall result from the patient's perspective (NOT with treatment or recovery process, but the end result only)

6.4.5.2 Post-vote discussion

The Chair, facilitators and I discussed the voting results prior to revealing these to participants. As the process had selected only seven domains for inclusion in the COS we decided that there was little value to be gained by proceeding with the original plan of small and whole group discussion of domain grouping. That process had envisaged several more domains with potential overlap being included in the COS as a result of the vote.

Instead, after revealing the voting results and final COS, the Chair invited reflections from participants on the COS achieved and on the steps taken to reach the COS. In the rest of this section, I summarise the main discussion points.

Participants' thoughts on the final COS

On first impressions the range of domains included in the COS seemed to cover key priorities from the preceding discussions: *"It looks good. I mean…those categories obviously do contain a lot of the other things and I suppose it's a question of how you measure all that stuff, but…it looks to me like from all the discussion today, it covers most of those categories"* (Patient).

'Emotional / mood impact to self' and 'hobbies and recreation' were two domains that had not reached consensus via the Delphi and had failed to be salvaged in the early step of the consensus meeting. These generated some discussion:

> ***"Therapist 1:*** *I just wondered what people thought now at this stage about hobbies. I know we took it out very early on and we do have something to do with work, but we didn't keep in anything to do with the patients' feeling of self – I can't remember*

> *the wording for that one – or hobbies and I just wondered what people thought about that?*
>
> ***Patient:*** *That's true. The 'emotional health' and the 'hobbies' were all tied together weren't they and it's true, there isn't anything in there.*
>
> ***Therapist 2:*** *I certainly think it would have been more holistic…having the emotional aspect in there."*

There was some discussion about why the domain had not passed the 'salvage' vote. There might have been a split vote between these two domains: *"I think it probably got lost because there were a few things ['hobbies and recreation' and 'emotional / mood impact to self'] that everybody [was] looking at the same time and that maybe split the vote early on in the process"* (Surgeon).

There appeared to be a time-related aspect to the prioritisation of a domain such as 'emotional / mood impact to self', with a patient expressing that their mental health (which in itself is far broader than the definition of the domain that was voted on) was only impacted during a protracted recovery period: *"It depends on the timeline as well. Obviously initially when you first have your surgery, mental health isn't too unstable but from my personal experience, as time's gone on I was off for six months and I'm still not sort of back to full fitness by a long way, so my mental health…it had a massive effect on me"* (Patient).

Such a consideration certainly explained the voting decision of some of the clinicians:

> *"I think that all surgeons and all therapists should be looking at mental health and hobbies and sleep and things that we have discounted, but the reason why I didn't vote it very high was because our brief was a core set of outcomes for clinical research relating to hand fractures and injuries…There was an occupational therapy research*

> *article…people can tolerate adjustments being accommodated, light duty on work for two months, and then the mental health kicks in and then we need to start focusing more on function and hobbies and meaning in life…generally hand fractures and injuries are tending to be resolved before that point is reached"* (Therapist)

The COS covered several bases across the seven outcome domains that were selected though, and in a sense this was thought to offer potential insight into a general mental health picture for patients without directly assessing the domain: *"I'm actually fascinated by the fact that we ended up with a couple of comfort questions, a couple of capability questions, one impairment rating…objective impairment…and one experience measure – satisfaction. That's pretty brilliant actually. The place mental health will come in is when the symptoms and incapability are disproportionate to the impairment. So it's in there"* (Surgeon).

Participants' thoughts on the consensus meeting process

Despite the various challenges of running an online consensus meeting and a few technical issues on the day, the meeting was able to run as planned and completed on time. Participants' comments were accordingly quite positive: *"Doing these consensus meetings in general is difficult; doing them online is nigh on impossible so to get consensus in a systematic way like that is really impressive"* (Surgeon).

Several participants commented on the importance of having all key stakeholder groups in a single meeting:

> *"I think…having all stakeholders in the one place has been really interesting to me. Hope it's valuable in the long term"* (Patient)

> *"I think it worked well and the thing I'm very impressed with is how well it works having the patients with us because they have a completely different point of view and they're at the receiving end…[they] come up with much more important things"* (Surgeon)

The small group discussion was highlighted as a good opportunity to encourage more people to air their views: *"I liked the small group discussion and I thought the whole flow of the afternoon has gone really well and it is really interesting, reflecting on everybody's part and we all tend to stress our own opinion but it's good to work together for sure"* (Surgeon).

Finally, participants acknowledged the convenience offered by an online meeting, in particular facilitating international participation:

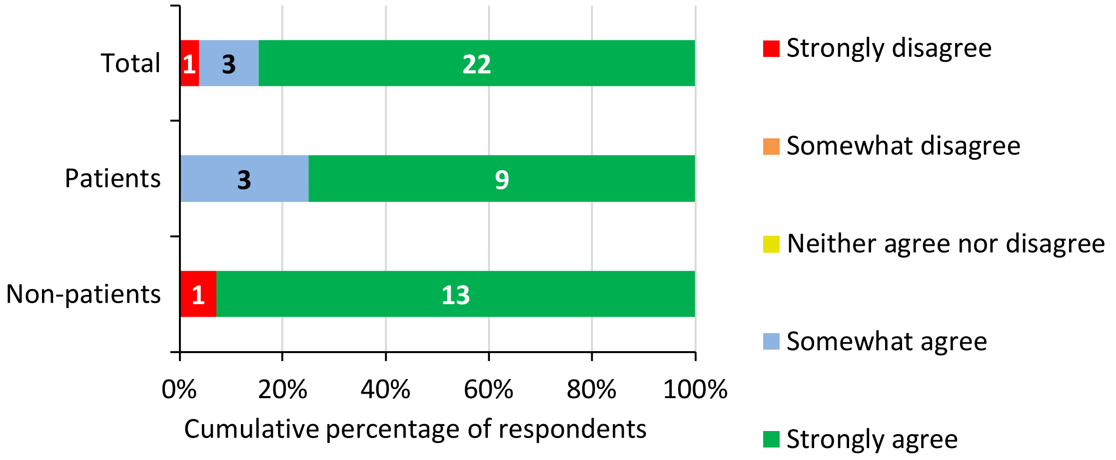
> *"I really like the online format…It's quite accessible for people and there's international members here"* (Therapist)

6.4.6 Feedback from participants

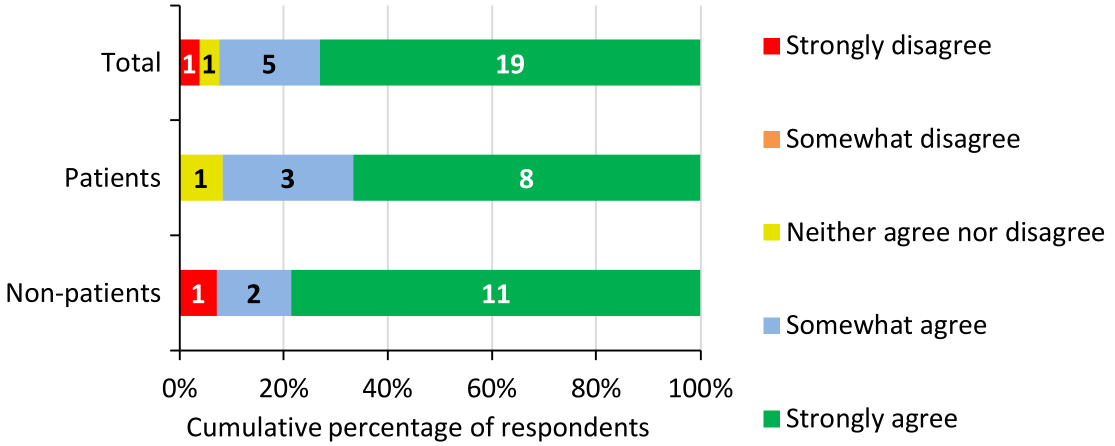
Since consensus meetings for COS development are typically face-to-face, the online format and specific methodology of this meeting made the process relatively novel. Although receiving some feedback in the closing discussion session of the meeting, I sought anonymised feedback on the meeting methodology as well as asked participants whether the final COS made sense and included domains of importance from their perspective. Of the 27 meeting participants, 26 provided feedback. Figure 6.5 shows the survey statements and level of agreement participants felt for each of these, with totals and a breakdown of patient versus non-patient participants.

On the whole, preparatory material for the meeting, the small group discussion session and overall timing of the meeting appeared very satisfactory to participants. The online platform for discussion and voting, as well as effectiveness of the whole group sessions appeared to work well for the majority too. Finally, it seemed that the vast majority of participants were in strong agreement that the final COS made sense and included critical outcomes from their perspective. This potentially reflects on the numerous steps taken in this study to seek key stakeholder views.

Figure 6.5 Cumulative bar charts of meeting participant feedback questionnaire ratings (numbers in bars represent number of participants for each rating)



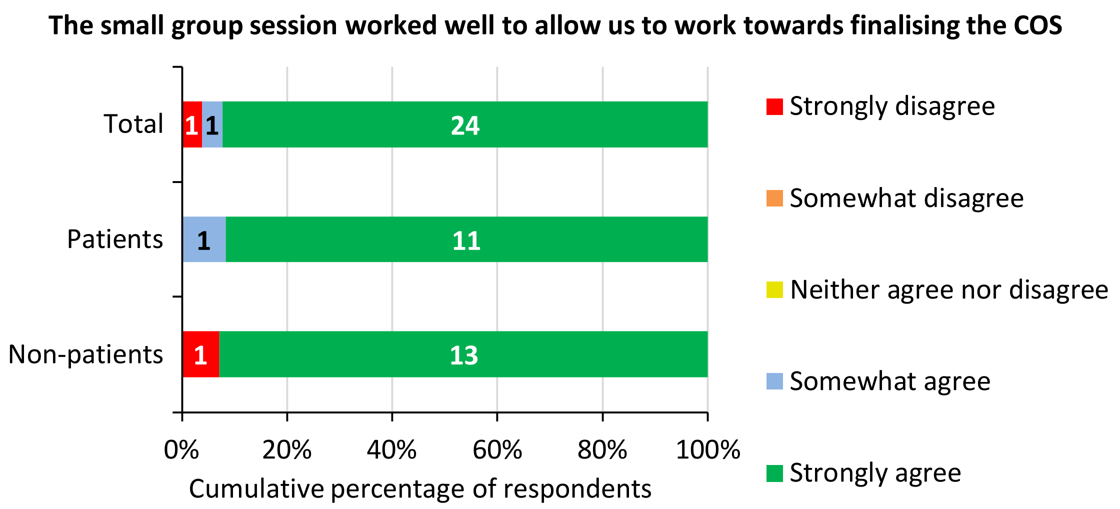
The pre-course pack was useful in helping me to prepare for the activities and discussions that took place in the meeting

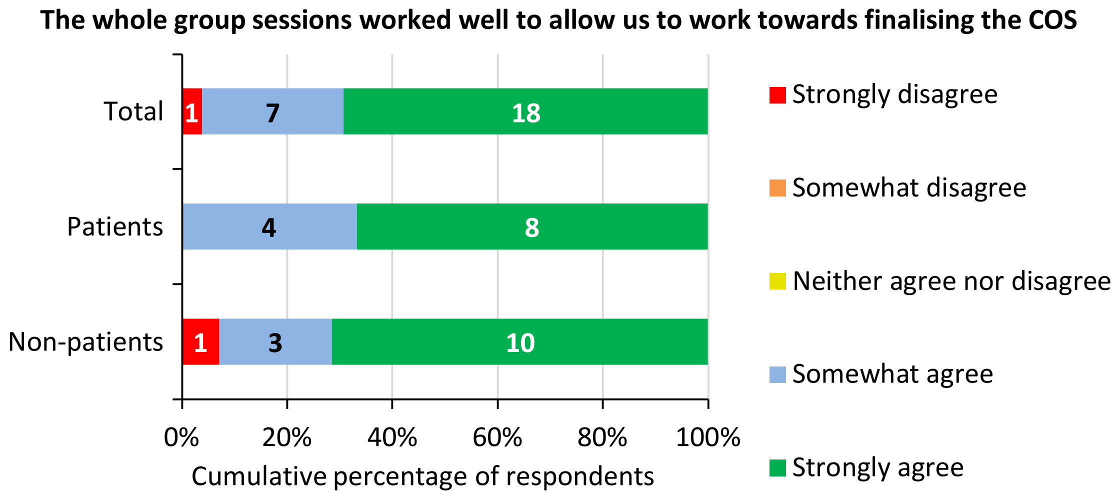


During the meeting I felt that I was able to express my views on the importance of outcomes and how they should be grouped in the COS

Figure 6.5 Cumulative bar charts of meeting participant feedback questionnaire ratings (cont'd)

(numbers in bars represent number of participants for each rating)





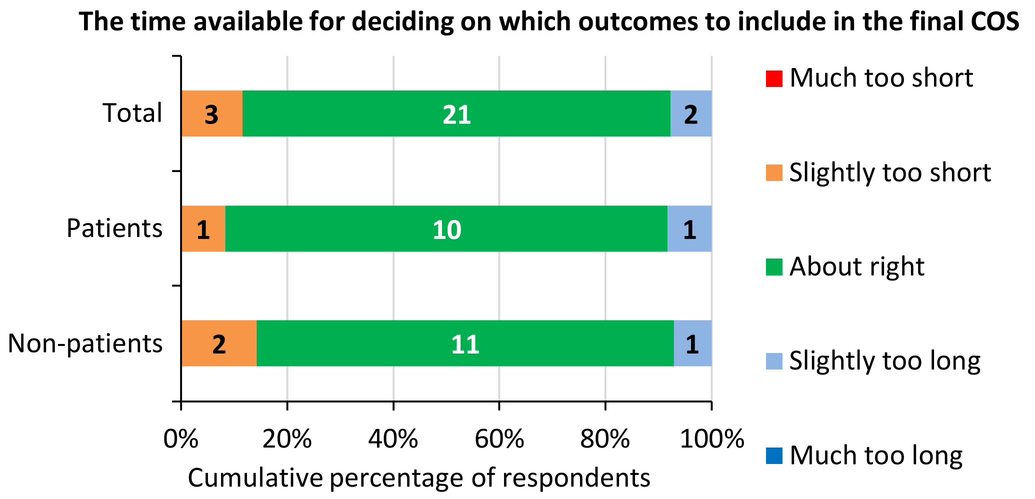
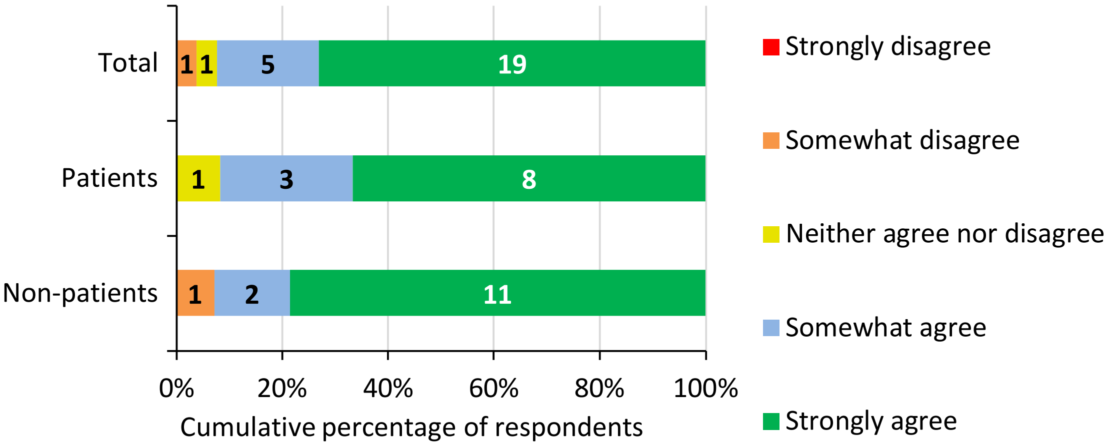


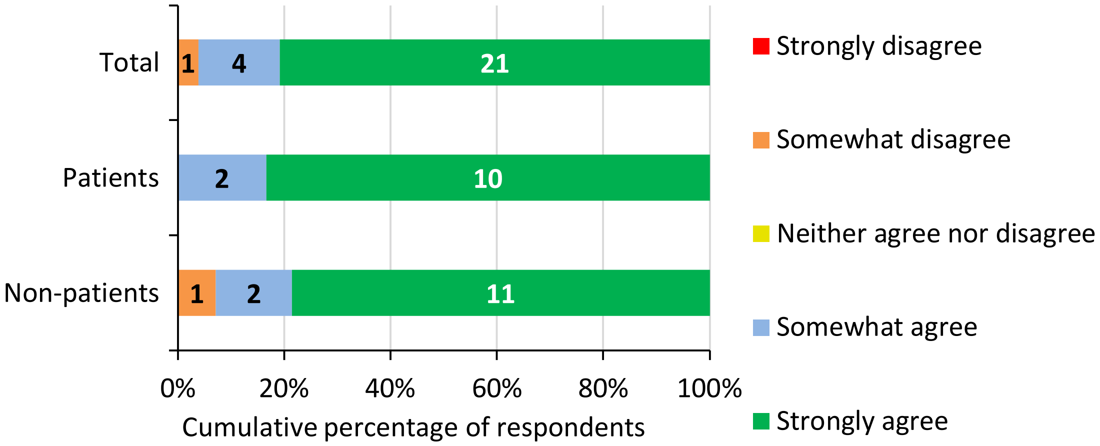
Figure 6.5 Cumulative bar charts of meeting participant feedback questionnaire ratings (cont'd)

(numbers in bars represent number of participants for each rating)

I felt comfortable and confident using the software to interact with others and casting my votes on the web-based polling system



By the end of the meeting I think that the COS produced made sense and included the critical outcomes from my perspective



6.5 Discussion

6.5.1 Key findings

The meeting panel reached consensus on seven outcome domains for inclusion in the COS. These domains achieved the high consensus threshold bar of ≥80% of meeting participants voting them as critically important to the COS. They covered a range of elements including:

- Patient comfort/pain
- Performing functional tasks (one relatively fundamental for maintaining independence and the other two perhaps working at a higher-order of functional level)
- A physiological domain
- Patient satisfaction

Although meeting participants had an opportunity to vote to salvage domains that did not reach consensus in or out via the Delphi, none of those 21 domains were retained for further discussion in the meeting.

Participants reflected on the selected domains and also commented on additional domains felt to be important to one or more of the stakeholder groups but not meeting the threshold for inclusion in the COS. There was acknowledgment that some research studies might include additional domains believed to be pertinent to the specific injury studied. This is entirely in keeping with the intended nature of a COS and its implementation (Williamson et al. 2017). Appendix 15 summarises the domains which did not reach the level required for inclusion in the COS but either highlighted as important to one or more stakeholder groups in the Delphi or nominated for the salvage vote in the consensus meeting. Overall, the meeting feedback demonstrated the vast majority of participants were satisfied with the process and final COS.

6.5.2 Reflections on the consensus meeting prioritisation process

The meeting aimed to determine which domains key stakeholders agreed upon as being of critical importance for inclusion in the COS. The overall COS aimed to retain the patient voice, so patients comprised the largest single stakeholder group at the meeting. This is atypical for COS consensus meetings; data from the COMET systematic review annual update reveals that patient participants comprise a median of ~20% of meeting participants (Gargon et al. 2021).

Despite having a pre-determined consensus threshold and binary voting decision on each domain, I endeavoured to understand the reasoning behind the decisions that participants made through qualitative analysis of the discussions. While not to the standard of the patient interview/focus group work for outcome generation (where data saturation was sought) this yielded some interesting insights into the decision-making process.

It was clear that several themes on the prioritisation of domains were associated with the issue of 'how' to measure domains, despite emphasising to participants that this should not factor in their decisions (Figure 6.3). This seemed particularly so for non-patient participants, probably due to their awareness of existing outcome measurement tools and familiarity with the importance of 'how' to measure outcomes in clinical research.

Whilst the domains were borne out of a synthesis of the systematic review and patient interviews/focus groups along with a few Delphi Round 1 participant suggestions, several meeting participants commented on the challenge of selecting a shortlist for the COS because of an apparent overlap between some domains. Another understandable difficulty resulted from the fact that all the domains being discussed had already been highlighted as 'very important' by the Delphi survey. There were no hard limits set for how many domains could be prioritised as critical for the COS. However, by tasking participants with categorising all these domains across three levels of ‘importance’ and stressing that the COS should consist of critical ones, there was an expectation that not all the domains would end up in the COS. Participants appeared inclined towards grouping of some domains during the prioritisation phase; potentially a strategy to retain as many of the domains as possible for inclusion in the COS. When grouping of domains was discouraged at that phase and participants were reminded to categorise each domain on its own merit, then the strategy appeared to adapt towards generally prioritising domains with a broader scope to include/cover many of the ones that were given lesser priority.

The aggregated prioritisations of the small groups fuelled further discussion in the pre-vote whole group session. This appeared to prove quite important for the subsequent vote, with two areas to highlight:

- Clarification of prioritisation of the various ‘pain / discomfort’ domains to avoid a ‘split vote’
- Discussion of the importance of ‘patient satisfaction with outcome / result’

The latter is particularly interesting because of the seven COS domains, six had scored an aggregate of nine points though the small group discussions. Neither of the domains at seven points made it past the voting threshold to enter the COS, although ‘eating and drinking’ was extremely close. The only domain to enter the COS with less than nine aggregate points was ‘patient satisfaction with outcome / result’. This scored just five aggregate points. Discussion identified that this domain had been assigned a lower priority by two of the small groups in large part because of concerns about how best to measure it. When participants were reminded to focus on the importance of what the domain is rather than how to measure it, the final vote better reflected the priority they assigned to the domain.

It is difficult to draw comparisons between the SO-HANDI meeting process and those of other COS developers due to the large variation in methodology and lack of a gold standard (Williamson et al. 2017). Furthermore, the restrictions of manuscript word counts usually mean that much detail about the process and findings of COS consensus meetings is not conveyed. Hence it is not possible to comment on how widespread the above observations are in such meetings.

6.5.3 Limitations

While I attempted to involve patient participants with a range of injuries and demographic characteristics, no patients with a ‘heavy manual’ occupation type were able to take part in the consensus meeting. The meeting being online-only also meant that any potential participants without access to appropriate equipment would be unable to take part. Unfortunately, it was not possible to facilitate access to the meeting by provision of a location with appropriate IT equipment due to COVID-19 restrictions. Finally, an inability/unfamiliarity with the necessary software to participate in the online meeting might have caused eligible participants to refuse to take part. To counter this problem, I reassured participants by providing practice sessions to familiarise them with the meeting software and voting platform that would be used. This probably served to largely prevent substantial technical issues arising at the start of the meeting, as most people unfamiliar with the software had elected to take part in a practice session.

A further pragmatic limitation secondary to the online medium was on the duration of the meeting. The Steering Group believed that the meeting should be planned for no more than a four-hour duration, including breaks. This is substantially less than all-day face-to-face meetings which have frequently been used for COS consensus meetings in the wider literature. The shorter duration led to some restrictions in the meeting structure. For example, whilst the James Lind Alliance workshop methodology outlined earlier (Section 6.3.1) consists of two rounds of small group discussion and ranking of items, I opted to have a single session. The overall approach still involved iterative prioritisation of outcome domains but resulted in a plan of less small group discussion time. The impact of this was ameliorated by flexibility in session timings on the day – we adjusted the original 45-minute small group discussion to 70 minutes based on contemporaneous feedback from facilitators.

Finally, while the clinician participant pool included international representatives the patient participants were all from the UK. The prioritisation of domains might therefore not be generalisable to a global context, where healthcare systems and cultural factors differ. This is a difficult challenge for COS development, due to various ethics approval hurdles for patient recruitment from different countries. Despite a recent trend in COS development to have involvement of participants from more countries, the degree of involvement varies greatly and participants in a given COS usually appear to be predominantly from only a small handful of countries (Gargon et al. 2021). There is a paucity of involvement of participants from low and middle income countries in COS development (Davis et al. 2018).

6.5.4 Strengths

A key strength of the meeting was the robust foundation of background information provided through the Delphi survey. In the process of developing the outcome domains for the Delphi, I established descriptors for all the domains which provided meeting participants a common language for discussion. I created a pre-meeting pack which summarised the Delphi results and clarified the meeting format. This allowed participants to gather their thoughts on the key tasks to be accomplished in the meeting.

Participant selection was mostly from the pool of Delphi survey participants who had completed all three rounds and therefore were familiar with the concept of COSs and the domains established through prior stages of the SO-HANDI project. I managed to involve patient participants with a reasonable range of injuries across the scope of the COS, and the number selected intentionally gave their stakeholder group the greatest voting power (12/27, 44% of votes). Meanwhile, the non-patient participants overall had a high level of experience in their respective fields, with even the least experienced of the clinicians having 13 years of experience of managing patients with hand fractures and joint injuries. Many (11/13, 85%) of the clinicians participating also had experience of running a research study themselves and were therefore able to consider selection of the domains for the COS from both a clinician and researcher perspective.

No domains from amongst those that had not reached consensus through the Delphi survey were voted through for salvage at the meeting, reflecting positively on the process and suggesting that the meeting participant panel was reasonably reflective of the overall Delphi population.

Small group discussion provided an opportunity for wider participation from even the quieter participants, as reflected in the participant feedback. I stratified allocation of stakeholders to the small groups to encourage discourse across stakeholder groups. This seemed intuitively important when trying to achieve consensus across different stakeholder groups. I believe that the whole group discussions, where some participants were perhaps less likely to voice their views, benefitted from the mix of stakeholders. Particularly in the past, some COS developers took an approach of separate consensus meetings for the different stakeholder groups, typically dividing patients and clinicians (Potter et al. 2015). The consequence is a challenge to somehow combine the separately formed priority lists into a unified COS, all without any direct interaction and discussion between the stakeholder groups.

A further strength was optimal use of the online nature of the meeting by including some international clinician participants. Face-to-face meetings can also have international participation, but these involve substantial travel time and costs for participants. Given the aim of a COS having uptake across global research in a health condition, it was important to have some level of international participation in the consensus meeting.

6.6 Conclusion

The consensus meeting comprised the final stage of the consensus prioritisation process. Seven outcome domains were selected for inclusion in the COS on hand fractures and joint injuries. All seven were selected from within the pool established as ‘very important’ by the Delphi survey across the key stakeholder groups of patients, hand therapists and hand surgeons.

Through an online format, the meeting was able to include international participants and arrive at a consensus through a pre-defined, systematic process. Transcripts from the audio recordings of the meetings were analysed and provided insight into the decision-making process by which participants prioritised domains.

The output from this meeting will form the basis of future work, in which the best way to measure the identified COS domains needs to be established. I discuss this and consider the final COS domains in comparison to other relevant research in Chapter 7.

Chapter 7 Discussion

7.1 Introduction

The aim of this PhD was to develop a COS for hand fractures and joint injuries in adults – the primary purpose was to establish what should be measured when considering the outcome of hand fracture and joint injury treatment. Rather than reiterating the key findings of each step in this process, which are already presented at the end of the corresponding chapters, here I:

- Reflect on my experience of COS development in relation to existing knowledge
- Compare and contrast to existing relevant work in the literature
- Offer consideration for potential outcome measurement instruments for the COS domains
- Outline avenues of future research and dissemination

7.2 Reflections on COS development process in relation to existing knowledge

7.2.1 The initial step – setting a scope

Despite general guidance in the form of the COMET Handbook there is no gold standard recommended approach, with many potential options in aspects of the COS development process (Williamson et al. 2017). Nevertheless, there is general agreement that a key initial step is to establish the scope (Williamson et al. 2012; Williamson et al. 2017). As described in Chapter 2, I experienced an initial challenge in setting the limits of the 'health condition' (i.e.

the specific set of injuries) targeted for this COS because of varying interpretation of the phrase ‘hand fractures and joint injuries’. This challenge can be avoided by setting a very narrow scope of ‘health condition’ but that would also result in a COS which is only applicable to a narrow subset of patients, thus reducing generalisability and applicability. Super-specialisation of COSs could ironically be its own form of research waste, when one of the benefits of a COS should be a reduction in such waste (Yordanov et al. 2018). Of course, if health conditions differ in a substantial way, such that the outcome domains of priority to stakeholders varies, then this likely requires more than one COS. Recent orthopaedic trauma COS development has however taken a more inclusive approach, with one recently published to cover all paediatric limb fractures (Marson et al. 2021).

The challenge might therefore lie in ensuring appropriate exploration of this issue, remaining aware of the possibility that as development proceeds the need for ‘splitting’ of an initial ‘broad’ COS may become apparent. I approached the careful balance in this COS by discussion with experts in the field and with a Steering Group. I proceeded to explore for potential differences in outcome prioritisation between two of the main injury groups (those with distal radius fractures versus all other hand fractures and joint injuries) in my qualitative work with patients and found no major discrepancy in the outcomes of interest to both groups (Chapter 4).

7.2.2 Identifying a longlist of outcome domains

The systematic review identified vast heterogeneity in the outcome domains selected in clinical research on hand fractures and joint injuries and thereby helping to establish the value of a COS in this field.

While widely accepted that a review alone offers insufficient exploration of stakeholders' views on potential outcome domains for a COS, there is still controversy over the best way to supplement the longlist created through the review process (Chevance et al. 2020; Maxwell and Beaton 2020; Williamson et al. 2020). I believe that the review offered a reasonable coverage of the outcome domains that might be important to clinicians and researchers. I therefore focused on a qualitative exploration of the views of patients with experience of hand and wrist injuries via interviews and focus groups. Through this I identified domains relevant to the patient stakeholder group and generated the wording of outcome domains and descriptors used in the Delphi survey. The latter was achieved through a synthesis of the review outcome domains (based on the WHO ICF) together with the qualitative study domains (based on wording used by patients and discussed in the focus groups).

Despite a rise in the proportion of COS development studies involving patient participants in recent years, it is often unclear precisely how patients have participated (Jones et al. 2017; Gargon et al. 2019b; Gargon et al. 2021). Where COS developers reported such details, it tended to show patients involvement in the prioritisation of outcome domains rather than their generation. A strength of this COS is that the patient voice was integral to the outcome generation phase as well.

While patients can sometimes find it challenging to understand the concept of outcome domains, qualitative methods involving analysis of the first-hand experience of patients should not be hindered by such an issue (Keeley et al. 2016). I developed a topic guide which focused on eliciting patients' experiences of injury, treatment and recovery, with interpretive analysis to generate the outcome domains based on their lived experience. I had opportunities in the interviews and focus groups to take a direct approach in exploring what outcomes mattered to patients; however, by emphasising the indirect approach patients had the arguably simpler task of taking me through their patient journey (Young and Bagley 2016). The approach also provided useful wording for the domain descriptors, with real-life examples for several domains drawn directly from the interviews and focus groups.

7.2.3 Prioritising outcome domains to include in the COS

Initial prioritisation involved a Delphi survey of a large panel of participants representing the key stakeholder groups. The shortlisted outcome domains and data from the survey were then used by a smaller group of representative stakeholders to guide final domain selection in a consensus meeting. The prioritisation phase of this COS was held online – the Delphi survey was planned as being online but the consensus meeting was originally planned as an in-person meeting. For the Delphi survey I was able to accommodate participants who preferred posted questionnaires.

There are several different approaches to the prioritisation phase of COS development – in the past many would simply finalise a COS based on Delphi results, but in recent years a growing proportion have held some form of final consensus meeting (Gargon et al. 2021). Even when a consensus meeting was held, the purpose of the meeting often differed. Sometimes Delphi results would determine the COS, with the meeting used only to add further domains as necessary (Damhuis et al. 2021). In other cases meeting panellists would work to filter down to a smaller shortlist of critical domains (Alkhaffaf et al. 2021). I opted for the latter approach as this respected the output of the Delphi by ensuring all ‘very important’ domains from the survey were considered at the meeting. However, it enabled meeting participants, who would have a forum for discussion on the reasons motivating prioritisation of domains, to select the final COS domains.

I believe the choices made for the Delphi survey were key to optimising the output. Presentation of individual stakeholder group feedback was beneficial, as hand therapists and surgeons frequently commented that seeing the patient participants’ rating distribution influenced a change in their own ratings. The Delphi involved a large number of participants and had a very low attrition rate, which lends greater confidence to the results (Gargon et al. 2019a).

The consensus meeting format and preparation were adapted to suit an online meeting. Preparation of participants with a clear information pack and offering ‘rehearsal’ sessions to familiarise participants with the online platform proved valuable investments of researcher time and likely translated into fewer technical issues on the day. The analysis benefitted greatly from the decision to record audio from the meeting. Through transcript analysis I developed an understanding of the decision-making process behind participants’ prioritisation of domains. While COS development papers often highlight an area of controversy or focused discussion in their consensus meetings, there is not usually an extensive exploration of the discussion. It is unclear whether this data is simply not reported, or rather not collected.

A limitation was that although some international participation was possible amongst the clinician stakeholder groups, the patients were all from the UK. This appears to be a widespread choice in COS development (Chevance et al. 2020). There are many challenges, including not merely superficial language barriers but also ensuring that the true underlying meaning of outcome domains is not lost when translating. Cultural factors and differences in healthcare systems may also have a role in how people from around the world prioritise outcomes. I believe the SO-HANDI COS has prioritised domains that would be applicable globally, as the domains are broad and capture fundamental concepts of function and the experience of patients with hand fractures and joint injuries.

The COS is relevant regardless of the intervention received by patients, as there are no intervention-specific outcome domains within the final COS. This is perhaps largely as a result of the overall patient participant pool involved in the outcome generation phase (qualitative study with interviews and focus groups) as well as the outcome prioritisation phase (Delphi and consensus meeting) having experienced a mix of surgical and non-surgical interventions. In addition, considering the overall scope of injuries included in the COS participants of the Delphi and meeting might have felt that intervention-specific outcomes could not be considered ‘core’.

However, while there did not appear to be distinct outcome domains arising with relation to surgical versus non-surgical interventions from the qualitative study, the systematic review did identify some unique outcomes in this regard. For example, ‘duration of surgery’ and complications such as post-operative infection would fall within the domains of ‘healthcare resources used’ and ‘treatment complications’, respectively. It is important to highlight that these two domains were not included in the final COS on the basis of being considered as domains that would likely be measured in studies by default. ‘Healthcare resources used’ did not reach consensus as ‘very important’ in the Delphi and was not salvaged at the early stage of the consensus meeting. ‘Treatment complications’ was quite highly prioritised in the small groups (receiving an aggregate of seven points) but did not reach consensus threshold for inclusion in the COS. I cannot be certain, but it appears that enough meeting participants believed that reporting of adverse events is performed by default in clinical trials and would effectively cover this domain.

7.2.4 Should ‘when to measure’ an outcome be considered in the ‘what to measure’ phase of COS development?

One particular theme became apparent by the conclusion of this study. The COMET Initiative has traditionally assigned consideration of the concept of ‘when to measure’ outcome domains to the ‘how to measure’ phase of COS development, which follows the ‘what to measure’ phase (Williamson et al. 2017). Although not disruptive to the SO-HANDI COS development processes, there was a recurring theme of participants raising the prospect that ‘when to measure’ the domains actually had a bearing on what domains should be selected for the COS in the first place.

In patient focus group discussion, it was noted that priorities can and do change over time: at various phases after the injury, through treatment and recovery, patients might have evolving priorities. For example, initially ‘pain’ would be of greater concern than later in the recovery process. Managing basic tasks to maintain independence, such as ‘self-care and personal hygiene’ and ‘eating and drinking’ might be very important in the early phase. As time passes their recovery might progress such that those basic functions are a given; instead, their priority might shift to achieving a higher level of function enabling social or work roles.

This theme was reflected in comments collected as part of the Delphi survey, with the commonest reason patients provided to explain a change in their ratings for domains being their evolving experience of the injury. Rather than altering their ratings because of how hand surgeons or therapists rated the domains, patients appeared to tend to still prioritise their personal experience of their injury – and because this was changing over the three to four months that the Delphi survey ran, so too did their ratings.

The theme of 'when to measure' domains was raised in the consensus meeting too, though only briefly and not proving disruptive when participants were working to prioritise domains to select for the COS.

The issue of 'when to measure' impacting on what domains should be selected for a COS in the first place does not appear to have arisen in COS development studies that I am aware of. It perhaps manifested for this COS because of the nature of the health conditions covered, which are acute in presentation. While many COSs are developed for chronic health problems where control of symptoms is the goal, perhaps it is logical that one on trauma brings this issue to the fore. Trauma has an evolving recovery process and patient experience of symptoms can change over a relatively short timeframe and have a permanency lacking in chronic conditions with a remitting/relapsing course. Counter to this theory, a COS on hip fractures did not highlight the timing of outcome measurement as being an issue when prioritising domain selection – however this project included only one patient representative and three carer representatives taking part in a survey, with the patient unable to attend the consensus meeting (Haywood et al. 2014).

From the systematic review it was clear that a broad range of time points were selected for outcome assessment in studies on hand fractures and joint injuries (Figure 3.2). The modal time bracket of six weeks to three months indicated that clinicians and researchers were not merely interested in the long-term (e.g. one-year) outcomes for patients but also earlier time points when domains of function, pain, etc, were unlikely to have reached a final steady state. I believe that the domains selected in the SO-HANDI COS strike a reasonable balance in addressing what are likely ‘early priorities’ and ‘longer-term priorities’ as they cover a range of bases such as ‘pain / discomfort’, movement, patient satisfaction and then different levels of function. The latter are on a spectrum from basic self-care and then function sufficient for fine hand use and even return to work. However, as this theme was only readily apparent by the conclusion of the domain selection process I was unable to directly address it in the methodology presented in this thesis. Nevertheless, this reflection might help shape the development of future COSs, particularly those in which it might be anticipated that the priorities of stakeholders are likely to change over time.

7.2.5 The impact of COVID-19

Precautions due to the COVID-19 pandemic led to an alteration from my original plans for face-to-face data collection. This impacted on the qualitative research and the consensus meeting.

Some patient interviews were completed before lockdown restrictions began, but all focus groups were conducted online. The same issue impacted a number of other studies during this time, with various researchers sharing their opinion on the different online platforms available and the degree of success as compared to traditional in-person focus groups (Aligato et al. 2021; Dos Santos Marques et al. 2021). Based on the responses of eligible participants who were offered the chance to join an in-person focus group (before lockdown restrictions), I believe that the online focus group approach in the COS development process resulted in greater participation than would have otherwise been achieved. This was likely due to the convenience factor – in my experience of approaching patients eligible to participate, the commonest reason not to take part in an in-person focus group was the extra effort to travel to a specified location. Informal feedback I received from focus group participants suggested that they viewed the experience fairly positively. The small group sizes meant that I was able to ensure that everyone participated and engaged with the discussion. It was reassuring to observe discussions between participants with minimal prompting from myself, much as I envisage would happen in a face-to-face focus group. One difference with the online focus groups though, is that there was not really the opportunity for simultaneous dialogues between participants (such as a pair of participants having a spontaneous side conversation while the main group has their own discussion) due to the limitations of audio pickup and clarity through the online medium.

There were challenges faced, with participant cancellations and 'no-shows' – this has been reported to be higher than for in-person focus groups (Daniels et al. 2019). An online-only approach also excluded those who lacked access or familiarity with the necessary technology to participate. I would therefore suggest that when possible, COS development studies consider implementing a combination of in-person and online qualitative data collection to maximise inclusivity whilst also optimising convenience to patient participants. The latter approach could theoretically facilitate data collection from international participants – however the logistics of multiple ethics approvals remains a challenge. I was unable to organise this and instead limited recruitment to patients from the UK.

In terms of the consensus meeting, sufficient time had passed that an online meeting format was quite a normal experience for clinicians, as well as many patients. I had taken numerous steps to try to optimise the meeting processes for an online format and was reassured by guidance published by COMET just prior to the meeting date (Gorst et al. 2021). This offered suggestions on how to run an online COS consensus meeting, which essentially fit with my existing plans. Again, feedback from participants (this time formalised) was reassuring in terms of the online format. From my perspective, I feel that the main limitation with the online meeting was that it had a relative lack of ‘mingling’ and ‘small talk’ between participants such as might occur during a coffee/tea or lunch break in a face-to-face meeting. I believe that this can help to develop a rapport and sense of familiarity amongst participants that could facilitate engagement during discussions. Nevertheless, the actual engagement during the consensus meeting appeared wholly adequate – whether achieved through the approach I selected, the nature of the participants or a combination of both, I cannot be certain.

Beyond merely the necessity of an online format for some of the research, the COVID-19 epidemic was a significant event for the population which could theoretically impact on the specific COS domains being prioritised. It is impossible to be certain of the impact the COVID-19 pandemic might have had on selection of COS domains. All participants of the study would have been affected by it, and in varying ways. However, I did not often encounter a reference to COVID-19 from participants explaining their prioritisation of domains, and don’t think the final COS would have been drastically different had the pandemic not occurred.

7.3 Findings in relation to existing relevant research

In this section I present a comparison of methodology and prioritised domains between this COS and prior research relevant to development of a COS for hand fractures and joint injuries. This offers the potential for some degree of triangulation of the 'critical' domains, though scoping and methodological differences limit this.

The SO-HANDI COS domains, along with the prioritised domains from existing relevant research are summarised in Table 7-1. For the WHO ICF Core Set, I have shown only those domains selected as part of the brief set; this is a more reasonable comparison of 'core' domains with the SO-HANDI COS than the comprehensive set of 117 domains.

The WHO Brief ICF Core Set has a far broader range of health conditions which it aims to cover and thereby has domains such as 'spinal cord and related structures' and 'structure of shoulder region' which seem of minimal relevance to a COS on hand fractures and joint injuries. The ICHOM standard set for hand and wrist conditions similarly aims for a broad range of health conditions, but they separated into five tracks. Hence, I can focus on the domains selected for those tracks relevant to hand fractures and joint injuries as defined for the SO-HANDI COS. Meanwhile, the core set of domains for distal radius fractures by Goldhahn et al (2014) focuses on only a subset of the injuries considered by the SO-HANDI COS.

Table 7-1 Summary of prioritised domains in other core/standard sets with scope including hand fractures and joint injuries

WHO Brief ICF Core Set for hand conditions (Rudolf et al. 2012)	**Goldhahn et al (2014) core set of domains for distal radius fractures**	**ICHOM standard set for hand and wrist conditions – thumb, finger and wrist** (Wouters et al. 2021)	**SO-HANDI COS**
Emotional functions Touch function Sensory functions related to temperature & other stimuli Sensation of pain Mobility of joint functions Stability of joint functions Muscle power functions Control of voluntary movement functions Protective functions of the skin Spinal cord and related structures Structure of shoulder region Structure of upper extremity Carrying out daily routine Lifting and carrying objects Fine hand use Hand and arm use Self-care Domestic life Interpersonal interactions & relationships Work and employment	**Core** Pain Function Participation in normal life roles (patient-specific) Complications **Optional** Performance (hand function/dexterity tests) Range of motion Proof of concept (intervention-specific) Patient satisfaction	**All tracks** Pain Hand function/activities of daily life Health-related quality of life Return to work Satisfaction with treatment results Complications Revision (surgery) **Thumb track only** Grip and pinch strength Range of motion* **Finger track only** Range of motion Grip and pinch strength* **Wrist track only** Grip strength Range of motion* * only if outcomes are to be measured (i.e. status change expected) at the one-year time point	Self-hygiene / personal care Pain / discomfort during activity Pain / discomfort during rest Fine hand use / dexterity Return to usual work / job Range of movement Patient satisfaction with outcome / result

7.3.1 Prioritised domains across core/standard sets relating to the hand and wrist

Despite these differences in scope, there is a sense of overlap in some key areas across the core/standard sets. All included domains covering concepts of pain, function and return to life roles, with return to work being common to all four. Range of movement is common to all except the core set of domains for distal radius fractures, where it was highlighted as a recommended ‘optional’ domain. There are clearly some fundamental domains related to injuries of the hand or wrist which, despite the variation in methods as described in Chapter 1, have led to consistency in the final recommended sets of domains. This is considering the concepts of ‘function’ and ‘life roles’ at a very broad level – the challenge was to define what these meant in relation to hand fractures and joint injuries, so that one can then consider how best to measure the domains effectively.

One approach to defining the specifics of these concepts was to use more granular domains such as in the WHO Brief ICF Core Set. Indeed, much of the Core Set shows great overlap with the 20 domains that reached consensus agreement as ‘very important’ in the Delphi survey. However, the consensus meeting achieved consensus prioritisation from amongst these.

One domain obviously lacking in the WHO Brief ICF Core Set compared to the others was patient satisfaction. Of course, the Core Set was constrained by the WHO ICF framework which does not include a ‘patient satisfaction’ domain. I recognised the issue of constraints such as this, given the various domains generated through the systematic review which were not covered by the WHO ICF framework. This partly prompted my drive to generate outcome domains from the patient interviews and focus groups through an inductive approach rather than using the WHO ICF framework.

The fact that ‘patient satisfaction’ was a recommended ‘optional’ for the Goldhahn et al (2014) core set of domains for distal radius fractures but incorporated as part of the ICHOM standard set and the SO-HANDI COS may reflect a growing recognition of the importance of this domain over time (Anhang Price et al. 2014).

7.3.2 Detailed comparison with ICHOM standard set for hand and wrist conditions

Given the differences in injury scope for the WHO Brief ICF Core Set and Goldhahn et al (2014) core set of domains for distal radius fractures, the ICHOM standard set for hand and wrist conditions offered the most direct comparison to the SO-HANDI COS. The ICHOM standard set was produced by a working group consisting of 22 hand surgeons, therapists and researchers. They proceeded with determining how the various outcome domains should be measured, but that is outside the scope of comparison presented here.

7.3.2.1 Generation of outcome domains

Generation of the ICHOM standard set domains was achieved solely through systematic review of the literature and discussion by the working group. They developed 18 outcome domains as compared to the 74 generated through the SO-HANDI systematic review process. The difference in numbers is explained by different approaches for qualitative analysis of the data. It is not clear what framework was used by the ICHOM working group but their output was a longlist of domains generally broader than the SO-HANDI COS, the latter being based on the more granular WHO ICF framework.

Beyond the systematic review, I conducted an extensive exploration of the patient perspective and domains of relevance to patients. Through thematic analysis I developed 35 outcome domains based solely on patient interviews and focus groups – the ICHOM standard set had no equivalent process.

Despite a lack of patient input at the outcome domain generation step, the ICHOM working group found that 95% of 1,060 surveyed patients indicated that the longlist they developed 'broadly captured' all the important outcomes for adults with hand and wrist conditions (Wouters et al. 2021). Their survey result was not surprising given my experience when synthesising the outcome domains generated by systematic review and qualitative work with patients. The Delphi outcome domains I developed showed large overlap, with only six domains derived from just one source:

- From the systematic review only – 'stability of joints' and 'thinking and memory'
- From the patient interviews/focus groups only – 'keyboard and mouse / touchscreen use', 'looking after dependents within household', 'emotional / mood impact to friends and family' and 'added expenses due to injury'

Of these six domains, the only ones prioritised as 'consensus in' through the Delphi survey were 'stability of joints' and 'keyboard and mouse / touchscreen use'. However, it is important to consider that patients being asked whether a longlist of domains 'broadly captures' the important outcomes for their hand and wrist conditions is not the same as finding out whether the list of domains was comprehensive. Even if the goal was to determine whether the longlist covers all domains that are 'important' enough to potentially be in the final core set, the phrasing chosen may have given misleading results for several reasons:

- The threshold of ‘important’ could vary between patients
- ‘Broadly captures’ was an ambiguous term open to interpretation by survey respondents
- The concept of ‘outcomes’ could be difficult to grasp for those unfamiliar with it, so relying on them to judge whether any additional ‘important’ domains are missing relied on a clear explanation of background information

Analysis of patient interviews and focus groups generated additional domains, so this was a useful step in order to achieve a comprehensive longlist of outcome domains as the starting point for the SO-HANDI consensus prioritisation work. It also aligned with COMET recommendations for COS development (Williamson et al. 2017). Perhaps most importantly though, by completing the analysis of patient interviews and focus groups I was able to generate wording for the domains and their descriptors which captured the patient perspective – an advantage of qualitative research which has been noted by others (Keeley et al. 2016). This subsequently impacted on the prioritisation phase – when Delphi survey respondents were judging the importance of outcome domains in the longlist, they needed clarity on the meaning of the various domains. As an example of the difference in outcome domain and descriptor wording, compare the SO-HANDI COS domain of ‘appearance’ with the ICHOM standard set domain of ‘aesthetics’:

> *“Appearance – how the hand / wrist looks after treatment is complete. This includes various aspects such as how one feels about the appearance themselves, concern about how others view the injured hand / wrist, etc, e.g. change in shape of injured hand / wrist, scars, swelling”* (SO-HANDI COS)

> *“Aesthetics – the branch of philosophy dealing with the nature of the beautiful. It includes beauty, aesthetic experience, aesthetic judgment, aesthetic aspects of medicine, etc.”* (ICHOM standard set, Wouters et al., 2021)

The ICHOM working group’s domain definitions were drawn from standard professional/academic bodies and were not in lay language. The same domain definitions were presented to patients participating in the evaluative survey (personal correspondence from members of the ICHOM working group).

7.3.2.2 Prioritisation of outcome domains

The prioritisation step in the ICHOM standard set was completed by the working group, with no patient involvement. This runs counter to the COMET Initiative guidance of achieving a consensus for the domains core to all key stakeholders, which should include patients (Williamson et al. 2017). To mitigate the impact of a lack of patient input at the domain selection step, the ICHOM working group surveyed patients to find out which domains they felt were most important from amongst their longlist of 18. For the finger track 260 patients responded, for the thumb track there were 191 patients, and for the wrist track there were 375 patients. However, it is not clear how many had direct experience of hand fractures and joint injuries as per the scope of the SO-HANDI COS, as compared to other hand or wrist conditions. For the thumb and wrist tracks, patients’ prioritisations aligned with those of the working group for 6 of 9 domains and 7 of 9 domains, respectively. For the finger track however, this dropped to an overlap of just 4 of 9 domains. One of the examples of domains prioritised by patients and not the working group was ‘fine hand use’, the stated definition of which confirmed a match to the SO-HANDI domain ‘fine hand use / dexterity’. However, ICHOM did not alter their standard set based on the input of patients in this survey.

In contrast, the priorities of the patient perspective were particularly emphasised in the SO-HANDI methodology by assigning their stakeholder group the highest voting power of any in the consensus meeting. I do not suggest that this was the reason that the SO-HANDI COS included the ‘fine hand use / dexterity’ domain, particularly given that many clinicians rated it as ‘very important’ throughout the Delphi survey too. Nevertheless, it seems intuitive that in order to prioritise, “outcome domains that capture the patient’s perspective” (Wouters et al. 2021), it is crucial to involve patients in the selection of the final set of domains.

7.4 Potential outcome measurement instruments for the COS domains, future research and dissemination of the COS

This thesis has aimed to deliver a COS of what domains to select in all future clinical research on hand fractures and joint injuries. To achieve standardisation of outcomes in clinical research it is also necessary to define how to measure these COS domains, which also includes defining when to measure them. This would therefore be the next phase of COS development and constitutes a substantial body of work. In the interim, to guide those selecting outcomes for future research studies, I now consider potential candidate outcome measurement instruments.

7.4.1 Potential outcome measurement instruments for COS domains

A number of studies have examined the validity of some of the existing upper limb PROMs commonly used in research. Dacombe et al (2016) reported that two PROMs had evidence based on hand and wrist trauma populations: the Disabilities of the Arm, Shoulder and Hand (DASH) score and the Patient-Rated Wrist Evaluation (PRWE). However, there is concern that the intended scope of these measures is not specific to the SO-HANDI population. The DASH assesses the upper limb more broadly, with questions addressing problems with the shoulder and elbow as well as the hand. The PRWE has been used in a wrist trauma population, but not in patients with hand injuries.

Since that 2016 review though, the COSMIN Initiative (briefly described in Chapter 1) has produced a checklist to guide evaluation of measurement instruments on the basis of set criteria of reliability, validity and responsiveness. Wormald et al (2019) used the COSMIN approach and focused on the properties of the six PROMs most commonly used in clinical research on hand conditions, while excluding any wrist-specific PROMs. Although this included evidence from trauma and non-trauma hand condition populations, their systematic COSMIN-driven assessment of PROMs offers reasonable grounds to select interim suggestions on how to measure some of the COS domains. They concluded there is incomplete evidence to support use of one specific PROM over another in hand surgery research when held to contemporary PROM standards, and therefore could not make a definitive recommendation. DASH, QuickDASH (an abbreviated version of DASH) and the Michigan Hand Questionnaire (MHQ) appeared to have the most published data evaluating their clinimetric properties overall, though this evidence was far from comprehensive across all the COSMIN domains. Very limited evidence for most COSMIN domains was also found for PROMs in the setting of adults with distal radius fractures (British Orthopaedic Association and British Society for Surgery of the Hand 2018).

Wormald et al (2019) also analysed the PROMs to determine which broad domains were being covered – it appears that all three of the most evidenced PROMs would offer significant coverage of the SO-HANDI COS domains. Closer inspection of specific items in these PROMs reveals that while there were questions relating to ‘fine hand use / dexterity’ as per the SO-HANDI COS domain descriptor and to ‘pain / discomfort during activity’, the remaining domains had varying coverage across the PROMs. Figure 7.1 reflects a summary of my views on this coverage based on the specific PROM items and how they relate to each of the domains. Specific items from each PROM are quoted to help convey the reasoning for the judgments reached.

Figure 7.1 Summary of interpretation of SO-HANDI COS domain coverage by three PROMs with most evidence supporting use in hand surgery research

COS Domain	DASH	QuickDASH	MHQ
Self-hygiene / personal care	+ “Rate ability to: wash or blow dry your hair; wash your back”	–/+ “Rate ability to wash your back”	–/+ “Using both hands, how difficult is it to wash your hair”
Pain / discomfort during activity	+ “Rate severity of arm, shoulder or hand pain when you performed any specific activity”	–	+ “How often did pain interfere with daily activities (e.g. eating, bathing)?”
Pain / discomfort during rest	–/+ “Rate severity of arm, shoulder or hand pain” “How much difficulty sleeping because of pain in arm, shoulder or hand”	–/+ “Rate severity of arm, shoulder or hand pain” “How much difficulty sleeping because of pain in arm, shoulder or hand”	–/+ “How often did you have pain?” “Severity of pain” “How often did pain interfere with sleep?”
Fine hand use / dexterity	+ “Rate ability to: write; turn a key”	–	+ “How difficult is it to: pick up a coin; turn a key in a lock”
Return to usual work / job	++ DASH work module: “Any difficulty in: usual technique at work; doing usual work because of pain; doing work as well as you would like; spending usual amount of time doing work”	++ DASH work module: “Any difficulty in: usual technique at work; doing usual work because of pain; doing work as well as you would like; spending usual amount of time doing work”	+ “In your normal work (including housework and school work) because of hand/wrist problems how often: unable to do work; shorten work day, take it easy at work, accomplish less; take longer?”
Range of movement	–/+ “Rate the severity of stiffness in your arm, shoulder or hand”	–	–/+ “How well do your fingers move?” “How well does your wrist move?”
Patient satisfaction with outcome / result	–	–	–/+ “Satisfaction with: overall function; finger motion; wrist motion; hand strength, hand pain; hand sensation”

Key:
– indicates no coverage of domain
+ indicates partial coverage focused on domain
–/+ indicates some coverage but inadequate in nature
++ indicates good coverage focused on domain

Several domains covered by DASH are not represented in the abbreviated QuickDASH. Despite some coverage of most domains by both DASH and MHQ, they appear to do so in ways which only partially cover key aspects of the domain. For example, they both include an item relating to 'pain / discomfort during activity', but this does not explore the domain well enough. They lack details, such as pain during light versus heavy activity, enquiring about pain severity and frequency, etc. Meanwhile, 'pain / discomfort during rest' was queried tangentially by asking about interference of sleep due to pain – though not directly asking about pain at rest, this appears to have some degree of domain coverage.

The 'range of movement' domain is also only tangentially covered by asking patients to rate their own stiffness or movement. This is not a shortcoming of PROM design, but rather an inherent issue with this domain within the setting of the SO-HANDI COS – based on the interview/focus group work with patients, as well as comments during the consensus meeting, it is clear stakeholders desired data on this domain as reflected by performance-based measurement. Hence, a PROM alone would not satisfactorily cover the domain; traditional measurements, such as through use of goniometers, would likely also be required.

Wormald et al (2019) reported that the MHQ covered the domain of 'satisfaction'. On looking at the specific items relating to this domain, I would argue that the coverage is not adequate in relation to the SO-HANDI COS domain of 'patient satisfaction with outcome / result'. The MHQ items specifically highlight satisfaction with certain parameters (overall function, motion, strength, pain and sensation). While perhaps considered reasonable to build up a picture of 'overall satisfaction with the result' based on these criteria, it does constrain responses – patients may have a different opinion on the overall result than is conveyed by these narrow parameters.

The one domain which appears to have good coverage is ‘return to usual work / job’, through the optional work module for DASH. I marked the equivalent set of MHQ items down slightly because the question stem indicates that when answering patients should take into account housework and schoolwork, which was not the focus of the domain in the SO-HANDI COS.

ICHOM highlighted limitations of existing PROMs for their ‘satisfaction with treatment results’ and ‘return to work’ domains, creating their own questionnaires to address these two areas. The former featured:

- A broader question directly asking patients to rate their satisfaction with the treatment result (rather than confining the rating to very specific aspects such as function, pain, etc)
- Asking whether patients, in hindsight, would have opted for the same treatment based on their experience
- Asking whether patients would recommend the treatment to their friends and family

They had numerous items in a questionnaire on return to work, but essentially collected patient responses on employment status, hours of normal work, whether modified duties/hours were being worked, when work was first resumed and at what point normal duties were resumed and confidence in ability to return to work. The ICHOM questionnaires were developed by the working group and have not been assessed for reliability or validity so far (Wouters et al. 2021).

A broader point is that although DASH and MHQ seem to have items which cover the SO-HANDI COS domains to varying degrees, they are scored in specific ways:

- The validated output of DASH is a single overall score, and therefore several items bearing no relevance to the SO-HANDI COS domains would also impact on the overall score. One cannot simply extract the items specifically relevant to the COS
- MHQ has both total scoring and separate scoring of the six scales which comprise it, but not all items even within each scale are necessarily relevant to the COS.

7.4.2 Future research

According to COSMIN criteria for selection of outcome measurement instruments for a COS, there needs to be sufficient evidence that prospective measurement tools not only address the COS domains but do so in the specific adult hand fracture and joint injury population (Prinsen et al. 2016). Hence, there remains a great deal of work to ensure strongly evidenced and validated measurement tools are used to capture the key domains identified in the SO-HANDI COS. The descriptors developed offer clarification on the underlying construct of the domains, which is a benefit when identifying how best to measure them.

Given the standard set by COSMIN, I would recommend a key stakeholder consultation based on evidence such as that summarised by Wormald et al (2019) as a starting point, to decide whether existing instruments are adequate. Where they are not, further work would involve the de novo development of appropriate tools. This is a challenging and lengthy process. For example, although ICHOM developed a tool which on the surface appears to address ‘patient satisfaction with outcome / result’, the concept of measuring satisfaction with the result of treatment has raised controversy, pointing to the many potential confounding influences that could invalidate such a measure (Ring and Leopold 2015) and indicating the complex nature of this domain is not amenable to a single measure such as a visual analogue scale (Graham et al. 2015).

Beyond the outcome measurements instruments, appropriate time points of outcome assessment should also be determined. This again would need input from all key stakeholders. Finally, the completed work would need to undergo translation and validation.

7.4.3 Dissemination of COS

Uptake of a COS is required in order to reap the benefits of developing one. Dissemination is a fundamental aspect of promoting uptake. COSs are now endorsed by major funders of clinical research (e.g. National Institute for Health Research, Versus Arthritis, Association of Medical Research Charities, Health Research Board, Horizon 2020) (COMET Initiative 2022c), trialists (Standard Protocol Items: Recommendations for Interventional Trials guidelines (SPIRIT Group 2022)), Cochrane (Cochrane Community Blog 2022), NICE (National Institute for Health and Care Excellence, 2015), some journal editors (e.g. Core Outcomes in Women's and Newborn Health Initiative (CROWN Initiative 2019)) and patients and the public (COMET People and Patient Participation, Involvement and Engagement working group (COMET Initiative 2022b)). This incentivises researchers to use a relevant COS when available.

Through publication and presentation, I aim to raise awareness of the SO-HANDI COS domains. Furthermore, I shall contact relevant professional societies for hand surgeons and therapists about the COS – endorsement by such bodies will aid dissemination and increase the likelihood of outcome selection in future clinical research being guided by the COS domains.

7.5 Conclusion

In conclusion, this thesis defines what outcome domains to measure as a COS on hand fractures and joint injuries in adults. In doing so, I have persevered to explore, prioritise and incorporate the perspective of patients. Beyond an initial exercise to set the scope and a systematic review of treatment outcomes, patient participants were meaningfully involved at all steps. A qualitative exploration of the patient journey through interviews and focus groups enabled generation of outcome domains considered relevant by this key stakeholder group. Consensus prioritisation work through the Delphi survey and a final consensus meeting achieved agreement on the seven COS domains, with patients comprising 44% (12/27) of the voting power at the meeting.

This thesis also identified key COS development methodological considerations for future COS developers. There may be a role for coinciding exploration of when outcomes are to be measured together with the prioritisation step of what domains should be included in the COS, rather than relegating the former to the 'how to measure' phase. The interview and focus group work with patients not only provided novel domains compared to the systematic review but also facilitated the wording for domains and descriptors in language familiar to the patient stakeholder group. Finally, the specific format of the online consensus meeting appeared to work well for the SO-HANDI COS development and might be deemed a useful foundation for the format of other COS developers' online meetings.

Important work remains to define how to measure the SO-HANDI COS domains. Nevertheless, the work presented here is a critical step in the standardisation of meaningful, patient-centred outcome assessment selection in clinical research on these injuries.

References

Al-Jabri, S., Rowe, F.J. and Kirkham, J.J. 2021. Core outcome set for three ophthalmic conditions: a healthcare professional and patient consensus on core outcome sets for amblyopia, ocular motility and strabismus (COSAMS Study). *BMJ Open* 11(5), p. e042403. Available at: http://bmjopen.bmj.com/content/11/5/e042403.abstract.

Aligato, M.F., Endoma, V., Wachinger, J., Landicho-Guevarra, J., Bravo, T.A., Guevarra, J.R., Landicho, J., McMahon, S.A. and Reñosa, M.D.C. 2021. 'Unfocused groups': lessons learnt amid remote focus groups in the Philippines. *Family Medicine and Community Health* 9(Suppl 1), p. e001098. Available at: http://fmch.bmj.com/content/9/Suppl_1/e001098.abstract.

Alkhaffaf, B., Metryka, A., Blazeby, J.M., Glenny, A.-M., Adeyeye, A., Costa, P.M., Diez del Val, I., Gisbertz, S.S., Guner, A., Law, S., Lee, H.-J., Li, Z., Nakada, K., Reim, D., Vorwald, P., Baiocchi, G.L., Allum, W., Chaudry, M.A., Griffiths, E.A., Williamson, P.R., Bruce, I.A. and Group, on behalf of the G.I.W. 2021. Core outcome set for surgical trials in gastric cancer (GASTROS study): international patient and healthcare professional consensus. *British Journal of Surgery* . Available at: https://doi.org/10.1093/bjs/znab192.

Alwin, D.F. and Krosnick, J.O.N.A. 1991. The Reliability of Survey Attitude Measurement: The Influence of Question and Respondent Attributes. *Sociological Methods & Research* 20(1), pp. 139–181. Available at: https://doi.org/10.1177/0049124191020001005.

Ammann, B., Satink, T. and Andresen, M. 2012. Experiencing occupations with chronic hand disability: narratives of hand-injured adults. *Hand Therapy* 17(4), pp. 87–94. Available at: https://doi.org/10.1177/1758998312471253.

Andreasson, I., Kjellby-Wendt, G., Fagevik-Olsen, M., Karlsson, J. and Carlsson, G. 2019. Life has become troublesome - my wrist bothers me around the clock: an interview study relating to daily life with a malunited distal radius fracture. *Disabil Rehabil* , pp. 1–7. Available at: internal-pdf://112.99.116.90/Andreasson 2018.pdf.

Anhang Price, R., Elliott, M.N., Zaslavsky, A.M., Hays, R.D., Lehrman, W.G., Rybowski, L., Edgman-Levitan, S. and Cleary, P.D. 2014. Examining the role of patient experience surveys in measuring health care quality. *Medical care research and review : MCRR* 71(5), pp. 522–

554. Available at: https://pubmed.ncbi.nlm.nih.gov/25027409.

Baker, S.E. and Edwards, R. eds. 2012. *How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research*. National Centre for Research Methods. Available at: http://eprints.ncrm.ac.uk/2273/4/how_many_interviews.pdf.

Ball, C., Pratt, A.L. and Nanchahal, J. 2013. Optimal functional outcome measures for assessing treatment for Dupuytren's disease: a systematic review and recommendations for future practice. *BMC Musculoskeletal Disorders* 14(1), p. 131. Available at: https://doi.org/10.1186/1471-2474-14-131.

Bamford, R. and Walker, D.-M. 2010. A qualitative investigation into the rehabilitation experience of patients following wrist fracture. *Hand Therapy* 15(3), pp. 54–61. Available at: https://doi.org/10.1258/ht.2010.010013.

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

Beaton, D.E., Wright, J.G., Katz, J.N. and Group, U.E.C. 2005. Development of the QuickDASH: comparison of three item-reduction approaches. *JBJS* 87(5), pp. 1038–1046.

Begg, C.B. and Berlin, J.A. 1989. Publication bias and dissemination of clinical research. *Journal of the National Cancer Institute* 81(2), pp. 107–115. Available at: https://www.ncbi.nlm.nih.gov/pubmed/2642556.

Bergner, M. 1985. Measurement of Health Status. *Medical Care* 23(5), pp. 696–704. Available at: http://www.jstor.org/stable/3764999.

Bialocerkowski, A.E. 2002. Difficulties associated with wrist disorders--a qualitative study. *Clin Rehabil* 16(4), pp. 429–440. Available at: https://www.ncbi.nlm.nih.gov/pubmed/12061478.

Biggane, A.M., Brading, L., Ravaud, P., Young, B. and Williamson, P.R. 2018. Survey indicated that core outcome set development is increasingly including patients, being conducted internationally and using Delphi surveys. *Trials* 19(1), p. 113. Available at: https://www.ncbi.nlm.nih.gov/pubmed/29454368.

Black, N. 2013. Patient reported outcome measures could help transform healthcare. *BMJ : British Medical Journal* 346, p. f167. Available at: http://www.bmj.com/content/346/bmj.f167.abstract.

Boers, M., Beaton, D.E., Shea, B.J., Maxwell, L.J., Bartlett, S.J., Bingham, C.O., Conaghan, P.G., D'Agostino, M.A., de Wit, M.P., Gossec, L., March, L., Simon, L.S., Singh, J.A., Strand, V., Wells, G.A. and Tugwell, P. 2019. OMERACT Filter 2.1: Elaboration of the Conceptual Framework for Outcome Measurement in Health Intervention Studies. *The Journal of Rheumatology* 46(8), pp. 1021–1027. Available at: http://www.jrheum.org/content/46/8/1021.abstract.

Boers, M., Brooks, P., Strand, C. V. and Tugwell, P. 1998. The OMERACT filter for Outcome Measures in Rheumatology. *The Journal of rheumatology* 25(2), pp. 198–199.

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

Braun, V. and Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3(2), pp. 77–101. Available at: https://www.tandfonline.com/doi/abs/10.1191/1478088706qp063oa.

Brealey, S., Scally, A., Hahn, S., Thomas, N., Godfrey, C. and Coomarasamy, A. 2005. Accuracy of radiographer plain radiograph reporting in clinical practice: a meta-analysis. *Clinical Radiology* 60(2), pp. 232–241. Available at: https://www.sciencedirect.com/science/article/pii/S0009926004002703.

British Orthopaedic Association and British Society for Surgery of the Hand 2018. *Best practice for management of distal radial fractures*. Available at: https://www.bssh.ac.uk/_userfiles/pages/files/professionals/Radius/Blue Book DRF Final Document.pdf [Accessed: 20 May 2022].

Brookes, S.T., Macefield, R.C., Williamson, P.R., McNair, A.G., Potter, S., Blencowe, N.S., Strong, S. and Blazeby, J.M. 2016. Three nested randomized controlled trials of peer-only or multiple stakeholder group feedback within Delphi surveys during core outcome and information set development. *Trials* 17(1), p. 409. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27534622.

Bryman, A. 2004. *Social research methods*. 2nd ed. Oxford university press.

Buskov, L., Abild, A., Christensen, A., Holm, O., Hansen, C. and Christensen, H. 2013. Radiographers and trainee radiologists reporting accident radiographs: A comparative plain film-reading performance study. *Clinical Radiology* 68(1), pp. 55–58. Available at: https://www.sciencedirect.com/science/article/pii/S0009926012003534.

Chalmers, I. and Glasziou, P. 2009. Avoidable waste in the production and reporting of research evidence. *The Lancet* 374(9683), pp. 86–89. Available at: https://doi.org/10.1016/S0140-6736(09)60329-9.

Chan, A.-W., Hróbjartsson, A., Haahr, M.T., Gøtzsche, P.C. and Altman, D.G. 2004. Empirical Evidence for Selective Reporting of Outcomes in Randomized TrialsComparison of Protocols to Published Articles. *JAMA* 291(20), pp. 2457–2465. Available at: https://doi.org/10.1001/jama.291.20.2457.

Chan, J. and Spencer, J. 2004. Adaptation to hand injury: an evolving experience. *Am J Occup Ther* 58(2), pp. 128–139. Available at: internal-pdf://112.182.159.161/Chan-2004-Adaptation to hand injury_ an evolvi.pdf.

Chan, J. and Spencer, J. 2005. Contrasting perspectives on pain following hand injury. *J Hand Ther* 18(4), pp. 429–35, quiz 436. Available at: internal-pdf://107.225.159.224/Chan-2005-Contrasting perspectives on pain fol.pdf.

Cheng, Y.H. 2008. Against the odds or odds-on chance: pathways of coming back among hand-injured workers in the Chinese context. *Qual Health Res* 18(6), pp. 843–852. Available at: https://www.ncbi.nlm.nih.gov/pubmed/18503025.

Chevance, A., Tran, V.-T. and Ravaud, P. 2020. Controversy and Debate Series on Core Outcome Sets. Paper 1: Improving the generalizability and credibility of core outcome sets (COS) by a large and international participation of diverse stakeholders. *Journal of Clinical*

Epidemiology 125, pp. 206-212.e1. Available at: https://doi.org/10.1016/j.jclinepi.2020.01.004.

Chung, K.C., Burns, P.B. and Davis Sears, E. 2006. Outcomes Research in Hand Surgery: Where Have We Been and Where Should We Go? *The Journal of Hand Surgery* 31(8), pp. 1373–1379. Available at: http://www.sciencedirect.com/science/article/pii/S0363502306006964.

Cieza, A., Geyh, S., Chatterji, S., Kostanjsek, N., Üstün, B. and Stucki, G. 2005. ICF linking rules: An update based on lessons learned. *Journal of Rehabilitation Medicine* 37(4), pp. 212–218. doi: 10.1080/16501970510040263.

Clarke, M. 2007. Standardising outcomes for clinical trials and systematic reviews. *Trials* 8(1), p. 39. Available at: https://doi.org/10.1186/1745-6215-8-39.

Clarke, M. and Williamson, P.R. 2016. Core outcome sets and systematic reviews. *Systematic reviews* 5, p. 11. Available at: https://www.ncbi.nlm.nih.gov/pubmed/26792080.

Cochrane Community Blog 2022. *Cochrane and COMET: working together to improve core outcome sets*. Available at: https://community.cochrane.org/news/cochrane-and-comet-working-together-improve-core-outcome-sets.

Coenen, M., Kus, S., Rudolf, K.D., Muller, G., Berno, S., Dereskewitz, C. and MacDermid, J. 2013. Do patient-reported outcome measures capture functioning aspects and environmental factors important to individuals with injuries or disorders of the hand? *J Hand Ther* 26(4), pp. 332–42; quiz 342. Available at: internal-pdf://205.143.159.255/Coenen-2013-Do patient-reported outcome measur.pdf.

COMET Initiative 2020. *DelphiManager*. Available at: https://www.comet-initiative.org/delphimanager/.

COMET Initiative 2022a. *COMET Database*. Available at: http://www.comet-initiative.org/studies/search.

COMET Initiative 2022b. *COMET PoPPIE Working Group*. Available at: http://www.comet-initiative.org/ppi/poppie.

COMET Initiative 2022c. *COS Uptake and Endorsement*. Available at: http://www.comet-initiative.org/cosuptake [Accessed: 20 May 2022].

Copeland, A., Gallo, L., Weber, C., Moltaji, S., Gallo, M., Murphy, J., Axelrod, D. and Thoma, A. 2020. Reporting Outcomes and Outcome Measures in Thumb Carpometacarpal Joint Osteoarthritis: A Systematic Review. *Journal of Hand Surgery* . Available at: https://doi.org/10.1016/j.jhsa.2020.05.024.

COSMIN Initiative 2022. *COSMIN Initiative*. Available at: https://www.cosmin.nl/about/.

Court-Brown, C.M., Biant, L., Bugler, K.E. and McQueen, M.M. 2014. Changing epidemiology of adult fractures in Scotland. *Scottish Medical Journal* 59(1), pp. 30–34. Available at: https://doi.org/10.1177/0036933013518148.

CROWN Initiative 2019. *About CROWN*. Available at: http://www.crown-initiative.org/14-2/about/.

Dacombe, P.J., Amirfeyz, R. and Davis, T. 2016. Patient-Reported Outcome Measures for Hand and Wrist Trauma. *Hand* 11(1), pp. 11–21. Available at: https://doi.org/10.1177/1558944715614855.

Dalkey, N. 1969. An experimental study of group opinion: The Delphi method. *Futures* 1(5), pp. 408–426. Available at: https://www.sciencedirect.com/science/article/pii/S001632876980025X.

Dalkey, N. and Helmer, O. 1963. An experimental application of the Delphi method to the use of experts. *Management science* 9(3), pp. 458–467.

Damhuis, S.E., Bloomfield, F.H., Khalil, A., Daly, M., Ganzevoort, W. and Gordijn, S.J. 2021. A Core Outcome Set and minimum reporting set for intervention studies in growth restriction in the NEwbOrN: the COSNEON study. *Pediatric Research* 89(6), pp. 1380–1385. Available at: https://doi.org/10.1038/s41390-020-01119-5.

Daniels, N., Gillen, P., Casson, K. and Wilson, I. 2019. STEER: Factors to Consider When Designing Online Focus Groups Using Audiovisual Technology in Health Research. *International Journal of Qualitative Methods* 18, p. 1609406919885786. Available at: https://doi.org/10.1177/1609406919885786.

Davis, K., Gorst, S.L., Harman, N., Smith, V., Gargon, E., Altman, D.G., Blazeby, J.M., Clarke, M., Tunis, S. and Williamson, P.R. 2018. Choosing important health outcomes for comparative effectiveness research: An updated systematic review and involvement of low and middle income countries. *PLOS ONE* 13(2), p. e0190695. Available at: https://doi.org/10.1371/journal.pone.0190695.

De Meyer, D., Kottner, J., Beele, H., Schmitt, J., Lange, T., Van Hecke, A., Verhaeghe, S. and Beeckman, D. 2019. Delphi procedure in core outcome set development: rating scale and consensus criteria determined outcome selection. *Journal of Clinical Epidemiology* 111, pp. 23–31. Available at: http://www.sciencedirect.com/science/article/pii/S0895435618310874.

de Putter, C.E., Selles, R.W., Polinder, S., Panneman, M.J.M., Hovius, S.E.R. and Van Beeck, E.F. 2012. Economic impact of hand and wrist injuries: Health-care costs and productivity costs in a population-based study. *Journal of Bone and Joint Surgery American* 94(9), p. e56(1-7). Available at: https://www.ncbi.nlm.nih.gov/pubmed/22552678.

Deshmukh, S., Karantana, A., Montgomery, A., Leighton, P., Jerosch-Herold, C., Rodrigues, J., Trickett, R., Costa, M. and Griffin, X. 2022. *COMET Database - Developing a core outcome set for hand fractures and joint injuries in adults*. Available at: https://comet-initiative.org/studies/details/1237 [Accessed: 20 May 2022].

Deshmukh, S.R., Mousoulis, C., Marson, B., Grindlay, D., Leighton, P., Montgomery, A. and Karantana, A. 2019. *Developing a core outcome set for hand fractures and joint injuries in adults: a systematic review of treatment outcomes*. Available at: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42019126299.

Dias, J., Brealey, S., Cook, L., Fairhurst, C., Hinde, S., Leighton, P., Choudhary, S., Costa, M., Hewitt, C. and Hodgson, S. 2020. Surgical fixation compared with cast immobilisation for adults with a bicortical fracture of the scaphoid waist: the SWIFFT RCT. *Health technology assessment (Winchester, England)* 24(52), p. 1. doi: 10.3310/hta24520.

Djulbegovic, B. and Guyatt, G.H. 2017. Progress in evidence-based medicine: a quarter century on. *The Lancet* 390(10092), pp. 415–423. Available at: http://www.sciencedirect.com/science/article/pii/S0140673616315926.

Dodd, S., Clarke, M., Becker, L., Mavergames, C., Fish, R. and Williamson, P.R. 2018. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *Journal of Clinical Epidemiology* 96, pp. 84–92. Available at: https://www.ncbi.nlm.nih.gov/pubmed/29288712.

Dos Santos Marques, I.C., Theiss, L.M., Johnson, C.Y., McLin, E., Ruf, B.A., Vickers, S.M., Fouad, M.N., Scarinci, I.C. and Chu, D.I. 2021. Implementation of virtual focus groups for qualitative data collection in a global pandemic. *American journal of surgery* 221(5), pp. 918–922. Available at: https://pubmed.ncbi.nlm.nih.gov/33070983.

Dwan, K., Altman, D.G., Arnaiz, J.A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E., Easterbrook, P.J., Von Elm, E., Gamble, C., Ghersi, D., Ioannidis, J.P.A., Simes, J. and Williamson, P.R. 2008. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS one* 3(8), pp. e3081–e3081. Available at: https://www.ncbi.nlm.nih.gov/pubmed/18769481.

Dwan, K., Altman, D.G., Cresswell, L., Blundell, M., Gamble, C.L. and Williamson, P.R. 2011. Comparison of protocols and registry entries to published reports for randomised controlled trials. *The Cochrane database of systematic reviews* 2011(1), p. MR000031. doi: 10.1002/14651858.MR000031.pub2.

Farrell, B., Kenyon, S. and Shakur, H. 2010. Managing clinical trials. *Trials* 11(1), p. 78. Available at: https://doi.org/10.1186/1745-6215-11-78.

Feehan, L.M. and Sheps, S.B. 2006. Incidence and demographics of hand fractures in British Columbia, Canada: a population-based study. *The Journal of hand surgery* 31(7), pp. 1068–1074. doi: 10.1016/j.jhsa.2006.06.006.

Ferrans, C.E., Zerwic, J.J., Wilbur, J.E. and Larson, J.L. 2005. Conceptual model of health-related quality of life. *Journal of nursing scholarship : an official publication of Sigma Theta Tau International Honor Society of Nursing* 37(4), pp. 336–342. doi: 10.1111/j.1547-5069.2005.00058.x.

Fish, R., MacLennan, S., Alkhaffaf, B. and Williamson, P.R. 2020. “Vicarious thinking” was a key driver of score change in Delphi surveys for COS development and is facilitated by feedback of results. *Journal of clinical epidemiology* 128, pp. 118–129. doi: 10.1016/j.jclinepi.2020.09.028.

Fitch, K., Bernstein, S., Aguilar, M.D., Burnand, B. and Lacalle, J. 2001. The RAND/UCLA Appropriateness Method User’s Manual.

Flick, U. 2018. *An introduction to qualitative research*. 6th ed. SAGE Publications.

Francis, J.J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M.P. and Grimshaw, J.M. 2010. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology & Health* 25(10), pp. 1229–1245. Available at: https://doi.org/10.1080/08870440903194015.

Gargon, E., Crew, R., Burnside, G. and Williamson, P.R. 2019a. Higher number of items associated with significantly lower response rates in COS Delphi surveys. *Journal of Clinical Epidemiology* 108, pp. 110–120. Available at: http://www.sciencedirect.com/science/article/pii/S0895435618307765.

Gargon, E., Gorst, S.L., Matvienko-Sikar, K. and Williamson, P.R. 2021. Choosing important health outcomes for comparative effectiveness research: 6th annual update to a systematic review of core outcome sets for research. *PLOS ONE* 16(1), p. e0244878. Available at: https://doi.org/10.1371/journal.pone.0244878.

Gargon, E., Gorst, S.L. and Williamson, P.R. 2019b. Choosing important health outcomes for comparative effectiveness research: 5th annual update to a systematic review of core outcome sets for research. *PLOS ONE* 14(12), p. e0225980. Available at: https://doi.org/10.1371/journal.pone.0225980.

Gargon, E., Williamson, P.R., Altman, D.G., Blazeby, J.M., Tunis, S. and Clarke, M. 2017. The COMET Initiative database: Progress and activities update (2015). *Trials* 18(1), p. 54. Available at: https://pubmed.ncbi.nlm.nih.gov/28159003.

Garratt, A., Schmidt, L., Mackintosh, A. and Fitzpatrick, R. 2002. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 324(7351), p. 1417. Available at: http://www.bmj.com/content/324/7351/1417.1.abstract.

Givon, M.M. and Shapira, Z. 1984. Response to Rating Scales: A Theoretical Model and Its Application to the Number of Categories Problem. *Journal of Marketing Research* 21(4), pp. 410–419. Available at: http://www.jstor.org/stable/3151467.

Goldhahn, J., Beaton, D., Ladd, A., Macdermid, J. and Hoang-Kim, A. 2014. Recommendation for measuring clinical outcome in distal radius fractures: A core set of domains for standardized reporting in clinical practice and research. *Archives of Orthopaedic and Trauma Surgery* 134(2), pp. 197–205. Available at: https://www.ncbi.nlm.nih.gov/pubmed/23728832.

Gorst, S., Barrington, H., Brookes, S., Chalmers, J., Devane, D., Fledderus, A., Grosskleg, S., Hall, D., Harman, N., Hoffmann, C., Katiri, R., Maeso, R., Saldanha, I., Tong, A. and Williamson, P. 2021. *Online consensus meetings for COS development: issues to consider*. Available at: https://www.cometinitiative.org/Downloads/Issues to consider for online consensus meetings.pdf.

Gorst, S.L., Gargon, E., Clarke, M., Blazeby, J.M., Altman, D.G. and Williamson, P.R. 2016. Choosing Important Health Outcomes for Comparative Effectiveness Research: An Updated Review and User Survey. *PLOS ONE* 11(1), p. e0146444. Available at: https://doi.org/10.1371/journal.pone.0146444.

Graham, B., Green, A., James, M., Katz, J. and Swiontkowski, M. 2015. Measuring Patient Satisfaction in Orthopaedic Surgery. *JBJS* 97(1). Available at: https://journals.lww.com/jbjsjournal/Fulltext/2015/01070/Measuring_Patient_Satisfaction_in_Orthopaedic.12.aspx.

Guest, G., Bunce, A. and Johnson, L. 2006. How many interviews are enough? An experiment with data saturation and variability. *Field methods* 18(1), pp. 59–82.

Guest, G., Namey, E. and McKenna, K. 2016. How Many Focus Groups Are Enough? Building an Evidence Base for Nonprobability Sample Sizes. *Field Methods* 29(1), pp. 3–22. Available at: https://doi.org/10.1177/1525822X16639015.

Gustafsson, M., Persson, L.-O. and Amilon, A. 2000. A qualitative study of stress factors in the early stage of acute traumatic hand injury. *Journal of Advanced Nursing* 32(6), pp. 1333–1340. Available at: https://doi.org/10.1046/j.1365-2648.2000.01630.x.

Gustafsson, M., Persson, L.O. and Amilon, A. 2002. A qualitative study of coping in the early stage of acute traumatic hand injury. *J Clin Nurs* 11(5), pp. 594–602. Available at: internal-pdf://0065576975/Gustafsson-2002-A qualitative study of coping.pdf.

Guyatt, G.H., Oxman, A.D., Kunz, R., Atkins, D., Brozek, J., Vist, G., Alderson, P., Glasziou, P., Falck-Ytter, Y. and Schünemann, H.J. 2011. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of clinical epidemiology* 64(4), pp. 395–400. doi: 10.1016/j.jclinepi.2010.09.012.

Handoll, H.H. and Vaghela, M. V 2004. Interventions for treating mallet finger injuries.

Cochrane Database of Systematic Reviews (3), p. CD004574. Available at: https://www.ncbi.nlm.nih.gov/pubmed/15266538.

Hanna, P. 2012. Using internet technologies (such as Skype) as a research medium: a research note. *Qualitative Research* 12(2), pp. 239–242. Available at: https://doi.org/10.1177/1468794111426607.

Hasson, F., Keeney, S. and McKenna, H. 2000. Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing* 32(4), pp. 1008–1015. Available at: https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x.

Haywood, K.L., Griffin, X.L., Achten, J. and Costa, M.L. 2014. Developing a core outcome set for hip fracture trials. *The bone & joint journal* 96-B(8), pp. 1016–1023. doi: 10.1302/0301-620X.96B8.33766.

Heath, J., Williamson, H., Williams, L. and Harcourt, D. 2018. "It's just more personal": Using multiple methods of qualitative data collection to facilitate participation in research focusing on sensitive subjects. *Applied Nursing Research* 43, pp. 30–35. Available at: http://www.sciencedirect.com/science/article/pii/S0897189718302039.

Hoang-Kim, A., Ladd, A., Moore, D., Crisco, J., Weiss, A.-P., Marks, M. and Wolf, J. 2022. *COMET Database - Development of a Core set of Outcome Measures for the Thumb Carpometacarpal Osteoarthritis*. Available at: https://www.comet-initiative.org/Studies/Details/1137.

Holt, A. 2010. Using the telephone for narrative interviewing: a research note. *Qualitative Research* 10(1), pp. 113–121. Available at: https://doi.org/10.1177/1468794109348686.

HOME Initiative 2022. *Harmonising Outcome Measures for Eczema (HOME).* Available at: http://www.homeforeczema.org/about/about.aspx.

Hove, L.M. 1993. Fractures of the hand. Distribution and relative incidence. *Scandinavian journal of plastic and reconstructive surgery and hand surgery* 27(4), pp. 317–319.

Hudak, P.L., Amadio, P.C., Bombardier, C., Beaton, D., Cole, D., Davis, A., Hawker, G., Katz, J.N., Makela, M., Marx, R.G., Punnett, L. and Wright, J. 1996. Development of an upper extremity outcome measure: The DASH (disabilities of the arm, shoulder, and head). *American Journal of Industrial Medicine* 29(6), pp. 602–608. Available at:

https://doi.org/10.1002/(SICI)1097-0274(199606)29:6%3C602::AID-AJIM4%3E3.0.CO.

Humphrey-Murto, S., Varpio, L., Gonsalves, C. and Wood, T.J. 2017. Using consensus group methods such as Delphi and Nominal Group in medical education research*. *Medical Teacher* 39(1), pp. 14–19. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27841062.

Iacono, V. Lo, Symonds, P. and Brown, D.H.K. 2016. Skype as a Tool for Qualitative Research Interviews. *Sociological Research Online* 21(2), pp. 1–12. Available at: https://ideas.repec.org/a/sro/srosro/2015-177-2.html.

Idzerda, L., Rader, T., Tugwell, P. and Boers, M. 2014. Can We Decide Which Outcomes Should Be Measured in Every Clinical Trial? A Scoping Review of the Existing Conceptual Frameworks and Processes to Develop Core Outcome Sets. *The Journal of Rheumatology* 41(5), pp. 986 LP – 993. Available at: http://www.jrheum.org/content/41/5/986.abstract.

Ioannidis, J.P.A., Greenland, S., Hlatky, M.A., Khoury, M.J., Macleod, M.R., Moher, D., Schulz, K.F. and Tibshirani, R. 2014. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet* 383(9912), pp. 166–175.

James Lind Alliance 2017. *Common conditions affecting the hand and wrist priority setting partnership*. Available at: http://www.jla.nihr.ac.uk/priority-setting-partnerships/common-conditons-affecting-the-hand-and-wrist/downloads/JLA-Final-Summary.pdf [Accessed: 6 July 2019].

James Lind Alliance 2021. *The James Lind Alliance Guidebook*. Available at: https://www.jla.nihr.ac.uk/jla-guidebook/downloads/JLA-Guidebook-Version-10-March-2021.pdf.

Jerrhag, D., Englund, M., Karlsson, M.K. and Rosengren, B.E. 2017. Epidemiology and time trends of distal forearm fractures in adults - a study of 11.2 million person-years in Sweden. *BMC musculoskeletal disorders* 18(1), p. 240. Available at: https://pubmed.ncbi.nlm.nih.gov/28576135.

Jones, J. and Hunter, D. 1995. Consensus methods for medical and health services research. *BMJ (Clinical research ed.)* 311(7001), pp. 376–380. Available at: https://www.ncbi.nlm.nih.gov/pubmed/7640549.

Jones, J.E., Jones, L.L., Keeley, T.J.H., Calvert, M.J. and Mathers, J. 2017. A review of patient and carer participation and the use of qualitative research in the development of core outcome sets. *PLOS ONE* 12(3), p. e0172937. Available at: https://doi.org/10.1371/journal.pone.0172937.

Karl, J.W., Olson, P.R. and Rosenwasser, M.P. 2015. The epidemiology of upper extremity fractures in the United States, 2009. *Journal of Orthopaedic Trauma* 29(8), pp. e242–e244. Available at: https://www.ncbi.nlm.nih.gov/pubmed/25714441.

Keeley, T., Williamson, P., Callery, P., Jones, L.L., Mathers, J., Jones, J., Young, B. and Calvert, M. 2016. The use of qualitative methods to inform Delphi surveys in core outcome set development. *Trials* 17(1), p. 230. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27142835.

Kennedy, D., Furniss, D. and Chism-Balangue, T. 2022. *COMET Database - Development of a Core Outcome Measurement Set for the Evaluation of Scars in the Hand & Wrist (COSSH).* Available at: https://www.comet-initiative.org/Studies/Details/1503.

Kingston, G.A., Judd, D.J. and Gray, M.A. 2014. The experience of living with a traumatic hand injury in a rural and remote location: an interpretive phenomenological study. *Rural Remote Health* 14(3), p. 2764. Available at: https://www.ncbi.nlm.nih.gov/pubmed/25018127.

Kingston, G.A., Judd, J. and Gray, M.A. 2015. The experience of medical and rehabilitation intervention for traumatic hand injuries in rural and remote North Queensland: a qualitative study. *Disabil Rehabil* 37(5), pp. 423–429. Available at: https://www.ncbi.nlm.nih.gov/pubmed/24856789 https://www.tandfonline.com/doi/abs/10.3109/09638288.2014.923526.

Kirkham, J.J., Altman, D.G., Chan, A.W., Gamble, C., Dwan, K.M. and Williamson, P.R. 2018. Outcome reporting bias in trials: A methodological approach for assessment and adjustment in systematic reviews. *BMJ (Online)* 362, p. k3802. Available at: https://www.ncbi.nlm.nih.gov/pubmed/30266736.

Kirkham, J.J., Dwan, K.M., Altman, D.G., Gamble, C., Dodd, S., Smyth, R. and Williamson, P.R. 2010. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ (Online)* 340(7747), pp. 637–640. Available at: https://www.ncbi.nlm.nih.gov/pubmed/20156912.

Kirkham, J.J., Gargon, E., Clarke, M. and Williamson, P.R. 2013. Can a core outcome set improve the quality of systematic reviews? - a survey of the Co-ordinating Editors of Cochrane review groups. *Trials* 14(1), p. 21. Available at: https://doi.org/10.1186/1745-6215-14-21.

Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E., Moher, D., Schmitt, J., Tugwell, P., Tunis, S. and Williamson, P.R. 2016. Core Outcome Set–STAndards for Reporting: The COS-STAR Statement. *PLOS Medicine* 13(10), p. e1002148. Available at: https://doi.org/10.1371/journal.pmed.1002148.

Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Tunis, S., Williamson, P.R., Aldcroft, A., Bagley, H., Connolly, B., Gorst, S., Kirkham, J.J., Li, T., Matvienko-Sikar, K., Thomas, V., Tunis, S., Williamson, P.R. and Group, for the C.-S. 2019. Core Outcome Set-STAndardised Protocol Items: the COS-STAP Statement. *Trials* 20(1), p. 116. Available at: https://doi.org/10.1186/s13063-019-3230-x.

Kirkham, J.J., Tunis, S., Clarke, M., Williamson, P.R., Altman, D.G., Blazeby, J.M. and Davis, K. 2017. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLOS Medicine* 14(11), p. e1002447. Available at: https://www.ncbi.nlm.nih.gov/pubmed/29145404.

Kirwan, J., Heiberg, T., Hewlett, S., Hughes, R., Kvien, T., Ahlmen, M., Boers, M., Minnock, P., Saag, K., Shea, B., Suarez Almazor, M. and Taal, E. 2003. Outcomes from the Patient Perspective Workshop at OMERACT 6. *The Journal of rheumatology* 30(4), pp. 868–872.

Kitzinger, J. 1995. Qualitative research. Introducing focus groups. *BMJ (Clinical research ed.)* 311(7000), pp. 299–302. Available at: https://www.ncbi.nlm.nih.gov/pubmed/7633241.

Kloppenburg, M., Bøyesen, P., Smeets, W., Haugen, I., Liu, R., Visser, W. and van der Heijde, D.M. 2014. Report from the OMERACT Hand Osteoarthritis Special Interest Group: Advances and Future Research Priorities. *The Journal of Rheumatology* 41(4), pp. 810 LP – 818. Available at: http://www.jrheum.org/content/41/4/810.abstract.

Kus, S., Dereskewitz, C., Wickert, M., Schwab, M., Eisenschenk, A., Steen, M. and Rudolf, K.-D. 2011. Validation of the comprehensive international classification of functioning, disability and health (ICF) core set for hand conditions. *Hand Therapy* 16(3), pp. 58–66.

Kus, S., Oberhauser, C. and Cieza, A. 2012. Validation of the brief international classification of functioning, disability, and health (ICF) core set for hand conditions. *Journal of Hand Therapy* 25(3), pp. 274–287. doi: 10.1016/j.jht.2012.02.003.

Lai, C.H. 2004. Motivation in hand-injured patients with and without work-related injury. *J Hand Ther* 17(1), pp. 6–17. Available at: internal-pdf://254.159.163.153/Lai-2004-Motivation in hand-injured patients w.pdf.

Laugharne, E., Bhavsar, D. and Rajaratnam, V. 2013. The distribution of hand fractures: a British perspective. *European Journal of Plastic Surgery* 36(6), pp. 367–370. Available at: https://doi.org/10.1007/s00238-012-0775-2.

Li, Y.K., Alolabi, N., Kaur, M.N. and Thoma, A. 2015. A systematic review of utilities in hand surgery literature. *Journal of Hand Surgery* 40(5), pp. 997–1005. Available at: https://www.ncbi.nlm.nih.gov/pubmed/25771479.

Lingard, L., Albert, M. and Levinson, W. 2008. Grounded theory, mixed methods, and action research. *BMJ* 337, p. a567. Available at: http://www.bmj.com/content/337/bmj.39602.690162.47.abstract.

Lissitz, R.W. and Green, S.B. 1975. Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of applied psychology* 60(1), p. 10.

MacDermid, J.C., Turgeon, T., Richards, R.S., Beadle, M. and Roth, J.H. 1998. Patient Rating of Wrist Pain and Disability: A Reliable and Valid Measurement Tool. *Journal of Orthopaedic Trauma* 12(8), pp. 577–86. Available at: https://journals.lww.com/jorthotrauma/Fulltext/1998/11000/Patient_Rating_of_Wrist_Pain_and_Disability__A.9.aspx.

Macefield, R.C., Jacobs, M., Korfage, I.J., Nicklin, J., Whistance, R.N., Brookes, S.T., Sprangers, M.A.G. and Blazeby, J.M. 2014. Developing core outcomes sets: Methods for identifying and including patient-reported outcomes (PROs). *Trials* 15(1), p. 49. Available at: https://www.ncbi.nlm.nih.gov/pubmed/24495582.

MacIntyre, N.J. and Dewan, N. 2016. Epidemiology of distal radius fractures and factors predicting risk and prognosis. *Journal of Hand Therapy* 29(2), pp. 136–145. Available at: https://www.ncbi.nlm.nih.gov/pubmed/27264899.

MacLennan, S., Kirkham, J., Lam, T.B.L. and Williamson, P.R. 2018. A randomized trial comparing three Delphi feedback strategies found no evidence of a difference in a setting with high initial agreement. *Journal of clinical epidemiology* 93, pp. 1–8. doi: 10.1016/j.jclinepi.2017.09.024.

Malterud, K. 2001. Qualitative research: standards, challenges, and guidelines. *The Lancet* 358(9280), pp. 483–488. Available at: http://www.sciencedirect.com/science/article/pii/S0140673601056276.

Marson, B.A., Manning, J.C., James, M., Craxford, S., Deshmukh, S.R., Perry, D.C. and Ollivere, B.J. 2021. Development of the CORE-Kids core set of outcome domains for studies of childhood limb fractures. *The Bone & Joint Journal* 103-B(12), pp. 1821–1830. Available at: https://doi.org/10.1302/0301-620X.103B.BJJ-2020-2321.R2.

Maxwell, L.J. and Beaton, D.E. 2020. Controversy and Debate Series on Core Outcome Sets. Paper 2: Debate on Paper 1 from the perspective of OMERACT [Outcome Measures in Rheumatology]. *Journal of Clinical Epidemiology* 125, pp. 213–215. Available at: https://doi.org/10.1016/j.jclinepi.2020.05.012.

McColl, E., Jacoby, A., Thomas, L., Soutter, J., Bamford, C., Steen, N., Thomas, R., Harvey, E., Garratt, A. and Bond, J. 2001. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health technology assessment (Winchester, England)* 5(31), pp. 1–256. doi: 10.3310/hta5310.

McCoyd, J.L.M. and Kerson, T.S. 2006. Conducting Intensive Interviews Using Email: A Serendipitous Comparative Opportunity. *Qualitative Social Work* 5(3), pp. 389–406. Available at: https://doi.org/10.1177/1473325006067367.

McDowell, I. 2006. *Measuring Health: A guide to rating scales and questionnaires*. New York: New York: Oxford University Press. doi: 10.1093/acprof:oso/9780195165678.001.0001.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J.A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J.J., Devereaux, P.J., Dickersin, K., Egger, M., Ernst, E., Gøtzsche, P.C., Grimshaw, J., Guyatt, G., Higgins, J., Ioannidis, J.P.A., Kleijnen, J., Lang, T., Magrini, N., McNamee, D., Moja, L., Mulrow, C., Napoli, M., Oxman, A., Pham, B., Rennie, D., Sampson, M., Schulz, K.F., Shekelle, P.G., Tovey, D. and Tugwell, P. 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine* 151(4), pp. 264–269. doi: 10.7326/0003-4819-151-4-200908180-00135.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M. and de Vet, H.C.W. 2010. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 19(4), pp. 539–549. Available at: https://pubmed.ncbi.nlm.nih.gov/20169472.

Mokkink, L.B., Terwee, C.B., Stratford, P.W., Alonso, J., Patrick, D.L., Riphagen, I., Knol, D.L., Bouter, L.M. and de Vet, H.C.W. 2009. Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Quality of Life Research* 18(3), pp. 313–333. Available at: https://doi.org/10.1007/s11136-009-9451-9.

Murphy, M.K., Black, N.A., Lamping, D.L., McKee, C.M., Sanderson, C.F., Askham, J. and Marteau, T. 1998. Consensus development methods, and their use in clinical guideline development. *Health technology assessment (Winchester, England)* 2(3), pp. i–iv, 1–88.

Nasser, J.S., Huetteman, H.E., Shauver, M.J. and Chung, K.C. 2018. Older Patient Preferences for Internal Fixation after a Distal Radius Fracture: A Qualitative Study from the Wrist and Radius Injury Surgical Trial. *Plast Reconstr Surg* 142(1), pp. 34e-41e. Available at: https://www.ncbi.nlm.nih.gov/pubmed/29952895.

National Institute for Health and Care Excellence (Great Britain) 2015. *Developing NICE guidelines: the manual*. National Institute for Health and Care Excellence (NICE).

National Institute for Health Research Health Technology Assessment 2022. *Health Technology Assessment Database*. Available at: https://www.crd.york.ac.uk/CRDWeb/HomePage.asp [Accessed: 20 May 2022].

Nunkoosing, K. 2005. The Problems With Interviews. *Qualitative Health Research* 15(5), pp. 698–706. Available at: https://doi.org/10.1177/1049732304273903.

O’Brien, L. and Presnell, S. 2010. Patient experience of distraction splinting for complex finger fracture dislocations. *J Hand Ther* 23(3), pp. 249–9; quiz 260. Available at: internal-pdf://0685155256/O’Brien-2010-Patient experience of distraction.pdf.

O'Neill, T.W., Cooper, C., Finn, J.D., Lunt, M., Purdie, D., Reid, D.M., Rowe, R., Woolf, A.D. and Wallace, W.A. 2001. Incidence of distal forearm fracture in British men and women. *Osteoporosis International* 12(7), pp. 555–558. doi: 10.1007/s001980170076.

Office for National Statistics - UK Statistics Authority 2016. *Office for National Statistics Population Estimates*. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates.

OMERACT 2022. *OMERACT Handbook*. Available at: https://omeracthandbook.org/handbook.

Page, M.J., McKenzie, J.E., Kirkham, J., Dwan, K., Kramer, S., Green, S. and Forbes, A. 2014. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *The Cochrane database of systematic reviews* 2014(10), p. MR000035. doi: 10.1002/14651858.MR000035.pub2.

Patton, M.Q. 2002. *Qualitative research and evaluation methods*. 3rd ed. SAGE Publications.

Poolman, R.W., Goslings, J.C., Lee, J.B., Statius Muller, M., Steller, E.P. and Struijs, P.A. 2006. Conservative treatment for closed fifth metacarpal (small finger) fractures. *Orthopedic Trauma Directions* 4(5), pp. 21–28. Available at: https://www.ncbi.nlm.nih.gov/pubmed/16034891.

Potter, S., Holcombe, C., Ward, J.A., Blazeby, J.M. and Group, B.S. 2015. Development of a core outcome set for research and audit studies in reconstructive breast surgery. *The British journal of surgery* 102(11), pp. 1360–1371. Available at: https://pubmed.ncbi.nlm.nih.gov/26179938.

Prinsen, C.A.C., Vohra, S., Rose, M.R., Boers, M., Tugwell, P., Clarke, M., Williamson, P.R. and Terwee, C.B. 2016. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. *Trials* 17(1), p. 449. Available at: https://doi.org/10.1186/s13063-016-1555-2.

Ratislavová, K. and Ratislav, J. 2014. Asynchronous email interview as a qualitative research method in the humanities. *Human Affairs* 24(4), pp. 452–460.

Remus, A., Smith, V. and Wuytack, F. 2021. 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 Medical Research Methodology* 21(1), p. 10. Available at: https://doi.org/10.1186/s12874-020-01197-3.

Ring, D. and Leopold, S.S. 2015. Editorial—Measuring Satisfaction: Can It Be Done? *Clinical Orthopaedics and Related Research®* 473(10), pp. 3071–3073. Available at: https://doi.org/10.1007/s11999-015-4485-5.

Robinson, R. 1993. Cost-benefit analysis. *BMJ (Clinical research ed.)* 307(6909), pp. 924–926. Available at: https://pubmed.ncbi.nlm.nih.gov/8241859.

Rowe, G. and Wright, G. 1999. The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting* 15(4), pp. 353–375. Available at: https://www.sciencedirect.com/science/article/pii/S0169207099000187.

Rowe, G., Wright, G. and Bolger, F. 1991. Delphi: A reevaluation of research and theory. *Technological Forecasting and Social Change* 39(3), pp. 235–251. Available at: https://www.sciencedirect.com/science/article/pii/004016259190039I.

Rudolf, K.-D.D., Kus, S., Chung, K.C., Johnston, M., LeBlanc, M. and Cieza, A. 2012. Development of the international classification of functioning, disability and health core sets for hand conditions results of the World Health Organization international consensus process. *Disability and Rehabilitation* 34(8), pp. 681–693. Available at: https://www.ncbi.nlm.nih.gov/pubmed/21978202.

Ruperto, N., Ravelli, A., Murray, K.J., Lovell, D.J., Andersson-Gare, B., Feldman, B.M., Garay, S., Kuis, W., Machado, C., Pachman, L., Prieur, A.-M., Rider, L.G., Silverman, E., Tsitsami, E., Woo, P., Giannini, E.H. and Martini, A. 2003. Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis. *Rheumatology (Oxford, England)* 42(12), pp. 1452–1459. doi: 10.1093/rheumatology/keg403.

Sackett, D.L., Rosenberg, W.M., Gray, J.A., Haynes, R.B. and Richardson, W.S. 1996. Evidence based medicine: what it is and what it isn't. *BMJ (Clinical research ed.)* 312(7023), pp. 71–72. Available at: https://pubmed.ncbi.nlm.nih.gov/8555924.

Sale, J.E.M., Frankel, L., Thielke, S. and Funnell, L. 2017. Pain and fracture-related limitations persist 6 months after a fragility fracture. *Rheumatol Int* 37(8), pp. 1317–1322. Available at: internal-pdf://190.184.163.215/Sale-2017-Pain and fracture-related limitation.pdf.

Sbaraini, A., Carter, S.M., Evans, R.W. and Blinkhorn, A. 2011. How to do a grounded theory study: a worked example of a study of dental practices. *BMC Medical Research Methodology* 11(1), p. 128. Available at: https://doi.org/10.1186/1471-2288-11-128.

Schier, J.S. and Chan, J. 2007. Changes in life roles after hand injury. *J Hand Ther* 20(1), pp. 57–68; quiz 69. Available at: internal-pdf://134.241.164.22/Schier-2007-Changes in life roles after hand i.pdf.

Schmitt, J., Langan, S., Stamm, T. and Williams, H.C. 2011. Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. *The Journal of investigative dermatology* 131(3), pp. 623–630. doi: 10.1038/jid.2010.303.

Schmitt, J., Spuls, P.I., Thomas, K.S., Simpson, E., Furue, M., Deckert, S., Dohil, M., Apfelbacher, C., Singh, J.A., Chalmers, J. and Williams, H.C. 2014. The Harmonising Outcome Measures for Eczema (HOME) statement to assess clinical signs of atopic eczema in trials. *Journal of Allergy and Clinical Immunology* 134(4), pp. 800–807. Available at: https://doi.org/10.1016/j.jaci.2014.07.043.

Schmitt, J., Williams, H. and Group, on behalf of the H.D. 2010. Harmonising Outcome Measures for Eczema (HOME).Report from the First International Consensus Meeting (HOME 1), 24 July 2010, Munich, Germany. *British Journal of Dermatology* 163(6), pp. 1166–1168. Available at: https://doi.org/10.1111/j.1365-2133.2010.10054.x.

Seitz, S. 2015. Pixilated partnerships, overcoming obstacles in qualitative interviews via Skype: a research note. *Qualitative Research* 16(2), pp. 229–235. Available at: https://doi.org/10.1177/1468794115577011.

Selb, M., Escorpizo, R., Kostanjsek, N., Stucki, G., Üstün, B. and Cieza, A. 2015. A guide on how to develop an International Classification of Functioning, Disability and Health Core Set. *European journal of physical and rehabilitation medicine* 51(1), pp. 105–117.

Selwyn, N. and Robson, K. 1998. Using e-mail as a research tool. *Social research update*

21(6)

Sinha, I.P., Smyth, R.L. and Williamson, P.R. 2011. Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials: Recommendations for the Future Based on a Systematic Review of Existing Studies. *PLoS Medicine* 8(1), p. e1000393. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3026691/pdf/pmed.1000393.pdf.

Smith, M.E., Auchincloss, J.M. and Ali, M.S. 1985. Causes and consequences of hand injury. *The Journal of Hand Surgery: British & European Volume* 10(3), pp. 288–292. Available at: http://www.sciencedirect.com/science/article/pii/S0266768185800450.

Smith, T., Noble, M., Noble, S., Wright, G., McLennan, D. and Plunkett, E. 2015. The English indices of deprivation 2015. *London: Department for Communities and Local Government* , pp. 1–123.

Smith, T.O., Hawker, G.A., Hunter, D.J., March, L.M., Boers, M., Shea, B.J., Christensen, R., Guillemin, F., Terwee, C.B., Williamson, P.R., Dodd, S., Roos, E.M., Loeser, R.F., Schnitzer, T.J., Kloppenburg, M., Neogi, T., Ladel, C.H., Kalsi, G., Kaiser, U., Buttel, T.W., Ashford, A.E., Mobasheri, A., Arden, N.K., Tennant, A., Hochberg, M.C., de Wit, M., Tugwell, P. and Conaghan, P.G. 2019. The OMERACT-OARSI Core Domain Set for Measurement in Clinical Trials of Hip and/or Knee Osteoarthritis. *The Journal of rheumatology* 46(8), pp. 981–989. doi: 10.3899/jrheum.181194.

SPIRIT Group 2022. *SPIRIT Checklist - Item 12: Outcomes*. Available at: https://www.spirit-statement.org/outcomes-2/.

Spuls, P.I., Gerbens, L.A.A., Simpson, E., Apfelbacher, C.J., Chalmers, J.R., Thomas, K.S., Prinsen, C.A.C., von Kobyletzki, L.B., Singh, J.A., Williams, H.C. and Schmitt, J. 2017. Patient-Oriented Eczema Measure (POEM), a core instrument to measure symptoms in clinical trials: a Harmonising Outcome Measures for Eczema (HOME) statement. *The British journal of dermatology* 176(4), pp. 979–984. doi: 10.1111/bjd.15179.

Stucki, G., Boonen, A., Tugwell, P., Cieza, A. and Boers, M. 2007. The World Health Organisation International Classification of Functioning, Disability and Health: a conceptual model and interface for the OMERACT process. *The Journal of Rheumatology* 34(3), pp. 600 LP – 606. Available at: http://www.jrheum.org/content/34/3/600.abstract.

Tausch, A.P. and Menold, N. 2016. Methodological Aspects of Focus Groups in Health Research: Results of Qualitative Interviews With Focus Group Moderators. *Global Qualitative Nursing Research* 3, pp. 1–12. Available at: https://doi.org/10.1177/2333393616630466.

Thompson, P.W., Taylor, J. and Dawson, A. 2004. The annual incidence and seasonal variation of fractures of the distal radius in men and women over 25 years in Dorset, UK. *Injury* 35(5), pp. 462–466. doi: 10.1016/S0020-1383(03)00117-7.

Troianello, T., Yancosek, K. and Rhee, P.C. 2017. Unilateral hand training on functional performance in patients with upper extremity trauma. *J Hand Ther* . Available at: internal-pdf://115.203.164.38/Troianello-2017-Unilateral hand training on fu.pdf.

Tugwell, P., Boers, M., Brooks, P., Simon, L., Strand, V. and Idzerda, L. 2007. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 8, p. 38. Available at: https://pubmed.ncbi.nlm.nih.gov/18039364.

Tunis, S.R., Clarke, M., Gorst, S.L., Gargon, E., Blazeby, J.M., Altman, D.G. and Williamson, P.R. 2016. Improving the relevance and consistency of outcomes in comparative effectiveness research. *Journal of comparative effectiveness research* 5(2), pp. 193–205. doi: 10.2217/cer-2015-0007.

Turnbull, A.E., Dinglas, V.D., Friedman, L.A., Chessare, C.M., Sepulveda, K.A., Bingham, C.O. 3rd and Needham, D.M. 2018. A survey of Delphi panelists after core outcome set development revealed positive feedback and methods to facilitate panel member participation. *Journal of clinical epidemiology* 102, pp. 99–106. doi: 10.1016/j.jclinepi.2018.06.007.

Van de Ven, A.H. and Delbecq, A.L. 1972. The nominal group as a research instrument for exploratory health studies. *American journal of public health* 62(3), pp. 337–342. Available at: https://pubmed.ncbi.nlm.nih.gov/5011164.

van Onselen, E.B.H., Karim, R.B., Hage, J.J. and Ritt, M.J.P.F. 2003. Prevalence and distribution of hand fractures. *Journal of hand surgery (Edinburgh, Scotland)* 28(5), pp. 491–5. Available at: https://www.ncbi.nlm.nih.gov/pubmed/12954264.

van Staa, T.P., Dennison, E.M., Leufkens, H.G.M. and Cooper, C. 2001. Epidemiology of fractures in England and Wales. *Bone* 29(6), pp. 517–522. Available at:

http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L33124293%0Ahttp://dx.doi.org/10.1016/S8756-3282(01)00614-7%0Ahttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=87563282&id=doi:10.1016%2FS8756-3282%2801%2900614-7&atitle=Epidemiology.

Vargus-Adams, J.N. and Martin, L.K. 2009. Measuring what matters in cerebral palsy: a breadth of important domains and outcome measures. *Archives of physical medicine and rehabilitation* 90(12), pp. 2089–2095. doi: 10.1016/j.apmr.2009.06.018.

Verver, D., Timmermans, L., Klaassen, R.A., Van der Vlies, C.H., Vos, D.I. and Schep, N.W.L. 2017. Treatment of extra-articular proximal and middle phalangeal fractures of the hand: a systematic review. *Strategies in Trauma and Limb Reconstruction* 12(2), pp. 63–76. Available at: https://www.ncbi.nlm.nih.gov/pubmed/28260179.

Watson, N.J., Martin, S.A. and Keating, J.L. 2018. The impact of wrist fracture, surgical repair and immobilization on patients: a qualitative study. *Clin Rehabil* 32(6), pp. 841–851. Available at: https://doi.org/10.1177/0269215518754614.

Weinstock-Zlotnick, G. and Mehta, S.P. 2016. A structured literature synthesis of wrist outcome measures: An evidence-based approach to determine use among common wrist diagnoses. *Journal of Hand Therapy* 29(2), pp. 98–110. doi: 10.1016/j.jht.2016.03.001.

Weldring, T. and Smith, S.M.S. 2013. Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health services insights* 6, pp. 61–68. Available at: https://pubmed.ncbi.nlm.nih.gov/25114561.

Williamson, P.R., Altman, D.G., Bagley, H., Barnes, K.L., Blazeby, J.M., Brookes, S.T., Clarke, M., Gargon, E., Gorst, S., Harman, N., Kirkham, J.J., McNair, A., Prinsen, C.A.C., Schmitt, J., Terwee, C.B. and Young, B. 2017. The COMET Handbook: version 1.0. *Trials* 18(S3), p. 280. Available at: https://www.ncbi.nlm.nih.gov/pubmed/28681707.

Williamson, P.R., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E. and Tugwell, P. 2012. Developing core outcome sets for clinical trials: issues to consider. *Trials* 13(1), p. 132. Available at: https://www.ncbi.nlm.nih.gov/pubmed/22867278.

Williamson, P.R., Blazeby, J.M., Brookes, S.T., Clarke, M., Terwee, C.B. and Young, B. 2020. Controversy and Debate Series on Core Outcome Sets. Paper 4: Debate on Paper 1 from the perspective of COMET [Core Outcome Measures in Effectiveness Trials]. *Journal of clinical epidemiology* 125, pp. 222–224. doi: 10.1016/j.jclinepi.2020.05.014.

Wilson, I.B. and Cleary, P.D. 1995. Linking Clinical Variables With Health-Related Quality of Life: A Conceptual Model of Patient Outcomes. *JAMA* 273(1), pp. 59–65. Available at: https://doi.org/10.1001/jama.1995.03520250075037.

Woodyatt, C.R., Finneran, C.A. and Stephenson, R. 2016. In-Person Versus Online Focus Group Discussions: A Comparative Analysis of Data Quality. *Qualitative Health Research* 26(6), pp. 741–749. Available at: https://doi.org/10.1177/1049732316631510.

World Health Organization 2001. *International Classification of Functioning, Disability, and Health*. Available at: https://apps.who.int/iris/handle/10665/42407 [Accessed: 29 July 2022].

World Health Organization 2002. *Towards a Common Language for Functioning, Disability and Health: ICF*. Available at: https://www.who.int/classifications/icf/icfbeginnersguide.pdf.

World Health Organization 2022. *International Clinical Trials Registry Platform*. Available at: https://www.who.int/ictrp/en/.

Wormald, J.C.R., Geoghegan, L., Sierakowski, K., Price, A., Peters, M., Jain, A. and Rodrigues, J.N. 2019. Site-specific Patient-reported Outcome Measures for Hand Conditions: Systematic Review of Development and Psychometric Properties. *Plastic and reconstructive surgery. Global open* 7(5), p. e2256. doi: 10.1097/GOX.0000000000002256.

Wouters, R.M., Jobi-Odeneye, A.O., de la Torre, A., Joseph, A. and Hovius, S.E.R. 2021. A Standard Set for Outcome Measurement in Patients With Hand and Wrist Conditions: Consensus by the International Consortium for Health Outcomes Measurement Hand and Wrist Working Group. *The Journal of hand surgery* . doi: 10.1016/j.jhsa.2021.06.004.

Yordanov, Y., Dechartres, A., Atal, I., Tran, V.-T., Boutron, I., Crequit, P. and Ravaud, P. 2018. Avoidable waste of research related to outcome planning and reporting in clinical trials. *BMC medicine* 16(1), p. 87. Available at: https://www.ncbi.nlm.nih.gov/pubmed/29886846.

Young, B. and Bagley, H. 2016. Including patients in core outcome set development: issues to consider based on three workshops with around 100 international delegates. *Research Involvement and Engagement* 2(1), p. 25. Available at: https://doi.org/10.1186/s40900-016-

0039-6.

Zarin, D.A., Tse, T., Williams, R.J., Califf, R.M. and Ide, N.C. 2011. The ClinicalTrials.gov Results Database — Update and Key Issues. *New England Journal of Medicine* 364(9), pp. 852–860. Available at: https://doi.org/10.1056/NEJMsa1012065.

Appendices

Appendix 1 Combinations of regions fractured in multiple fracture events

Number of concurrent fractures per fracture event	Combination of regions fractured	Number of cases	% of total fracture events (n=295)
2	CC	6	2.0
	CH	13	4.4
	CW	41	14
	HH	176	60
	HW	18	6.1
	WW	10	3.4
3	CHH	2	0.7
	CWW	1	0.3
	HHH	16	5.4
	HHW	3	1.0
	HWW	1	0.3
4	CHHH	1	0.3
	HHHH	5	1.7
	HHHW	1	0.3
5	CHHHW	1	0.3

Key: C – carpus H – hand (phalanx/metacarpus) W – wrist (distal radius/ulna)

The number of times the letter occurs in the 'combination' denotes the number of fractures involving a bone in that region of the hand/wrist (e.g. CHHHW means that in a single fracture event a patient sustained one carpus fracture, three phalangeal/metacarpal fractures and one distal radius/ulna fracture)

Appendix 2 Examples of search strategies utilised

Example search strategies are provided for Ovid MEDLINE and PubMed. In the latter, to improve clarity the broad separation of search term categories is delineated by highlighting (blue for 'anatomy terms', yellow for 'generic injury terms' within scope of COS, green for 'hand/wrist-specific injury terms' and pink for excluding 'paediatric-only' studies. The other databases searched were Ovid Embase, Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, PEDro and Ovid PsycINFO.

Ovid MEDLINE

1. exp Hand/
2. hand.ti.
3. hands.mp.
4. exp Hand Bones/
5. phalan*.mp.
6. finger.mp.
7. fingers.mp.
8. thumb.mp.
9. thumbs.mp.
10. metacarp*.mp.
11. wrist.mp.
12. wrists.mp.
13. carpus.mp.
14. carpi.mp.
15. carpal.mp.
16. carpals.mp.
17. scapho*.mp.
18. hamate.mp.
19. hamates.mp.
20. lunate.mp.
21. lunates.mp.
22. triquet*.mp.
23. trapeziu*.mp.
24. trapezoi*.mp.
25. pisiform.mp.

26. pisiforms.mp.
27. exp Radius/ and distal.mp
28. distal radio*.mp.
29. distal radius.mp.
30. distal radial.mp.
31. radial styloid*.mp.
32. exp Collateral Ligament, Ulnar/
33. radial collateral.mp.
34. rcl.mp.
35. ulnar collateral.mp.
36. ucl.mp.
37. sagittal band.mp.
38. sagittal bands.mp.
39. beak ligament.mp.
40. beak ligaments.mp.
41. exp Palmar Plate/
42. volar plate.mp.
43. volar plates.mp.
44. exp Triangular Fibrocartilage/
45. triangular fibrocartilage.mp.
46. triangular fibrocartilages.mp.
47. triangular cartilage.mp.
48. triangular cartilages.mp.
49. triangular fibrocartilaginous.mp.
50. triangular ligament.mp.
51. triangular ligaments.mp.
52. tfcc.mp.
53. exp Hand Joints/
54. interphalangeal.mp.
55. metacarpophalangeal.mp.
56. carpometacarpal.mp.
57. druj.mp.
58. pericapitate.mp.
59. transcapitate.mp.
60. midcarpal.mp.
61. mesocarpal.mp.
62. mediocarpal.mp.
63. carpocarpal.mp.
64. transcarpal.mp.
65. intracarpal.mp.
66. perihamate.mp.

67. transhamate.mp.
68. hemihamate.mp.
69. perilunate.mp.
70. perilunar.mp.
71. translunate.mp.
72. midmetacarpal.mp.
73. transmetacarpal.mp.
74. midphalangeal.mp.
75. transphalangeal.mp.
76. peripisiform.mp.
77. periscaphoid.mp.
78. transscaphoid.mp.
79. peritrapezium.mp.
80. peritrapezial.mp.
81. transtrapezium.mp.
82. transtrapezial.mp.
83. pantrapezial.mp.
84. peritrapezoid.mp.
85. peritrapezoidal.mp.
86. peritriquetral.mp.
87. transtriquetrum.mp.
88. transtriquetral.mp.
89. cleland's ligament.mp.
90. cleland's ligaments.mp.
91. grayson's ligament.mp.
92. grayson's ligaments.mp.
93. extensor retinaculum.mp.
94. lateral band.mp.
95. lateral bands.mp.
96. lunotriquetral.mp.
97. natatory ligament.mp.
98. natatory ligaments.mp.
99. pisohamate.mp.
100. pisometacarpal.mp.
101. radiocapitate.mp.
102. radiolunotriquetral.mp.
103. radiopalmar.mp.
104. radioscaphocapitate.mp.
105. radioscapholunate.mp.
106. radiotriquetral.mp.
107. retinacular ligament.mp.

108. retinacular ligaments.mp.
109. scaphotrapeziotrapezoid.mp.
110. scaphotrapezoid.mp.
111. flexor pulley.mp.
112. flexor pulleys.mp.
113. annular pulley.mp.
114. annular pulleys.mp.
115. oblique pulley.mp.
116. oblique pulleys.mp.
117. trapeziocapitate.mp.
118. trapeziotrapezoid.mp.
119. triquetralcapitate.mp.
120. triquetralhamate.mp.
121. triquetrocapitate.mp.
122. triquetrohamate.mp.
123. ulnocapitate.mp.
124. ulnolunate.mp.
125. ulnotriquetral.mp.
126. (abductor digiti or abductor pollicis or adductor pollicis or anconeus or brachialis or brachioradialis or extensor carpi or extensor digiti or extensor digitorum or extensor indicis or extensor pollicis or flexor carpi or flexor digiti minimi or flexor digitorum or flexor pollicis or hypothenar or hypothenars or interosseous or interosseus or interossei or lumbrical or lumbricals or opponens digiti minimi or opponens pollicis or palmaris brevis or palmaris longus or pronator quadratus or pronator teres or supinator or supinators or thenar or thenars or parona or APL or ECRB or ECRL or ECU or ED or EDC or EDM or EIP or EPB or EPL or FCR or FCU or FDP or FDS or FPL or hand or wrist or finger or thumb).mp. and ((tendon or tendons).mp. or exp Tendons/)
127. central slip.mp.
128. central slips.mp.
129. extensor expansion.mp.
130. extensor expansions.mp.
131. extensor hood.mp.
132. extensor hoods.mp.
133. junctura tendinum.mp.
134. juncturae tendinum.mp.
135. palmaris brevis.mp.
136. palmaris longus.mp.
137. fractures, bone/ or exp fracture dislocation/ or exp fractures, avulsion/ or exp fractures, closed/ or exp fractures, comminuted/ or exp fractures, compression/ or exp fractures, malunited/ or exp fractures, multiple/ or exp fractures, open/ or exp

fractures, spontaneous/ or exp fractures, stress/ or exp intra-articular fractures/ or exp osteoporotic fractures/

138. Joint Instability/
139. Joint Dislocation/
140. Sprains and Strains/
141. exp Tendon Injuries/
142. injuries.fs
143. fractur*.mp.
144. trauma.mp.
145. non-union.mp.
146. nonunion.mp.
147. avulsio*.mp.
148. tear*.mp.
149. torn*.mp.
150. rupture*.mp.
151. sprain*.mp.
152. instability*.mp.
153. dislocation*.mp.
154. dislocated.mp.
155. subluxation*.mp.
156. subluxed.mp.
157. mallet*.mp.
158. exp Hand Injuries/
159. Forearm Injuries/ or exp Radius Fractures/
160. exp Wrist Injuries/
161. boutonnier*.mp.
162. colles*.ti,ab,kw and fracture*.mp.
163. smith*.ti,ab,kw and fracture*.mp.
164. bennett*.ti,ab,kw and fracture*.mp.
165. rolando*.ti,ab,kw and fracture*.mp.
166. barton*.ti,ab,kw and fracture*.mp.
167. ((jersey or rugby or sweater) and (finger* or fracture* or avulsion* or rupture* or tear*)).mp.
168. (boxer* and (fracture* or finger or fingers or knuckle*)).mp.
169. (gamekeeper* and (fracture* or avulsion* or rupture* or tear* or thumb or thumbs)).mp.
170. (skier* and (fracture* or avulsion* or rupture* or tear* or thumb or thumbs)).mp
171. stener.mp.
172. die-punch.mp.
173. or/1-136
174. or/137-157

175. or/158-172
176. 173 and 174
177. 175 or 176
178. 177 not ((Infant/ or Preschool/ or exp Child/ or Adolescent/) not exp Adult/)
179. 178 not review.pt
180. limit 181 to yr="2014 -Current"

PubMed

((("hand"[mh] OR "hand"[ti] OR "hands"[all fields] OR "hand bones"[mh] OR "phalanx"[all fields] OR "phalanxes"[all fields] OR "phalanges"[all fields] OR "phalangeal"[all fields] OR "finger"[all fields] OR "fingers"[all fields] OR "thumb"[all fields] OR "thumbs"[all fields] OR "metacarpus"[all fields] OR "metacarpal"[all fields] OR "metacarpals"[all fields] OR "wrist"[all fields] OR "wrists"[all fields] OR "carpus"[all fields] OR "carpi"[all fields] OR "carpal"[all fields] OR "carpals"[all fields] OR scapho*[all fields] OR "hamate"[all fields] OR "hamates"[all fields] OR "lunate"[all fields] OR "lunates"[all fields] OR triquetr*[all fields] OR trapez*[all fields] OR "pisiform"[all fields] OR "pisiforms"[all fields] OR ("radius"[mh] AND "distal"[all fields]) OR distal radio*[all fields] OR "distal radius"[all fields] OR "distal radial"[all fields] OR radial stylo*[all fields] OR "collateral ligament, ulnar"[mh] OR "radial collateral"[all fields] OR "rcl"[all fields] OR "ulnar collateral"[all fields] OR "ucl"[all fields] OR "sagittal band"[all fields] OR "sagittal bands"[all fields] OR "beak ligament"[all fields] OR "beak ligaments"[all fields] OR "palmar plate"[mh] OR "volar plate"[all fields] OR "volar plates"[all fields] OR "triangular fibrocartilage"[mh] OR "triangular fibrocartilage"[all fields] OR "triangular fibrocartilages"[all fields] OR "triangular cartilage"[all fields] OR "triangular cartilages"[all fields] OR "triangular fibrocartilaginous"[all fields] OR "triangular ligament"[all fields] OR "triangular ligaments"[all fields] OR "tfcc"[all fields] OR "hand joints"[mh] OR "interphalangeal"[all fields] OR "metacarpophalangeal"[all fields] OR "carpometacarpal"[all fields] OR "druj"[all fields] OR "pericapitate"[all fields] OR "transcapitate"[all fields] OR "midcarpal"[all fields] OR "mediocarpal"[all fields] OR "transcarpal"[all fields] OR "intracarpal"[all fields] OR "perihamate"[all fields] OR "transhamate"[all fields] OR "hemihamate"[all fields] OR "perilunate"[all fields] OR "perilunar"[all fields] OR "translunate"[all fields] OR "midmetacarpal"[all fields] OR "transmetacarpal"[all fields] OR "midphalangeal"[all fields] OR "transphalangeal"[all fields] OR "peripisiform"[all fields] OR "periscaphoid"[all fields] OR "transscaphoid"[all fields] OR "peritrapezium"[all fields] OR "peritrapezial"[all fields] OR "transtrapezium"[all fields] OR "transtrapezial"[all fields] OR "pantrapezial"[all fields] OR "peritrapezoid"[all fields] OR "peritrapezoidal"[all fields] OR "peritriquetral"[all fields] OR "transtriquetrum"[all fields] OR "transtriquetral"[all fields] OR "cleland’s ligament"[all fields] OR "cleland’s ligaments"[all fields] OR "grayson’s ligament"[all fields] OR "grayson’s ligaments"[all fields] OR "extensor retinaculum"[all fields] OR "lateral band"[all fields] OR "lateral bands"[all fields] OR "lunotriquetral"[all fields] OR "natatory ligament"[all fields] OR "natatory ligaments"[all fields] OR "pisohamate"[all fields] OR "pisometacarpal"[all fields] OR "radiocapitate"[all fields] OR "radiolunotriquetral"[all fields] OR "radiopalmar"[all fields] OR "radioscaphocapitate"[all fields] OR "radioscapholunate"[all fields] OR "radiotriquetral"[all fields] OR "retinacular ligament"[all fields] OR "retinacular ligaments"[all fields] OR "scaphotrapeziotrapezoid"[all fields] OR "scaphotrapezoid"[all fields] OR "flexor pulley"[all fields] OR "flexor pulleys"[all fields] OR "annular pulley"[all fields] OR "annular pulleys"[all fields] OR "oblique pulley"[all fields] OR "oblique pulleys"[all fields] OR "trapeziocapitate"[all fields] OR "trapeziotrapezoid"[all fields] OR "triquetrocapitate"[all fields] OR "triquetrohamate"[all fields] OR "ulnocapitate"[all fields] OR "ulnolunate"[all fields] OR "ulnotriquetral"[all fields] OR (("abductor digiti"[all fields] OR "abductor pollicis"[all fields] OR "adductor pollicis"[all fields] OR "anconeus"[all fields] OR "brachialis"[all fields] OR "brachioradialis"[all fields] OR "extensor carpi"[all fields] OR "extensor digiti"[all fields] OR "extensor digitorum"[all fields] OR "extensor indicis"[all fields] OR "extensor pollicis"[all fields] OR "flexor carpi"[all fields] OR "flexor digiti minimi"[all fields] OR "flexor digitorum"[all fields] OR "flexor pollicis"[all fields] OR "hypothenar"[all fields] OR "hypothenars"[all fields] OR "interosseous"[all fields] OR "interosseus"[all fields] OR "interossei"[all fields] OR "lumbrical"[all fields] OR "lumbricals"[all fields] OR "opponens digiti minimi"[all fields] OR "opponens pollicis"[all fields] OR "palmaris brevis"[all fields] OR "palmaris longus"[all fields] OR "pronator quadratus"[all fields] OR "pronator teres"[all fields] OR "supinator"[all fields] OR "supinators"[all fields] OR "thenar"[all fields] OR "thenars"[all fields] OR "parona"[all fields] OR "APL"[all fields] OR "ECRB"[all fields] OR "ECRL"[all fields] OR "ECU"[all fields] OR "ED"[all fields] OR "EDC"[all fields] OR "EDM"[all fields] OR "EIP"[all fields] OR "EPB"[all fields] OR "EPL"[all fields] OR "FCR"[all fields] OR "FCU"[all fields] OR "FDP"[all fields] OR "FDS"[all fields] OR "FPL"[all fields] OR "hand"[all fields] OR "wrist"[all fields] OR "finger"[all fields] OR "thumb"[all fields]) AND ("tendon"[all fields] OR "tendons"[all fields] OR "tendons"[mh])) OR "central slip"[all fields] OR "central slips"[all fields] OR "extensor expansion"[all fields] OR "extensor expansions"[all fields] OR "extensor hood"[all fields] OR "extensor hoods"[all fields] OR "junctura tendinum"[all fields] OR "juncturae tendinum"[all fields] OR "palmaris brevis"[all fields] OR "palmaris longus"[all fields]) **AND** ("fractures, bone"[mh] OR "sprains and strains"[mh] OR "joint instability"[mh] OR "joint dislocations"[mh] OR "tendon injuries"[mh] OR "fracture"[all fields] OR "fractures"[all fields] OR "fractured"[all fields] OR "trauma"[all fields] OR "non-union"[all fields] OR "nonunion"[all fields] OR "avulsion"[all fields] OR "avulsions"[all fields] OR "tear"[all fields] OR "tears"[all fields] OR "torn"[all fields] OR "rupture"[all fields] OR "ruptures"[all fields] OR "ruptured"[all fields] OR sprain*[all fields] OR instabilit*[all fields] OR "dislocation"[all fields] OR "dislocations"[all fields] OR "dislocated"[all fields] OR "subluxation"[all fields] OR "subluxations"[all fields] OR "subluxed"[all fields] OR "mallet"[all fields] OR "mallets"[all fields])) **OR** ("hand injuries"[mh] OR "forearm injuries"[mh] OR "wrist injuries"[mh] OR "colles’ fracture"[mh] OR boutonnier*[all fields] OR (colles*[tw] AND fracture*[all fields]) OR (smith*[tw] AND fracture*[all fields]) OR (bennett*[tw] AND fracture*[all fields]) OR (rolando*[tw] AND fracture*[all fields]) OR (barton*[tw] AND fracture*[all fields]) OR ((jersey[all fields] OR rugby[all fields] OR sweater[all fields]) AND (finger*[all fields] OR fracture*[all fields] OR avulsion*[all fields] OR rupture*[all fields] OR tear*[all fields])) OR (boxer*[all fields] AND (fracture*[all fields] OR finger[all fields] OR fingers[all fields] OR knuckle*[all fields])) OR (gamekeeper*[all fields] AND (fracture*[all fields] OR avulsion*[all fields] OR rupture*[all fields] OR tear*[all fields] OR thumb[all fields] OR thumbs[all fields])) OR (skier* AND (fracture*[all fields] OR avulsion*[all fields] OR rupture*[all fields] OR tear*[all fields] OR thumb[all fields] OR thumbs[all fields])) OR "stener"[all fields] OR "die-punch"[all fields])) **NOT** (("Infant"[mh] OR "Child"[mh] OR "Adolescent"[mh]) NOT "Adult"[mh]) **NOT** review[pt] **AND** ("2014/01/01"[PDat] : "2019/03/29"[PDat])

Appendix 3 Table of all patient-reported outcome measures with linked WHO ICF domains and frequency of use across study types

Patient-reported outcome measure	WHO ICF Outcome Domain(s)	DRF studies (n-121)	Non-DRF studies (n=39)	Total (n=160)
Visual analogue scale – pain	b280 Sensation of pain	85 (70%)	29 (74%)	114 (71%)
Disabilities of the Arm, Shoulder and Hand (DASH)	b126 Temperament and personality functions; b134 Sleep functions; b265 Touch function; b280 Sensation of pain; b640 Sexual functions; b710 Mobility of joint functions; b730 Muscle power functions; d230 Carrying out daily routine; d430 Lifting and carrying objects; d440 Fine hand use; d445 Hand and arm use; d470 Using transportation; d510 Washing oneself; d540 Dressing; d550 Eating; d640 Doing housework; d650 Caring for household objects; d750 Informal social relationships; d760 Family relationships; d850 Remunerative employment; d920 Recreation and leisure	84 (69%)	29 (74%)	113 (71%)
Patient-Rated Wrist Evaluation	b280 Sensation of pain; d410 Changing basic body position; d430 Lifting and carrying objects; d440 Fine hand use; d445 Hand and arm use; d510 Washing oneself; d530 Toileting; d540 Dressing; d550 Eating; d640 Doing housework; d850 Remunerative employment; d920 Recreation and leisure	83 (69%)	28 (72%)	111 (69%)

QuickDASH (Abbreviated version of DASH)	b134 Sleep functions; b265 Touch function; b280 Sensation of pain; d230 Carrying out daily routine; d430 Lifting and carrying objects; d440 Fine hand use; d445 Hand and arm use; d510 Washing oneself; d550 Eating; d640 Doing housework; d750 Informal social relationships; d760 Family relationships; d850 Remunerative employment; d920 Recreation and leisure	82 (68%)	28 (72%)	110 (69%)
EuroQol-5D-3L	b152 Emotional functions; b280 Sensation of pain; d230 Carrying out daily routine; d450 Walking; d510 Washing oneself; d540 Dressing; d850 Remunerative employment; d920 Recreation and leisure; nd-gh	81 (67%)	28 (72%)	109 (68%)
Modified Mayo Wrist Score	b280 Sensation of pain; b710 Mobility of joint functions; b730 Muscle power functions; d230 Carrying out daily routine; nc-overall satisfaction	80 (66%)	26 (67%)	106 (66%)
36-Item Short Form Survey (SF-36)	b126 Temperament and personality functions; b130 Energy and drive functions; b152 Emotional functions; d230 Carrying out daily routine; d640 Doing housework; d750 Informal social relationships; d760 Family relationships; d850 Remunerative employment; d920 Recreation and leisure; nd-ph; nd-gh	77 (64%)	19 (49%)	96 (60%)
Michigan Hand Outcomes Questionnaire (MHQ)	b126 Temperament and personality functions; b134 Sleep functions; b152 Emotional functions; b180 Experience of self	70 (58%)	25 (64%)	95 (59%)

	and time functions; b265 Touch function; b280 Sensation of pain; b710 Mobility of joint functions; b730 Muscle power functions; d430 Lifting and carrying objects; d440 Fine hand use; d540 Dressing; d510 Washing oneself; d550 Eating; d640 Doing housework; s730 Structure of upper extremity; d839 Education unspecified;d850 Remunerative employment; d920 Recreation and leisure; nc-overall satisfaction			
Numeric rating scale – pain	b280 Sensation of pain	54 (45%)	17 (44%)	71 (44%)
12-Item Short Form Survey (SF-12)	b126 Temperament and personality functions; b130 Energy and drive functions; b152 Emotional functions; b280 Sensation of pain; d230 Carrying out daily routine; d430 Lifting and carrying objects; d445 Hand and arm use; d455 Moving around; d640 Doing housework; d750 Informal social relationships; d760 Family relationships; d850 Remunerative employment; d920 Recreation and leisure; nd-ph; nd-gh	48 (40%)	11 (28%)	59 (37%)
Patient-Rated Wrist/Hand Evaluation (PRWHE)	b280 Sensation of pain; d410 Changing basic body position; d430 Lifting and carrying objects; d440 Fine hand use; d445 Hand and arm use; d510 Washing oneself; d530 Toileting; d540 Dressing; d550 Eating; d640 Doing housework; d850 Remunerative employment; d920 Recreation and leisure	48 (40%)	11 (28%)	59 (37%)

Canadian Occupational Performance Measure	d420 Transferring oneself; d460 Moving around in different locations; d470 Using transportation; d510 Washing oneself; d520 Caring for body parts; d530 Toileting; d540 Dressing; d550 Eating; d560 Drinking; d620 Acquisition of goods and services; d850 Remunerative employment; d855 Non-remunerative employment; d630 Preparing meals; d640 Doing housework; d870 Economic self-sufficiency; d920 Recreation and leisure; nc-overall satisfaction	46 (38%)	11 (28%)	57 (36%)
Hospital Anxiety and Depression Scale (HADS)	b152 Emotional functions	20 (17%)	3 (7.7%)	23 (14%)
Barthel Index	d230 Carrying out daily routine; d410 Changing basic body position; d440 Fine hand use; d445 Hand and arm use; d450 Walking; d455 Moving around; d510 Washing oneself; d530 Toileting; d540 Dressing; d550 Eating	11 (9.1%)	3 (7.7%)	14 (8.8%)
Center for Epidemiologic Studies Depression Scale (CES-D)	b130 Energy and drive functions; b134 Sleep functions; b152 Emotional functions	12 (9.9%)	1 (2.6%)	13 (8.1%)
Dreiser’s Functional Hand Index	b280 Sensation of pain; b710 Mobility of joint functions; b730 Muscle power functions; d430 Lifting and carrying objects; d440 Fine hand use; d550 Eating	7 (5.8%)	1 (2.6%)	8 (5%)
General Self-Efficacy Scale (GSE)	b126 Temperament and personality functions; b152 Emotional functions	5 (4.1%)	1 (2.6%)	6 (3.8%)

Hand20	b126 Temperament and personality functions; b180 Experience of self and time functions; b280 Sensation of pain; d230 Carrying out daily routine; d430 Lifting and carrying objects; d440 Fine hand use; d445 Hand and arm use; d510 Washing oneself; d520 Caring for body parts; d540 Dressing; d550 Eating; d640 Doing housework; d920 Recreation and leisure	4 (3.3%)	1 (2.6%)	5 (3.1%)
Illness Perception Questionnaire	b126 Temperament and personality functions; b130 Energy and drive functions; b152 Emotional functions; b164 Higher-level cognitive functions; b180 Experience of self and time functions	4 (3.3%)	1 (2.6%)	5 (3.1%)
Lawton Instrumental Activities of Daily Living Scale (IADL)	d440 Fine hand use; d470 Using transportation; d570 Looking after one's health; d620 Acquisition of goods and services; d630 Preparing meals; d640 Doing housework; d860 Basic economic transactions	2 (1.7%)	1 (2.6%)	3 (1.9%)
Mayo Elbow Performance Score	b280 Sensation of pain; b710 Mobility of joint functions; b715 Stability of joint functions; d510 Washing oneself; d520 Caring for body parts; d540 Dressing	2 (1.7%)	1 (2.6%)	3 (1.9%)
Pain Catastrophizing Scale	b152 Emotional functions; b160 Thought functions; b164 Higher-level cognitive functions	3 (2.5%)	0 (0%)	3 (1.9%)
Patient-Specific Functional Scale	nc-individualised rating scale	1 (0.8%)	2 (5.1%)	3 (1.9%)

Patient Evaluation Measure (Hand Health Profile)	b152 Emotional functions; b180 Experience of self and time functions; b265 Touch function; b280 Sensation of pain; b710 Mobility of joint functions; b730 Muscle power functions; d230 Carrying out daily routine; d440 Fine hand use; d850 Remunerative employment	2 (1.7%)	1 (2.6%)	3 (1.9%)
Patient Health Questionnaire-2 (PHQ-2)	b152 Emotional functions	2 (1.7%)	1 (2.6%)	3 (1.9%)
Patient Health Questionnaire-9 (PHQ-9)	b130 Energy and drive functions; b152 Emotional functions	2 (1.7%)	0 (0%)	2 (1.3%)
Patient Perception of Patient Centeredness (PPPC)	b164 Higher-level cognitive functions; e355 Health professionals; e450 Individual attitudes of health professionals	1 (0.8%)	0 (0%)	1 (0.6%)
Pain Self-Efficacy Questionnaire (PSEQ)	b280 Sensation of pain; d230 Carrying out daily routine; d640 Doing housework; d850 Remunerative employment; d920 Recreation and leisure	1 (0.8%)	0 (0%)	1 (0.6%)
Rapid Assessment of Physical Activity (3 levels)	b455 Exercise tolerance functions; b710 Mobility of joint functions; b730 Muscle power functions	1 (0.8%)	0 (0%)	1 (0.6%)
Self-Perception Profile	b117 Intellectual functions; b122 Global psychosocial functions; b126 Temperament and personality functions; b140 Attention functions; b152 Emotional functions; b164 Higher-level cognitive functions; b180 Experience of self and time functions; b760 Control of voluntary movement	1 (0.8%)	0 (0%)	1 (0.6%)

	functions; d710 Basic interpersonal interactions; d720 Complex interpersonal interactions; d910 Community life; d920 Recreation and leisure; s730 Structure of upper extremity			
Self-Perception Profile for Adolescents – Global self-worth subscale only	b117 Intellectual functions; b122 Global psychosocial functions; b126 Temperament and personality functions; b140 Attention functions; b152 Emotional functions; b164 Higher-level cognitive functions; b180 Experience of self and time functions; d710 Basic interpersonal interactions; d720 Complex interpersonal interactions; d910 Community life	1 (0.8%)	0 (0%)	1 (0.6%)
Short Musculoskeletal Functional Assessment	b130 Energy and drive functions; b134 Sleep functions; b140 Attention functions; b144 Memory functions; b152 Emotional functions; b280 Sensation of pain; b640 Sexual functions; b710 Mobility of joint functions; d410 Changing basic body position; d440 Fine hand use; d450 Walking; d455 Moving around; d475 Driving; d520 Caring for body parts; d530 Toileting; d540 Dressing; d620 Acquisition of goods and services; d640 Doing housework; d650 Caring for household objects; b710 Mobility of joint functions; b730 Muscle power functions, d850 Remunerative employment; d920 Recreation and leisure; nd-ph	2 (1.7%)	0 (0%)	2 (1.3%)

TNO-AZN Adult Quality of Life questionnaire (TAAQoL)	b130 Energy and drive functions; b134 Sleep functions; b140 Attention functions; b144 Memory functions; b152 Emotional functions; b280 Sensation of pain; b640 Sexual functions; d230 Carrying out daily routine; d410 Changing basic body position; d430 Lifting and carrying objects; d440 Fine hand use; d450 Walking; d455 Moving around; d750 Informal social relationships; d850 Remunerative employment; d920 Recreation and leisure	1 (0.8%)	0 (0%)	1 (0.6%)
Visual analogue scale – aesthetic appearance	b180 Experience of self and time functions	1 (0.8%)	0 (0%)	1 (0.6%)
Visual analogue scale – function	d440 Fine hand use; d445 Hand and arm use	1 (0.8%)	0 (0%)	1 (0.6%)

nd-gh – not defined (general health) within WHO ICF

nd-ph – not defined (physical health) within WHO ICF

nc – not covered within WHO ICF

Appendix 4 WHO ICF outcome domains presented according to the proportion of verbatim outcomes to which they were linked, per outcome reporting status category

WHO ICF outcome domain	Outcome reporting status category						
	None (n=66)	Minimal (n=94)	Partial (n=454)	Complete (n=756)	Unexpected (n=310)	N/A (n=91)	Total (1,777)
b710 Mobility of joint functions	26 (39%)	9 (9.6%)	204 (45%)	293 (39%)	83 (27%)	0 (0.0%)	619 (35%)
b280 Sensation of pain	12 (18%)	11 (12%)	87 (19%)	195 (26%)	47 (15%)	0 (0.0%)	354 (20%)
s730 Structure of upper extremity	10 (15%)	36 (38%)	90 (20%)	130 (17%)	56 (18%)	0 (0.0%)	323 (18%)
b730 Muscle power functions	4 (6.1%)	7 (7.4%)	66 (15%)	157 (21%)	33 (11%)	0 (0.0%)	271 (15%)
d850 Remunerative employment	3 (4.5%)	5 (5.3%)	57 (13%)	116 (15%)	37 (12%)	0 (0.0%)	220 (12%)
d920 Recreation and leisure	3 (4.5%)	4 (4.3%)	58 (13%)	101 (13%)	25 (8.1%)	0 (0.0%)	193 (11%)
d510 Washing oneself	3 (4.5%)	4 (4.3%)	51 (11%)	95 (13%)	20 (6.5%)	0 (0.0%)	175 (9.8%)
d640 Doing housework	2 (3.0%)	1 (1.1%)	52 (11%)	93 (12%)	18 (5.8%)	0 (0.0%)	168 (9.5%)
d440 Fine hand use	2 (3.0%)	4 (4.3%)	54 (12%)	87 (12%)	17 (5.5%)	0 (0.0%)	165 (9.3%)
d230 Carrying out daily routine	4 (6.1%)	5 (5.3%)	37 (8.1%)	89 (12%)	24 (7.7%)	0 (0.0%)	160 (9.0%)
d430 Lifting and carrying objects	2 (3.0%)	1 (1.1%)	49 (11%)	88 (12%)	18 (5.8%)	0 (0.0%)	160 (9.0%)
d445 Hand and arm use	4 (6.1%)	2 (2.1%)	45 (9.9%)	82 (11%)	19 (6.1%)	0 (0.0%)	153 (8.6%)
d540 Dressing	3 (4.5%)	4 (4.3%)	40 (8.8%)	80 (11%)	17 (5.5%)	0 (0.0%)	146 (8.2%)
d550 Eating	1 (1.5%)	1 (1.1%)	44 (9.7%)	82 (11%)	13 (4.2%)	0 (0.0%)	142 (8.0%)
b265 Touch function	0 (0.0%)	15 (16%)	41 (9.0%)	64 (8.5%)	12 (3.9%)	0 (0.0%)	133 (7.5%)

b134 Sleep functions	0 (0.0%)	1 (1.1%)	35 (7.7%)	55 (7.3%)	11 (3.5%)	0 (0.0%)	103 (5.8%)
nc-Complications/Adverse events	0 (0.0%)	0 (0.0%)	3 (0.7%)	3 (0.4%)	2 (0.6%)	91 (100%)	99 (5.6%)
b126 Temperament and personality functions	2 (3.0%)	1 (1.1%)	28 (6.2%)	49 (6.5%)	13 (4.2%)	0 (0.0%)	95 (5.3%)
d750 Informal social relationships	2 (3.0%)	1 (1.1%)	29 (6.4%)	51 (6.7%)	10 (3.2%)	0 (0.0%)	94 (5.3%)
d760 Family relationships	2 (3.0%)	1 (1.1%)	28 (6.2%)	51 (6.7%)	10 (3.2%)	0 (0.0%)	93 (5.2%)
e580 Health services, systems and policies	4 (6.1%)	1 (1.1%)	1 (0.2%)	27 (3.6%)	33 (11%)	0 (0.0%)	66 (3.7%)
b152 Emotional functions	4 (6.1%)	3 (3.2%)	17 (3.7%)	29 (3.8%)	11 (3.5%)	0 (0.0%)	66 (3.7%)
d470 Using transportation	0 (0.0%)	1 (1.1%)	19 (4.2%)	38 (5.0%)	6 (1.9%)	0 (0.0%)	64 (3.6%)
b640 Sexual functions	0 (0.0%)	1 (1.1%)	19 (4.2%)	33 (4.4%)	6 (1.9%)	0 (0.0%)	59 (3.3%)
d650 Caring for household objects	0 (0.0%)	1 (1.1%)	18 (4.0%)	33 (4.4%)	6 (1.9%)	0 (0.0%)	58 (3.3%)
nc-Overall satisfaction	3 (4.5%)	0 (0.0%)	8 (1.8%)	30 (4.0%)	13 (4.2%)	0 (0.0%)	55 (3.1%)
d410 Changing basic body position	1 (1.5%)	0 (0.0%)	16 (3.5%)	29 (3.8%)	7 (2.3%)	0 (0.0%)	54 (3.0%)
d530 Toileting	1 (1.5%)	0 (0.0%)	16 (3.5%)	29 (3.8%)	5 (1.6%)	0 (0.0%)	51 (2.9%)
nc-Bone healing	1 (1.5%)	4 (4.3%)	14 (3.1%)	15 (2.0%)	7 (2.3%)	0 (0.0%)	41 (2.3%)
d450 Walking	2 (3.0%)	3 (3.2%)	5 (1.1%)	12 (1.6%)	5 (1.6%)	0 (0.0%)	28 (1.6%)
nd-gh (general health)	2 (3.0%)	2 (2.1%)	3 (0.7%)	15 (2.0%)	5 (1.6%)	0 (0.0%)	28 (1.6%)
nd-ph (physical health)	2 (3.0%)	0 (0.0%)	4 (0.9%)	12 (1.6%)	7 (2.3%)	0 (0.0%)	26 (1.5%)
b180 Experience of self and time functions	1 (1.5%)	0 (0.0%)	10 (2.2%)	10 (1.3%)	4 (1.3%)	0 (0.0%)	26 (1.5%)
b130 Energy and drive functions	2 (3.0%)	0 (0.0%)	6 (1.3%)	9 (1.2%)	4 (1.3%)	0 (0.0%)	22 (1.2%)
nc-Technical (related to intervention)	2 (3.0%)	1 (1.1%)	3 (0.7%)	9 (1.2%)	5 (1.6%)	0 (0.0%)	20 (1.1%)
nc-Bone healing time	2 (3.0%)	1 (1.1%)	2 (0.4%)	10 (1.3%)	4 (1.3%)	0 (0.0%)	19 (1.1%)
d455 Moving around	2 (3.0%)	0 (0.0%)	6 (1.3%)	7 (0.9%)	2 (0.6%)	0 (0.0%)	18 (1.0%)

b760 Control of voluntary movement functions	2 (3.0%)	2 (2.1%)	9 (2.0%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	17 (1.0%)
b270 Sensory functions related to temperature and other stimuli	1 (1.5%)	5 (5.3%)	1 (0.2%)	8 (1.1%)	1 (0.3%)	0 (0.0%)	16 (0.9%)
b820 Repair functions of the skin	1 (1.5%)	0 (0.0%)	6 (1.3%)	4 (0.5%)	4 (1.3%)	0 (0.0%)	15 (0.8%)
b289 Sensation of pain, other specified and unspecified	1 (1.5%)	0 (0.0%)	1 (0.2%)	8 (1.1%)	2 (0.6%)	0 (0.0%)	12 (0.7%)
b455 Exercise tolerance functions	1 (1.5%)	0 (0.0%)	2 (0.4%)	5 (0.7%)	2 (0.6%)	0 (0.0%)	12 (0.7%)
d630 Preparing meals	0 (0.0%)	0 (0.0%)	4 (0.9%)	7 (0.9%)	1 (0.3%)	0 (0.0%)	12 (0.7%)
d839 Education unspecified	0 (0.0%)	0 (0.0%)	4 (0.9%)	6 (0.8%)	1 (0.3%)	0 (0.0%)	12 (0.7%)
nc-Satisfaction with intervention	0 (0.0%)	0 (0.0%)	1 (0.2%)	0 (0.0%)	10 (3.2%)	0 (0.0%)	11 (0.6%)
b830 Other functions of the skin	1 (1.5%)	0 (0.0%)	1 (0.2%)	8 (1.1%)	1 (0.3%)	0 (0.0%)	11 (0.6%)
d520 Caring for body parts	1 (1.5%)	0 (0.0%)	4 (0.9%)	6 (0.8%)	0 (0.0%)	0 (0.0%)	11 (0.6%)
b164 Higher-level cognitive functions	0 (0.0%)	0 (0.0%)	5 (1.1%)	3 (0.4%)	1 (0.3%)	0 (0.0%)	9 (0.5%)
d560 Drinking	1 (1.5%)	0 (0.0%)	3 (0.7%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	8 (0.5%)
d620 Acquisition of goods and services	0 (0.0%)	0 (0.0%)	3 (0.7%)	5 (0.7%)	0 (0.0%)	0 (0.0%)	8 (0.5%)
s810 Structure of areas of skin	1 (1.5%)	0 (0.0%)	1 (0.2%)	6 (0.8%)	0 (0.0%)	0 (0.0%)	8 (0.5%)
b140 Attention functions	0 (0.0%)	0 (0.0%)	5 (1.1%)	1 (0.1%)	0 (0.0%)	0 (0.0%)	6 (0.3%)
d420 Transferring oneself	0 (0.0%)	0 (0.0%)	2 (0.4%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	6 (0.3%)
d460 Moving around in different locations	0 (0.0%)	0 (0.0%)	2 (0.4%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	6 (0.3%)
d855 Non-remunerative employment	0 (0.0%)	0 (0.0%)	2 (0.4%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	6 (0.3%)
d870 Economic self-sufficiency	0 (0.0%)	0 (0.0%)	2 (0.4%)	4 (0.5%)	0 (0.0%)	0 (0.0%)	6 (0.3%)
nc-Blood tests	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (1.3%)	0 (0.0%)	4 (0.2%)
b715 Stability of joint functions	0 (0.0%)	0 (0.0%)	1 (0.2%)	3 (0.4%)	0 (0.0%)	0 (0.0%)	4 (0.2%)

d720 Complex interpersonal interactions	0 (0.0%)	0 (0.0%)	4 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (0.2%)
d845 Acquiring, keeping and terminating a job	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (0.3%)	1 (0.3%)	0 (0.0%)	3 (0.2%)
b117 Intellectual functions	0 (0.0%)	0 (0.0%)	3 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
b122 Global psychosocial functions	0 (0.0%)	0 (0.0%)	3 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
b160 Thought functions	0 (0.0%)	0 (0.0%)	1 (0.2%)	2 (0.3%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
b260 Proprioceptive function	0 (0.0%)	0 (0.0%)	3 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
d570 Looking after one's health	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.4%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
d710 Basic interpersonal interactions	0 (0.0%)	0 (0.0%)	3 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
d910 Community life	0 (0.0%)	0 (0.0%)	3 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (0.2%)
d475 Driving	0 (0.0%)	0 (0.0%)	1 (0.2%)	0 (0.0%)	1 (0.3%)	0 (0.0%)	2 (0.1%)
e565 Economic services, systems and policies	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)	1 (0.3%)	0 (0.0%)	2 (0.1%)
b144 Memory functions	0 (0.0%)	0 (0.0%)	2 (0.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (0.1%)
b156 Perceptual functions	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)	0 (0.0%)	0 (0.0%)	1 (0.1%)
d860 Basic economic transactions	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)	0 (0.0%)	0 (0.0%)	1 (0.1%)
nd-Patient adherence to treatment	1 (1.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)
nc-Individualised rating scale	1 (1.5%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.1%)

WHO ICF – World Health Organization International Classification of Functioning, Disability and Health

nd – not definable within WHO ICF

nc – not covered within WHO ICF

N/A – this category was reserved for complication/adverse event outcomes where not specifically mentioned in a study registration or methods. Rather than label these as ‘unexpected’ it was felt that this particular outcome group is widely assumed to be collected as part of a clinical intervention study. Where specific complications were specified as an outcome though, then their reporting status was categorised as with other verbatim outcomes

Appendix 5 Figure showing links formed between patient journey themes and outcome domains

Outcome domains	**Mechanism of injury**	**Delay in seeking medical attention**	**Attitudes regarding injury circumstances**	**Information from healthcare professionals**	**Treatment options**
Change in sensation					
Pain / discomfort		🟩		🟩	
Strength or muscle power					
Range of movement				🟩	
Confirmation of bone healing and alignment					
Confirmation of soft tissue healing					
Eating and drinking					
Self-hygiene and personal care					
Cooking / preparing meals					
Grasping and moving light objects					
Walking / mobility			🟩		
Using hands to change body position					
General household chores					
Shopping for groceries					
Sleeping					
Dressing					
Return to work / job					
Return to driving			🟩		
Fine hand use					
Hobbies and recreation	🟩		🟩		
Socialising / relationships with friends and family					
Carrying heavy objects					
Sexual activity					
Looking after dependents within household					
Keyboard and mouse / touchscreen use					
Fatigue / tiredness					
Emotional / mood impact to self			🟩	🟩	
Emotional / mood impact to friends and family					
Overall patient satisfaction				🟩	
Appearance					
Loss of income					
Added expenses due to injury					
Healthcare resources used		🟩			🟩
Treatment convenience / comfort					🟩
Treatment side effects or complications				🟩	🟩

Outcome domains	**Factors influencing treatment selection**	**Practical implications**	**Changes in social/family role and relationships**	**Detriment to independence and sense of vulnerability/loss of confidence**	**Priorities during recovery (early vs later phases)**
Change in sensation		🟩			
Pain / discomfort	🟩	🟩			🟩
Strength or muscle power		🟩		🟩	🟩
Range of movement	🟩	🟩			🟩
Confirmation of bone healing and alignment	🟩				🟩
Confirmation of soft tissue healing	🟩	🟩			
Eating and drinking		🟩		🟩	
Self-hygiene and personal care	🟩	🟩	🟩	🟩	🟩
Cooking / preparing meals		🟩	🟩	🟩	🟩
Grasping and moving light objects		🟩		🟩	🟩
Walking / mobility				🟩	
Using hands to change body position		🟩		🟩	
General household chores		🟩	🟩	🟩	
Shopping for groceries		🟩		🟩	
Sleeping		🟩			
Dressing		🟩		🟩	🟩
Return to work / job	🟩	🟩	🟩		
Return to driving	🟩	🟩	🟩	🟩	
Fine hand use		🟩			🟩
Hobbies and recreation	🟩	🟩	🟩	🟩	🟩
Socialising / relationships with friends and family		🟩	🟩	🟩	
Carrying heavy objects		🟩			🟩
Sexual activity			🟩		
Looking after dependents within household		🟩	🟩		
Keyboard and mouse / touchscreen use		🟩			
Fatigue / tiredness		🟩			
Emotional / mood impact to self	🟩	🟩	🟩	🟩	🟩
Emotional / mood impact to friends and family			🟩		
Overall patient satisfaction	🟩				
Appearance	🟩		🟩		
Loss of income	🟩	🟩	🟩		
Added expenses due to injury		🟩			
Healthcare resources used					
Treatment convenience / comfort	🟩	🟩			
Treatment side effects or complications	🟩				

Appendix 6 All initial codes generated during analysis of interview and focus group transcripts

- Acceptance of current state
- Accessibility to further information from healthcare professionals
- Acknowledging different perspectives and injuries
- Adapting to bypass altered function
- Adjunctive therapies
- Altered temperature sensation
- Anaesthetic regional block
- Anger
- Annoyed or irritated
- Anxiety or stress
- Appearance of treatment
- Appetite
- Applying make-up
- Applying moisturisers or creams
- Applying perfume
- Applying toothpaste to brush
- Availability of health services out-of-hours
- Avoiding water damaging splint or cast
- Awareness of NHS research or clinical research
- Bathing child(ren)
- Bearing weight on injured body part
- Further physio input desired
- Gardening
- Gauging own recovery
- General clinician time use
- General confidence
- General cooking
- General cutlery use to eat
- General dressing or undressing of oneself
- General feelings about research
- General function
- General impression of treatment
- General mobility impact
- General wrist movement range or stiffness
- Getting child(ren) ready for bed
- Grasping and moving objects
- Gratitude for healthcare received
- Gratitude for support from others
- Grip strength
- Guilt – general
- Guilt over doing more than medically advised
- Guilt over frustration at offer of help from others
- Pain-free
- Peeling vegetables
- Picking up heavy object
- Pinch strength
- Pins and needles sensation
- Placing items into shopping basket or trolley
- Playing piano
- Playing video or computer games
- Playing with child(ren)
- Playing with pet(s)
- Post-op infection
- Post-traumatic arthritis
- Pre-injury emotional state
- Previous injury
- Pronation-supination strength
- Public transport – work
- Pushing shopping trolley
- Putting on or taking off shirt
- Putting on or taking off socks
- Reassurance or advice from fellow patients
- Reduced income

Bone healing	Guilt over letting down work or colleagues	Regret
Boredom	Guilt over reliance on others	Relationship with spouse or partner
Brushing teeth	Hand dominance	Relationship with wider family
Brushing teeth of child(ren)	Hand dominance impact to function	Reliance on others
Buying new clothing or footwear due to functional impairment	Having stitches removed	Reliance on spouse
Carer role – looking after child(ren)	Having to attend hospital	Relief
Carer role – looking after grandchildren	Helping family member with garden work	Research interest through work activity
Carer role – looking after parent(s)	Helping parent(s) transfer to toilet or commode	Research relevant to health condition in family
Carer role – looking after pet(s)	Hobbies and recreation	Research relevant to own health condition
Carer role impact	Hockey	Rock-climbing
Caring for unwell family member	Holding a drink	Rowing
Carrying grocery items	Holding bin bag(s)	Rugby
Carrying heavy objects	Holding briefcase or suitcase or bag	Running
Carrying heavy paperwork or books	Holding filled bucket	Sauna
Change in body position	Holding filled cup	Scarring
Change in priorities at differing phases of recovery	Holding filled kettle	Seeking medical attention or delay
Change in sensation	Holding filled saucepan	Sexual activity
Changing child's nappy	Holding hairdryer	Shaking hands – socialising
Children emotional impact	Holding laptop	Shaving
Clothing sleeves not fitting	Holding light object – fruit or vegetable	Shoes and shoelaces
Combing or brushing or blowdrying hair	Holding plates or trays	Shopping
Comparing wrist and hand injuries	Holding shopping bag(s)	Sick pay
Concern about appearance of altered function to others	Hopefulness	Skateboarding

Concern about appearance of treatment to others	Household cleaning	Sleep disturbance due to pain
Concern about deterioration with age or long-term deficit	Household repair or maintenance	Sleeping
Concern about future injury	Hygiene or maintenance of splint or cast	Slip or trip
Concern about healing	Hyperaesthesia	Snowboarding
Concern over angering spouse	Importance of patient perspective in research	Socialising with non-family
Concern over worsening the injury	Information from healthcare professionals	Socialising with wider family
Confidence with mobilising	Injury sustained	Soft tissue healing
Conflicting information from healthcare professionals	Keeping arm elevated while sleeping	Splint or cast hindering function
Continuity of care was an issue	Keyboard use	Splint or cast hindering movement
Convenience in taking part in research	Kickboxing	Spouse or partner emotional impact
Cooking for family	Lifting bag of potatoes	Stand up from seated position
Coping with ongoing or persisting issues	Lifting child(ren)	Stirring a drink
Coronavirus-related	Lifting easel	Surgical wound
Cosmesis	Lifting fence panels	Surprise or shock
Cross-stitching or sewing	Lifting filled container	Swelling
Cutting food	Lifting hedgecutter	Swimming
Cutting own nails	Lifting iron	Taking parent(s) for health appointment
Cutting vegetables or meat	Lifting pint of milk	Teaching child(ren) how to write
Cycling	Lifting plants and plant pots	Thoughts about initial cause of injury
Dealing with rubbish and bins	Lifting weights	Thumb flexion-extension range or stiffness
Deformity	Lifting wheels or tyres	Time off work to attend hospital
Deformity of finger	Low mood	Time to return to normal day-to-day life
Deformity of knuckle	Maintaining dignity	Time to return to work
Deformity of thumb	Making bed	Tingling sensation

Deformity of wrist
Desire to maintain or regain independence
Determined
Diagnostic process
Difficulty finding comfortable sleep position
Discolouration or bruising
Discomfort or ache
Drawing
Dressing child(ren)
Dressing parent(s)
Driving
Driving child(ren)
Driving parent(s)
Dry skin
Dysaesthesia
Eating ready meals
Economic impact
Emotional impact
Enjoyed receiving support from others
Exercising at gym
Expectation of recovery speed
Extra cost for equipment or accessories related to treatment
Extra costs for grooming or personal hygiene
Extra costs for meals
Making hot drink for parent(s)
Malunion of fracture
Manual dexterity or co-ordination
Mechanism of injury
Memorable quote
Minimisation of the injury
Motivation
Movement range and stiffness
Moving object – table
Moving object – wheelbarrow
Moving object – wheelchair
Nausea
Non-injury related sleep disturbance
Non-injury site ache or strain from compensatory activity
Non-injury site pain
Non-union of fracture
Numbness
Occupation
Opening containers
Opening or closing door(s)
Opening toothpaste tube
Operative treatment cost
Optimism
Outcome time points
Tiredness or fatigue
Toileting
Touchscreen use
Travel – general
Travel – social events
Travel – work
Treatment – splint or cast
Treatment – surgery
Treatment options
Treatment risks
Treatment side effects or complications
Use of computer mouse
Use of healthcare resources
Using belt buckle
Using buttons on clothing
Using zips on clothing
Volunteer work
Vulnerability or loss of control
Walking
Walking – work
Walking dog
Washing dishes
Washing face
Washing hands

Extra costs for transport	Overall appearance of injured body part	Washing yourself
Factors against taking part in research	Pain	Washing yourself – shower
Factors influencing participation in research	Pain – analgesia required	Weights machine
Factors influencing questionnaire completion	Pain – concern of long-term pain	Wider family (outside household) emotional impact
Factors influencing treatment selection	Pain – dull	Work impact
Fastening or unfastening bra	Pain – giving way or not bearing weight	Working longer hours
Fastening travel restraint or seatbelt	Pain – mild	Working more slowly
Feeding child(ren)	Pain – moderate	Wrist circumduction range or stiffness
Feelings about how injury was sustained	Pain – severe	Wrist flexion-extension range or stiffness
Fine hand use	Pain – sharp	Wrist pronation-supination range or stiffness
Finger flexion-extension range or stiffness	Pain – shocklike	Wrist radial-ulnar deviation range or stiffness
Folding clothes or doing laundry	Pain – shooting	Wrist strength
Football	Pain – stabbing	Writing
Friends emotional impact	Pain – throbbing	X-ray appearance
Frightened or scared	Pain – uncertain what is safe level while recovering	Yoga
Frightened or scared of returning to area or activity of injury	Pain control provided	
Frustration	Pain on moving injured body part	

Appendix 7 First round Delphi items based on qualitative study and systematic review outcome domains

Delphi outcome item	**Qualitative study outcome domain**	**Systematic review WHO ICF outcome domain**
Change in sensation	Change in sensation	b265 Touch function b270 Sensory functions related to temperature and other stimuli b260 Proprioceptive function b156 Perceptual functions
Pain / discomfort	Pain / discomfort	b280 Sensation of pain b289 Sensation of pain, other specified and unspecified
Strength or muscle power	Strength or muscle power	b730 Muscle power functions b455 Exercise tolerance functions
Range of movement	Range of movement	b710 Mobility of joint functions
Confirmation of bone healing and alignment (fractures only)	Confirmation of bone healing and alignment (fractures only)	nc-Bone healing s730 Structure of upper extremity nc-Bone healing time

Confirmation of soft tissue healing	Confirmation of soft tissue healing	b820 Repair functions of the skin
Stability of joints		b715 Stability of joint functions
Grasping and moving light objects	Grasping and moving light objects	d445 Hand and arm use
Self-hygiene / personal care	Self-hygiene / personal care	d510 Washing oneself d530 Toileting d520 Caring for body parts d230 Carrying out daily routine nd-ph (physical health) d570 Looking after one's health
Cooking / preparing meals	Cooking / preparing meals	d630 Preparing meals
Dressing	Dressing	d540 Dressing
Eating and drinking	Eating and drinking	d550 Eating d560 Drinking

Shopping for groceries	Shopping for groceries	d620 Acquisition of goods and services d860 Basic economic transactions
Walking / mobility	Walking / mobility	d450 Walking d455 Moving around d460 Moving around in different locations
Using hands to change body position	Using hands to change body position	d410 Changing basic body position d420 Transferring oneself
General household chores	General household chores	d640 Doing housework d650 Caring for household objects
Sleeping	Sleeping	b134 Sleep functions
Thinking and memory		b160 Thought functions b144 Memory functions b117 Intellectual functions b164 Higher-level cognitive functions b140 Attention functions

		d839 Education unspecified b180 Experience of self and time functions
Fine hand use **/ dexterity**	Fine hand use	d440 Fine hand use b760 Control of voluntary movement functions
Hobbies and recreation	Hobbies and recreation	d920 Recreation and leisure
Return to **usual** work / job	Return to work / job	d850 Remunerative employment d855 Non-remunerative employment d845 Acquiring, keeping and terminating a job
Carrying or lifting heavy objects	Carrying or lifting heavy objects	d430 Lifting and carrying objects
Being able to control a vehicle for transport	Return to driving	d475 Driving d470 Using transportation
Keyboard and mouse / touchscreen use	Keyboard and mouse / touchscreen use	

Socialising / relationships with friends and family	Socialising / relationships with friends and family	d760 Family relationships d750 Informal social relationships d710 Basic interpersonal interactions d720 Complex interpersonal interactions b122 Global psychosocial functions d910 Community life
Fatigue / tiredness	Fatigue / tiredness	b130 Energy and drive functions
Looking after dependents within household	Looking after dependents within household	
Intimate / sexual activity	Sexual activity	b640 Sexual functions
Appearance	Appearance	s810 Structure of areas of skin
Emotional / mood impact to self	Emotional / mood impact to self	b152 Emotional functions b126 Temperament and personality functions nd-gh (general health)

Emotional / mood impact to friends and family	Emotional / mood impact to friends and family	
Overall patient satisfaction	Overall patient satisfaction	nc-Overall satisfaction nc-Satisfaction with intervention
Added expenses due to injury	Added expenses due to injury	
Loss of income	Loss of income	d870 Economic self-sufficiency e565 Economic services, systems and policies
Healthcare resources used	Healthcare resources used	e580 Health services, systems and policies
Treatment complications	Treatment complications	nc-Complications/Adverse events b830 Other functions of the skin
Treatment inconvenience / discomfort	Treatment inconvenience / discomfort	nd-Patient adherence to treatment

In the Delphi outcome item column, specific changes suggested through stakeholder feedback on wording are marked in bold and underlined.

Appendix 8 Record of additional outcome items suggested at end of first Delphi round with decision and rationale regarding inclusion for second round

Key: New domain added Example added to existing domain descriptor Not added

Suggested additional outcome	**Decision**	**Rationale for decision**
Pain during daytime activities	**Added domain 'pain/discomfort during activity'** Descriptor: Discomfort or pain in the hand or wrist specifically during activities (NOT at rest) e.g. ache, shooting pain, sharp pain, throbbing, discomfort/pain due to not being able to tolerate hot or cold sensation **&** **Added domain 'pain/discomfort during rest'** Descriptor: Discomfort or pain in the hand or wrist specifically during rest (i.e. with the hand/wrist not moving, so NOT during activities) e.g. ache, shooting pain, sharp pain, throbbing, discomfort/pain due to not being able to tolerate hot or cold sensation	This highlights a divide in the fairly broad 'pain/discomfort' domain and might help to highlight whether subset of activity vs rest is important when considering the 'pain/discomfort' domain

Movement velocity	**Added domain 'speed of movement'** Descriptor: How quick movements through the joints of the hand or wrist can be, including any issue of delay in being able to start the movement	In some ways movement velocity bridges the concepts of range of movement and strength/muscle power. However this, alongside initiation/hesitation, has its own elements and so has been added as a novel domain
Movement-initiation hesitation		
Satisfaction with the outcome	**Added domain 'patient satisfaction with outcome/result'** Descriptor: Satisfaction with the overall result from the patient's perspective (NOT with treatment or recovery process, but the end result only)	This is a subset of the 'overall patient satisfaction' domain but has been mentioned by more than one participant. It might be considered the key long-term aspect of 'overall patient satisfaction' as the result is what patients live with whereas the treatment and recovery phases (though of variable length) do eventually conclude
Are you (patient) happy with the final result?		
How your family and friends perceive your injured hand or digit	Added to 'appearance' domain descriptor	The research group concluded that these suggestions were mainly to do with what other people thought of the appearance of the patient's injury. This could be considered part of the existing 'appearance' domain if assuming that part of a person's perception of their own appearance is influenced by how they believe others view them
How strangers perceive your injured hand or digit		

Information provision of possible issues and what to do during the healing process	Added to 'overall patient satisfaction' domain descriptor as an example	Information provision is mentioned multiple times and is clearly important, but is not an outcome itself. Added to specified domain descriptor as it is an aspect of the treatment and recovery process
Information on suitable hand exercises when in a cast		
Interim timescales for regaining mobility/swelling/appearance/strength would be helpful and help to incentivise exercise without overdoing it.		
More information on what to expect regarding pain; weakness etc before leaving hospital		
Patients should be informed exactly what kind of injury they have and treatment/complications that could occur		
Hospital Outpatient visits (frequency; departments; waiting time)	Added to 'treatment inconvenience/discomfort' domain descriptor as an example	In terms of the healthcare resource perspective this is covered by the 'healthcare resources used' domain. However this was submitted by a patient and therefore it is the research group's opinion that

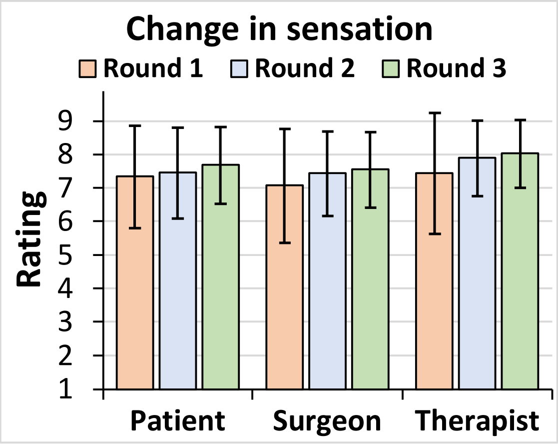
		this was meant more from an inconvenience perspective
Cold intolerance	Added to 'pain/discomfort' domain (and newly added pain domains) descriptor as an example	This is considered a specific example within the broader existing 'pain/discomfort' domain
Propriocepsis (especially in DRF)	Added to 'change in sensation' domain descriptor as an example	This is considered a specific example within the broader existing 'change in sensation' domain
Oedema	Added as specific example within the 'confirmation of soft tissue healing' domain	The example of swelling is within the 'appearance' domain, but we believe this refers to the sign as an outcome from the professional perspective as an indicator of soft tissue damage and healing
Grip and Pinch strength	Not added	This forms much of the 'strength or muscle power' domain and therefore little would be gained by creating an additional domain based on these
Participating in social activities with friends/family	Not added	This is covered within an existing domain 'socialising/relationships' with friends and family

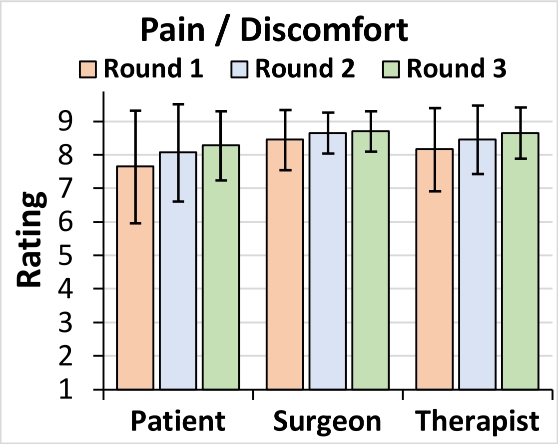
Fear of movement	Not added	This appears to refer to existing domains such as ‘emotional impact of the injury’ and perhaps ‘overall patient satisfaction’ (through the specific example of information given about safe movement and exercise during treatment and recovery)
Assess independence / dependence of others; (family; friends; including therapists) during the recovery process. Patients that rely 100% on the therapist for the recovery; tend to obtain lower outcome scores	Not added	The research group considers this as a ‘how to measure’ aspect of multiple existing domains. For each one, as part of the outcome the dependence on others cold also be assessed. As such, it is not considered a novel domain in its own right
Assess if the patient is the main economical provider of the house and has a formal job. Not being able to work can provide extra stress or interfere with the correct healing phases	Not added	Though this is clearly of relevance to patients with these injuries and the healthcare professionals treating them, it is a demographic characteristic rather than an outcome

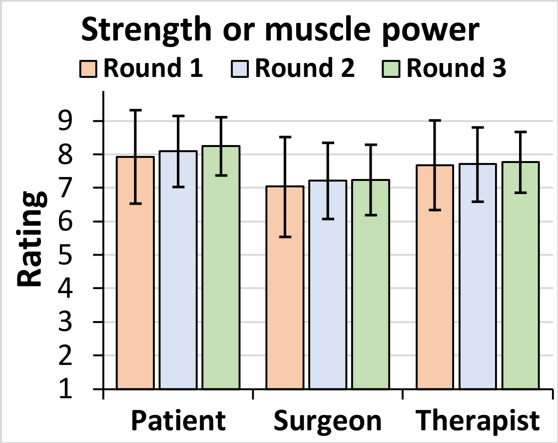
Functional use in/out of a splint/support	Not added	In the research group’s opinion, this suggestion reflects on ‘how to measure’ rather than ‘what to measure’. The ‘what to measure’ appears to fall under the existing domains of function (of which there are several) or potentially in the existing domain ‘treatment inconvenience/discomfort’
Caring for others	Not added	In the research group’s opinion this falls within the existing domain ‘looking after dependents within household’
I'd like to see more specific pain related questions eg type; timeframe; "Rate your pain" gives so little information	Not added	In the research group’s opinion, this suggestion reflects on ‘how to measure’ rather than ‘what to measure’
Xenomelia (When patient doesn’t want to associate a hand as a part of their body) for mangled hands	Not added	This is an outcome for injuries outside the scope of this COS and therefore has not been added
Surgery would have been better as first line of treatment as bones moved .whilst in a cast	Not added	This is not an outcome, but instead refers to existing domains such as ‘treatment complications’ and/or ‘confirmation of bone healing and alignment’

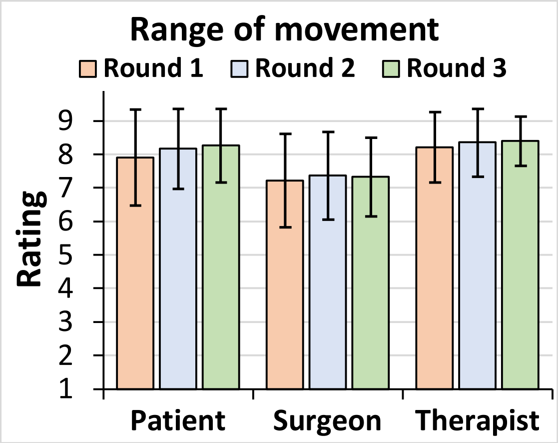
Bilateral injuries	Not added	This is a characteristic of the injury and whilst it may be of value to record when measuring outcomes, due to the impact on various outcomes compared to unilateral injuries, it is not itself an outcome

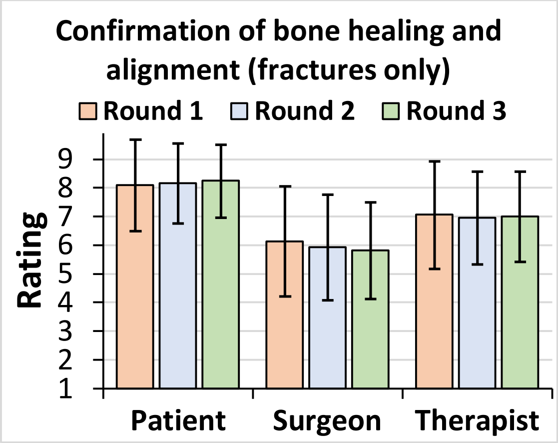
Appendix 9 Figures summarising Delphi survey mean ratings and standard deviations per stakeholder group across rounds

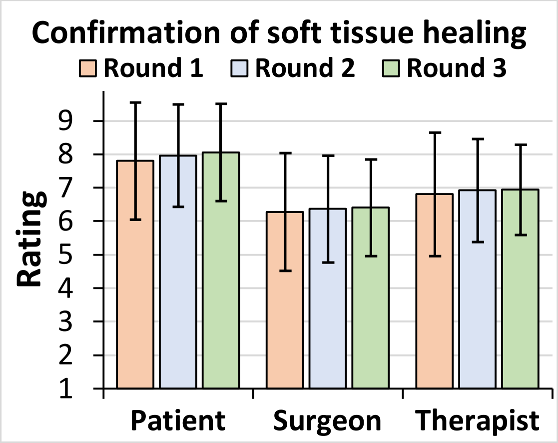

Change in sensation

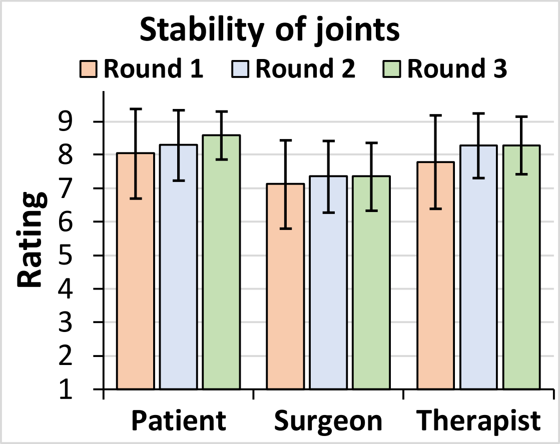

Pain / Discomfort

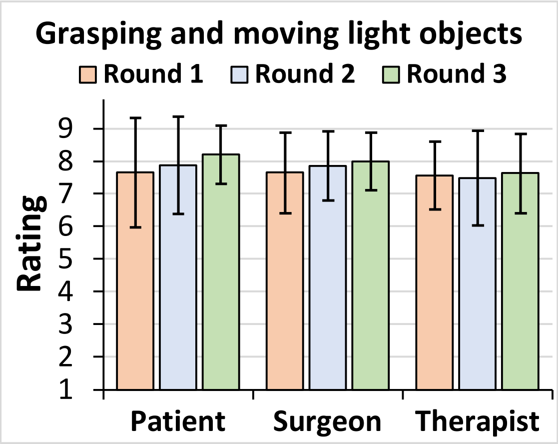

Strength or muscle power

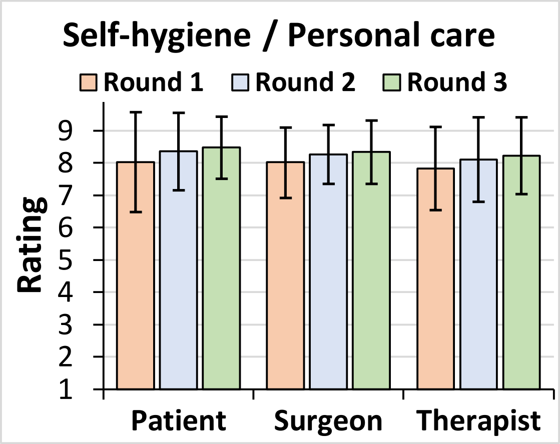

Range of movement

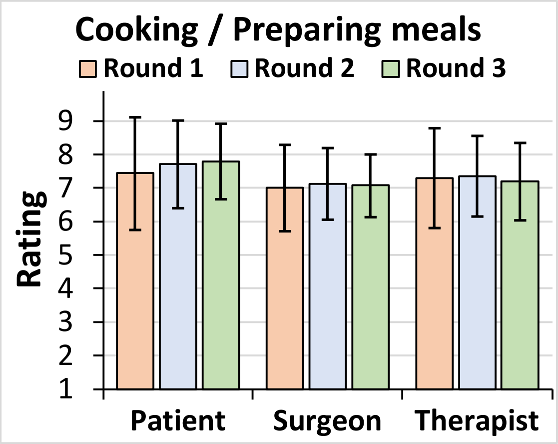

Confirmation of bone healing and alignment (fractures only)

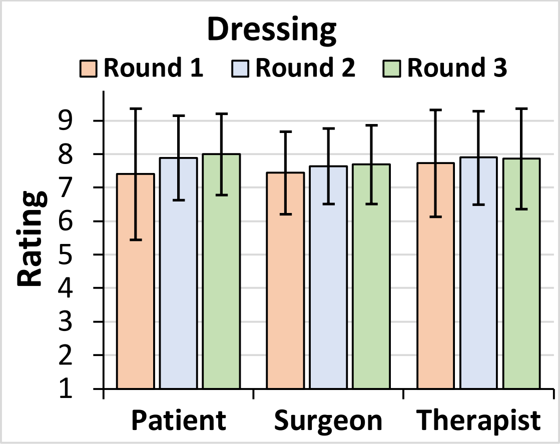

Confirmation of soft tissue healing

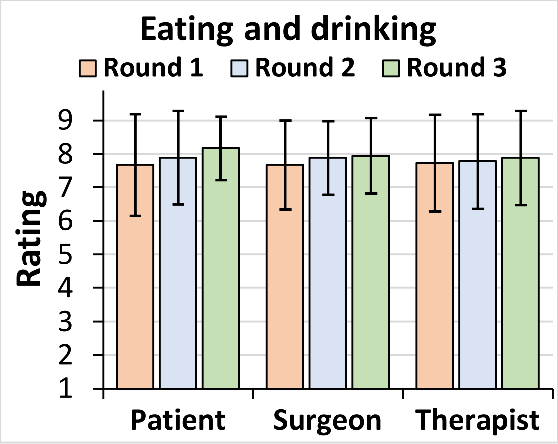

Stability of joints
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

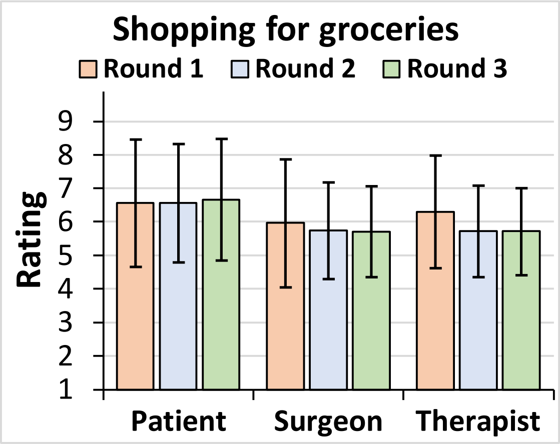

Grasping and moving light objects
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

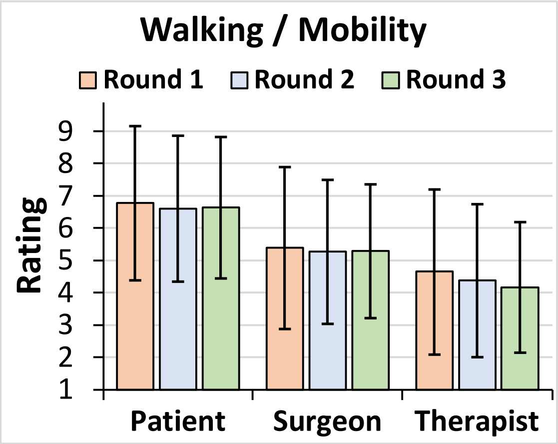

Self-hygiene / Personal care
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

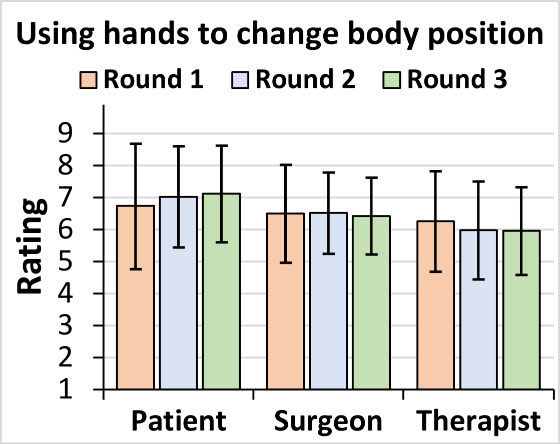

Cooking / Preparing meals
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

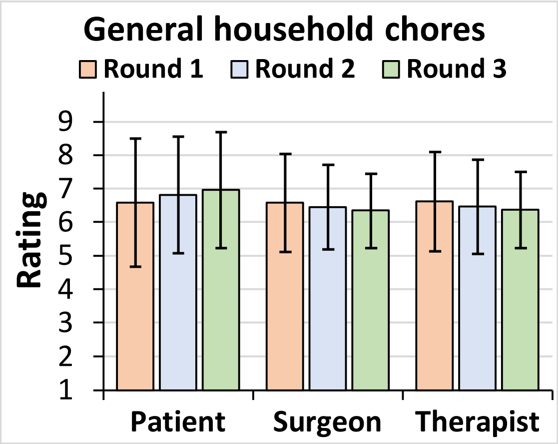

Dressing
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

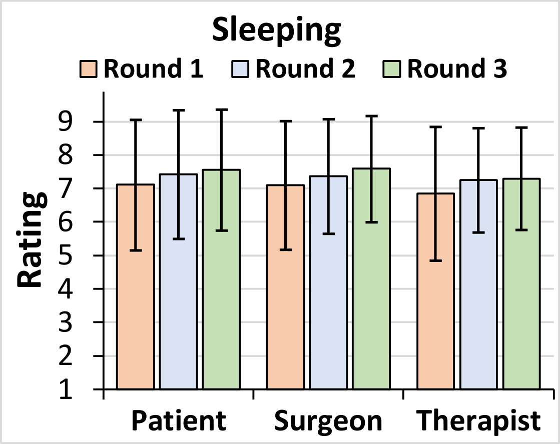

Eating and drinking
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

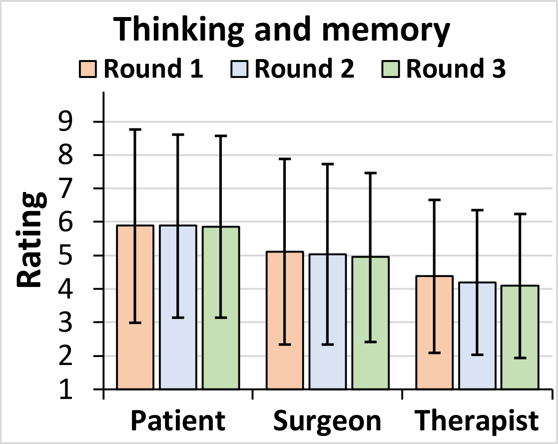

Shopping for groceries
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

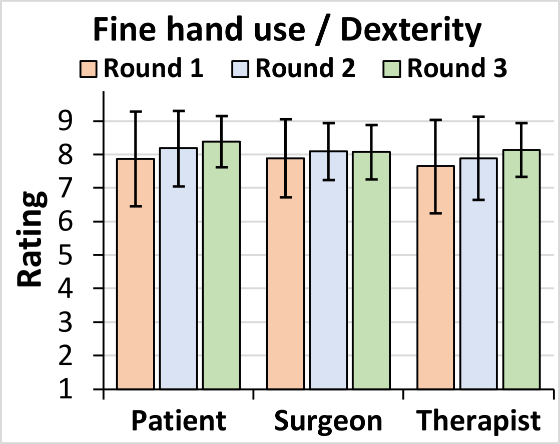

Walking / Mobility
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

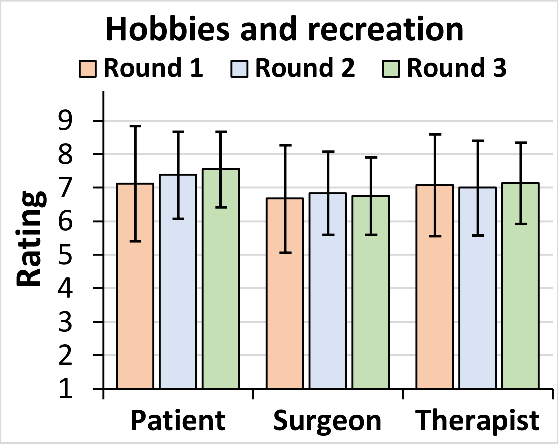

Using hands to change body position

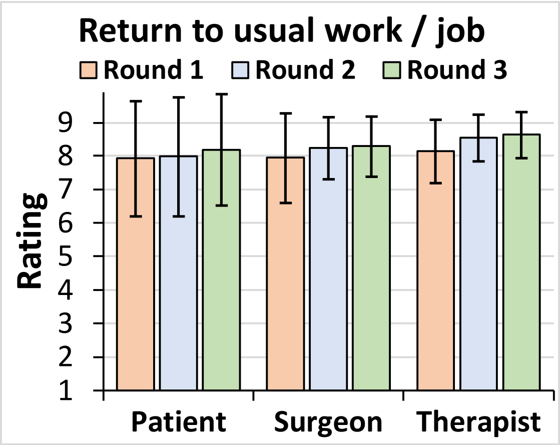

General household chores

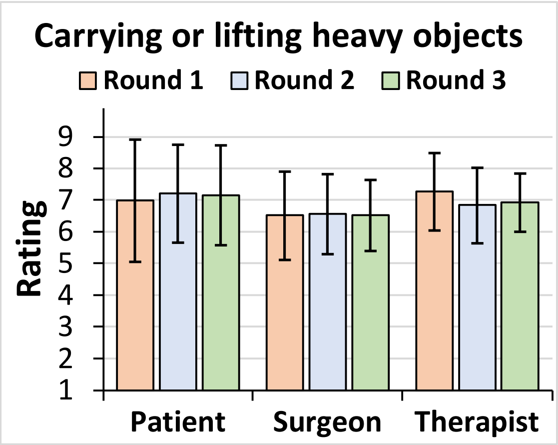

Sleeping

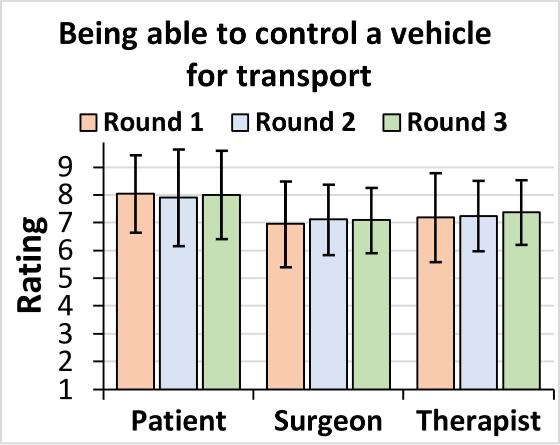

Thinking and memory

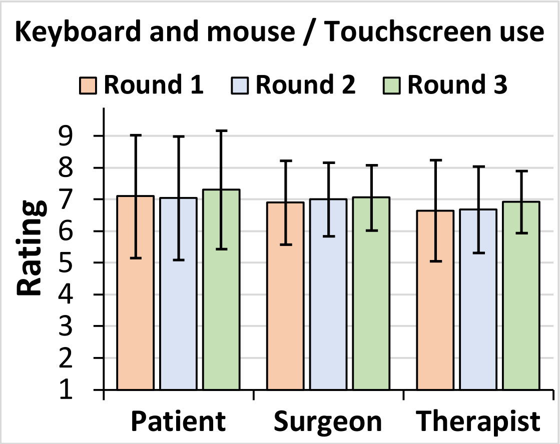

Fine hand use / Dexterity

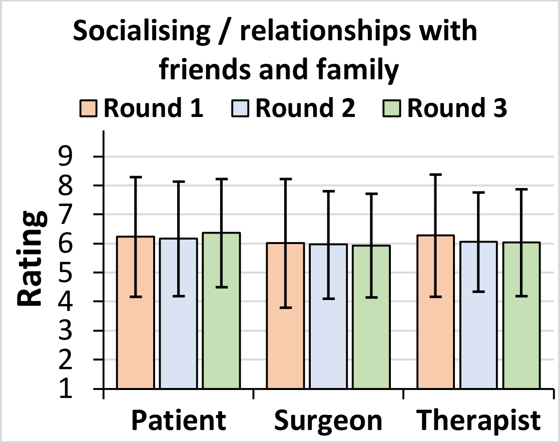

Hobbies and recreation

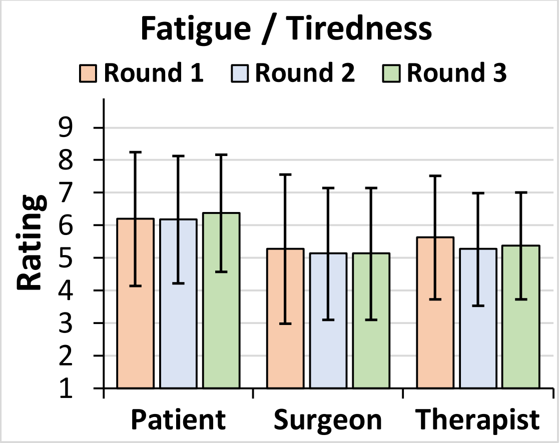

Return to usual work / job

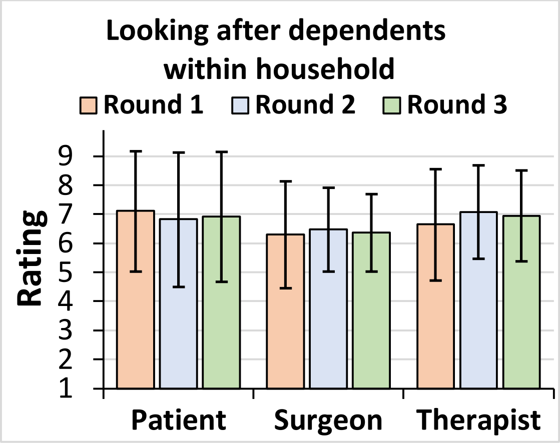

Carrying or lifting heavy objects

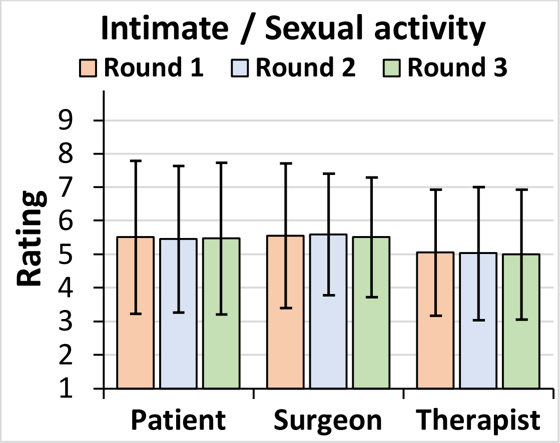

Being able to control a vehicle for transport
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

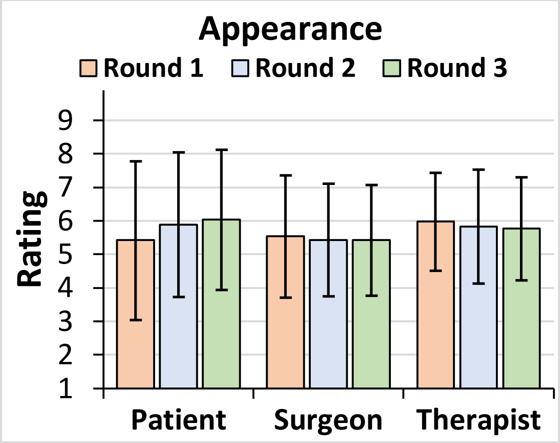

Keyboard and mouse / Touchscreen use
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

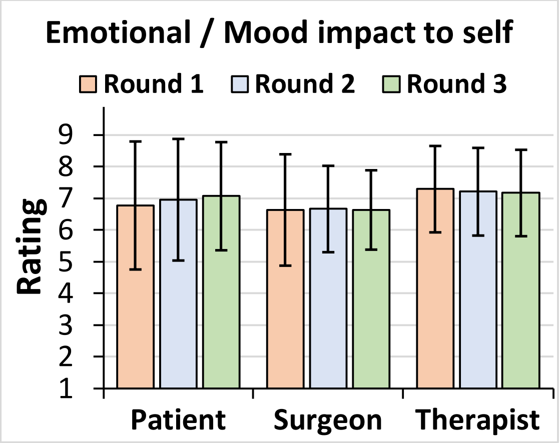

Socialising / relationships with friends and family
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

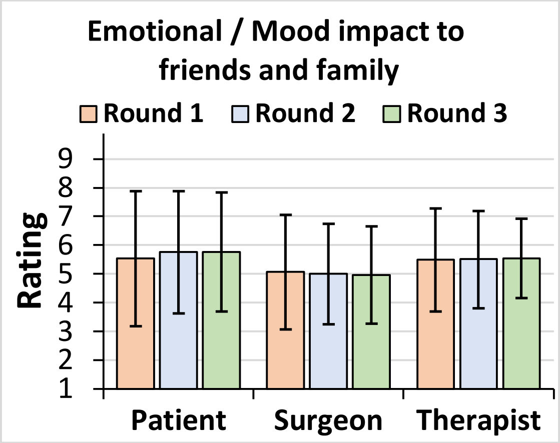

Fatigue / Tiredness
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

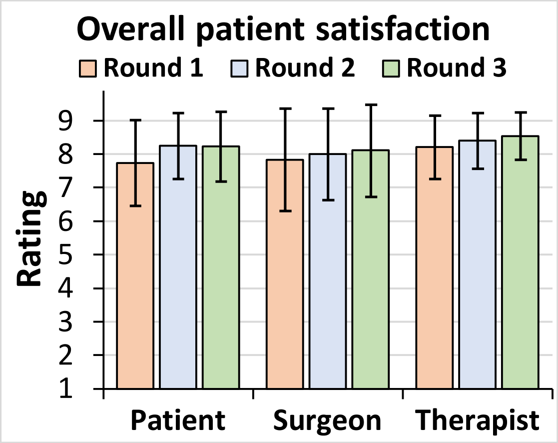

Looking after dependents within household
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

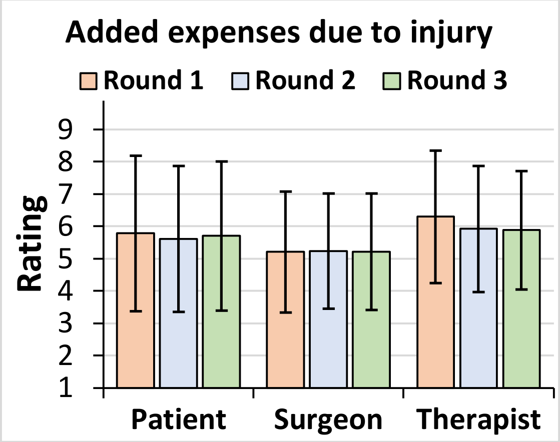

Intimate / Sexual activity
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

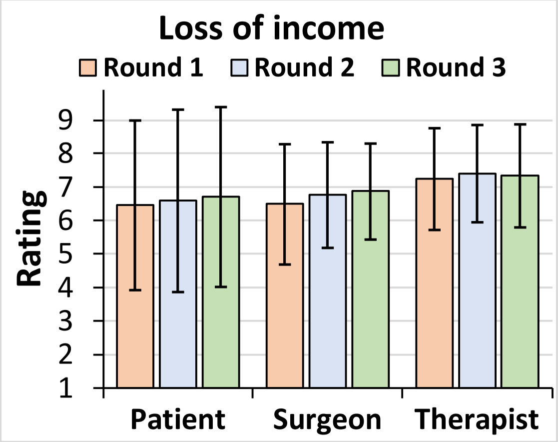

Appearance
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

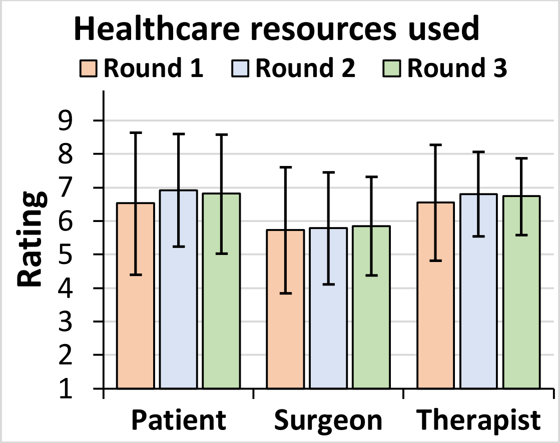

Emotional / Mood impact to self
Round 1
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

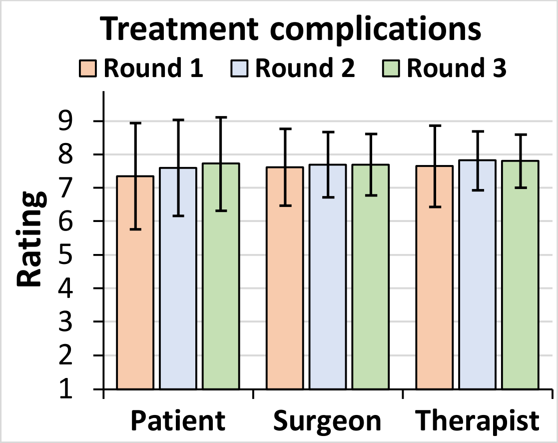

Emotional / Mood impact to friends and family
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist

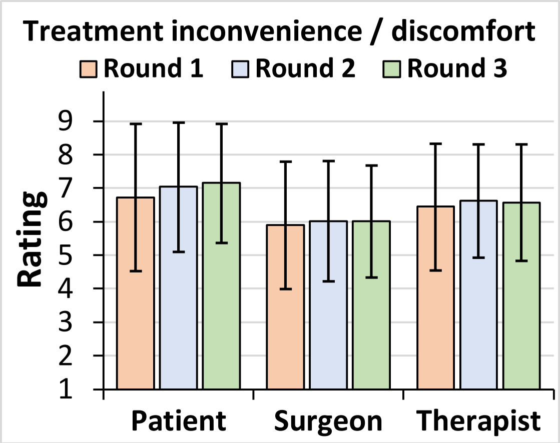

Overall patient satisfaction
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist

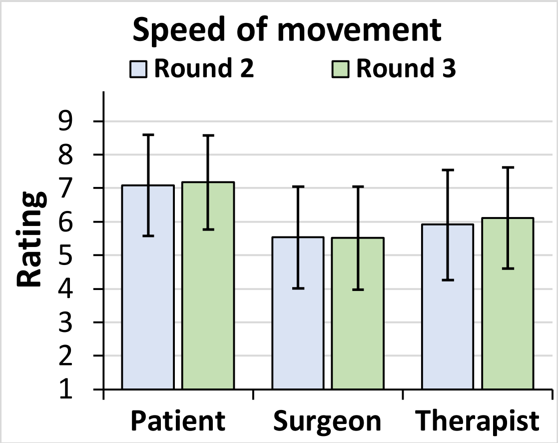

Added expenses due to injury
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist

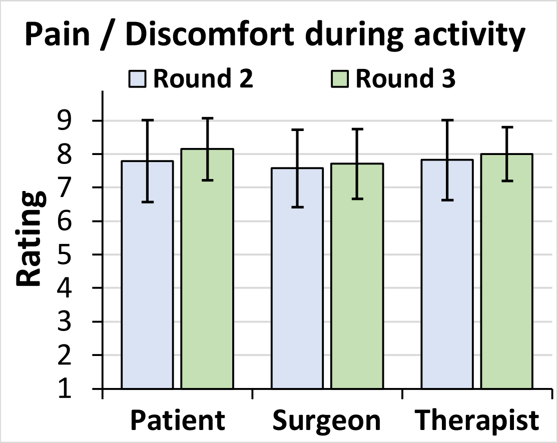

Loss of income
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist

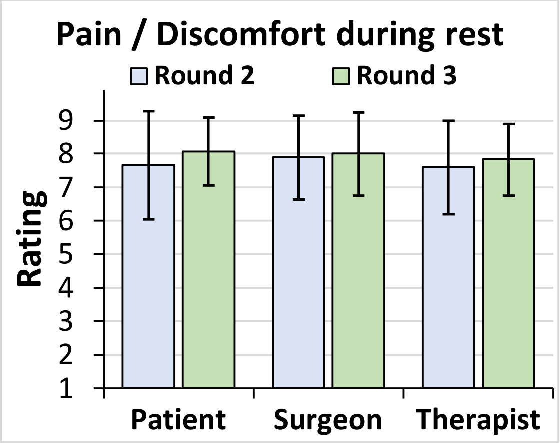

Healthcare resources used
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist

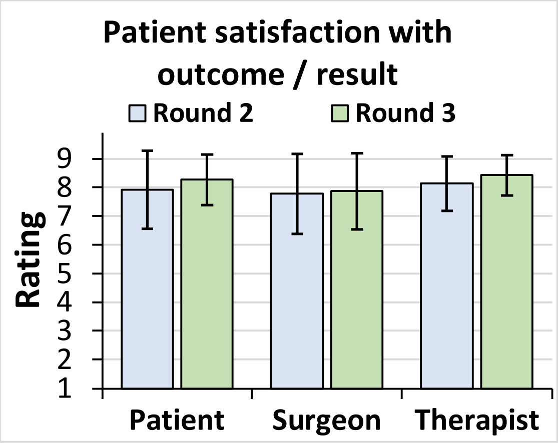

Treatment complications
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist


Treatment inconvenience / discomfort
Round 1
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist


Speed of movement
Round 2
Round 3
Rating
1
2
3
4
5
6
7
8
9
Patient
Surgeon
Therapist


Pain / Discomfort during activity
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist


Pain / Discomfort during rest
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist


Patient satisfaction with outcome / result
Round 2
Round 3
Rating
9
8
7
6
5
4
3
2
1
Patient
Surgeon
Therapist

Appendix 10 Magnitude and direction of Delphi ratings changes

Appendix Table 1 Number of patients with change in rating between Rounds 1 and 2 along with mean magnitude of change, per outcome domain

Outcome domain	Number that decreased rating	Mean decrease	Number with no rating change	Number that increased rating	Mean increase	Total number providing rating
Change in sensation	9	-1.8	31	13	1.5	53
Pain/discomfort	3	-2.7	34	18	1.7	55
Strength or muscle power	9	-1.6	33	13	1.8	55
Range of movement	5	-1.4	37	13	1.6	55
Confirmation of bone healing & alignment (fractures only)	5	-1.6	38	6	1.7	49
Confirmation of soft tissue healing	6	-1.8	35	10	1.7	51
Stability of joints	4	-1.5	38	10	1.8	52
Grasping and moving light objects	8	-2.6	31	16	2.1	55
Self-hygiene/personal care	6	-2.0	36	13	2.3	55
Cooking/preparing meals	5	-1.6	34	16	1.4	55
Dressing	5	-1.8	32	18	2.0	55
Eating and drinking	5	-2.4	34	16	1.5	55
Shopping for groceries	13	-1.5	32	10	2.0	55
Walking / mobility	10	-2.1	30	9	1.6	49
Using hands to change body position	7	-1.7	31	16	1.8	54
General household chores	8	-1.5	33	13	1.9	54
Sleeping	5	-1.6	35	15	1.7	55
Thinking and memory	13	-1.7	28	9	2.4	50
Fine hand use/dexterity	5	-1.8	37	13	2.0	55
Hobbies and recreation	7	-2.3	29	19	1.6	55
Return to usual work/job	6	-2.0	36	11	2.0	53
Carrying or lifting heavy objects	10	-1.2	29	15	1.8	54
Being able to control a vehicle for transport	5	-3.0	37	10	1.4	52
Keyboard and mouse / touchscreen use	13	-2.0	28	11	2.0	52
Socialising / relationships with friends and family	11	-2.1	31	10	2.0	52
Fatigue / tiredness	10	-2.1	31	10	1.8	51
Looking after dependents within household	9	-2.4	28	9	1.4	46
Intimate / sexual activity	6	-1.8	35	10	1.1	51
Appearance	10	-2.0	28	17	2.7	55
Emotional / mood impact to self	12	-2.3	28	15	2.5	55
Emotional / mood impact to friends and family	12	-2.3	28	14	2.9	54
Overall patient satisfaction	2	-1.0	34	18	1.6	54
Added expenses due to injury	11	-1.8	36	8	1.4	55
Loss of income	6	-2.7	34	11	2.3	51
Healthcare resources used	6	-1.3	39	10	3.0	55
Treatment complications	4	-2.5	37	11	1.9	52
Treatment inconvenience / discomfort	8	-2.0	29	18	1.8	55

Appendix Table 2 Number of patients with change in rating between Rounds 2 and 3 along with mean magnitude of change, per outcome domain

Outcome domain	**Number that decreased rating**	**Mean decrease**	**Number with no rating change**	**Number that increased rating**	**Mean increase**	**Total number providing rating**
Change in sensation	4	-1.3	38	12	1.6	54
Pain/discomfort	1	-1.0	46	7	1.9	54
Strength or muscle power	0	0.0	48	6	1.5	54
Range of movement	3	-1.3	45	6	1.7	54
Confirmation of bone healing & alignment (fractures only)	2	-1.0	46	3	2.3	51
Confirmation of soft tissue healing	3	-1.7	43	6	1.8	52
Stability of joints	0	0.0	44	10	1.6	54
Grasping and moving light objects	1	-1.0	46	7	2.9	54
Self-hygiene / personal care	1	-1.0	50	3	2.7	54
Cooking / preparing meals	4	-1.0	44	6	1.7	54
Dressing	1	-1.0	48	5	1.6	54
Eating and drinking	2	-1.0	45	7	2.6	54
Shopping for groceries	4	-1.3	40	10	1.3	54
Walking / mobility	5	-1.0	41	4	2.3	50
Using hands to change body position	2	-1.0	45	6	1.5	53
General household chores	1	-1.0	46	6	1.7	53
Sleeping	4	-1.5	42	8	1.9	54
Thinking and memory	2	-4.0	40	6	1.8	48
Fine hand use / dexterity	0	0.0	47	7	1.7	54
Hobbies and recreation	3	-1.0	44	7	2.0	54
Return to usual work / job	2	-2.0	41	10	1.6	53
Carrying or lifting heavy objects	6	-1.5	42	6	1.3	54
Being able to control a vehicle for transport	3	-1.3	46	5	2.2	54
Keyboard and mouse / touchscreen use	0	0.0	42	10	1.6	52
Socialising / relationships with friends and family	3	-1.7	39	10	1.8	52
Fatigue / tiredness	4	-1.0	38	9	1.9	51
Looking after dependents within household	3	-1.3	40	5	1.6	48
Intimate / sexual activity	1	-1.0	48	3	2.0	52
Appearance	2	-1.5	43	9	1.6	54
Emotional / mood impact to self	1	-2.0	47	6	1.7	54
Emotional / mood impact to friends and family	4	-2.3	40	8	1.3	52
Overall patient satisfaction	2	-2.0	51	1	1.0	54
Added expenses due to injury	2	-1.0	47	5	1.8	54
Loss of income	1	-2.0	44	6	1.7	51
Healthcare resources used	5	-2.6	44	5	1.4	54
Treatment complications	2	-1.5	45	6	1.8	53
Treatment inconvenience / discomfort	3	-2.0	41	10	1.4	54
Speed of movement	4	-1.5	41	8	1.5	53
Pain / discomfort during activity	1	-1.0	39	14	1.4	54
Pain / discomfort during rest	1	-1.0	43	10	2.3	54
Patient satisfaction with outcome / result	1	-2.0	39	13	1.5	53

Appendix Table 3 Number of surgeons with change in rating between Rounds 1 and 2 along with mean magnitude of change, per outcome domain

Outcome domain	Number that decreased rating	Mean decrease	Number with no rating change	Number that increased rating	Mean increase	Total number providing rating
Change in sensation	1	-1.0	38	16	1.3	55
Pain / discomfort	0	N/A	46	9	1.2	55
Strength or muscle power	3	-1.3	41	11	1.3	55
Range of movement	3	-1.3	41	11	1.1	55
Confirmation of bone healing & alignment (fractures only)	9	-1.9	41	4	1.5	54
Confirmation of soft tissue healing	4	-1.5	42	8	1.4	54
Stability of joints	3	-1.0	40	11	1.4	54
Grasping and moving light objects	2	-1.5	40	13	1.1	55
Self-hygiene / personal care	3	-1.0	38	14	1.2	55
Cooking / preparing meals	4	-1.3	40	11	1.1	55
Dressing	1	-1.0	43	11	1.1	55
Eating and drinking	0	N/A	45	10	1.2	55
Shopping for groceries	14	-1.6	33	8	1.3	55
Walking / mobility	11	-1.6	36	8	1.5	55
Using hands to change body position	9	-1.1	38	8	1.4	55
General household chores	12	-1.3	35	8	1.1	55
Sleeping	3	-1.0	42	10	1.8	55
Thinking and memory	5	-2.4	45	4	2.0	54
Fine hand use / dexterity	0	N/A	47	8	1.4	55
Hobbies and recreation	4	-1.0	43	8	1.6	55
Return to usual work / job	2	-1.0	41	12	1.5	55
Carrying or lifting heavy objects	6	-1.0	42	7	1.3	55
Being able to control a vehicle for transport	5	-1.2	40	10	1.5	55
Keyboard and mouse / touchscreen use	3	-1.3	45	7	1.3	55
Socialising / relationships with friends and family	10	-1.6	36	9	1.4	55
Fatigue / tiredness	10	-1.8	39	6	1.7	55
Looking after dependents within household	5	-1.0	42	8	1.9	55
Intimate / sexual activity	7	-1.3	41	6	1.8	54
Appearance	10	-1.3	40	5	1.4	55
Emotional / mood impact to self	8	-1.3	40	7	1.7	55
Emotional / mood impact to friends and family	7	-1.9	41	7	1.3	55
Overall patient satisfaction	2	-2.0	43	10	1.3	55
Added expenses due to injury	5	-1.6	44	6	1.5	55
Loss of income	3	-1.7	37	15	1.3	55
Healthcare resources used	6	-1.5	39	10	1.3	55
Treatment complications	2	-1.0	49	4	1.5	55
Treatment inconvenience / discomfort	6	-1.3	39	10	1.5	55

Appendix Table 4 Number of surgeons with change in rating between Rounds 2 and 3 along with mean magnitude of change, per outcome domain

Outcome domain	Number that decreased rating	Mean decrease	Number with no rating change	Number that increased rating	Mean increase	Total number providing rating
Change in sensation	1	-1.0	48	6	1.2	55
Pain / discomfort	0	0.0	52	3	1.0	55
Strength or muscle power	5	-1.2	44	6	1.2	55
Range of movement	5	-1.4	46	4	1.3	55
Confirmation of bone healing & alignment (fractures only)	7	-2.0	45	2	4.0	54
Confirmation of soft tissue healing	3	-2.0	46	5	1.6	54
Stability of joints	3	-1.0	48	3	1.0	54
Grasping and moving light objects	0	0.0	48	7	1.1	55
Self-hygiene / personal care	1	-2.0	48	6	1.0	55
Cooking / preparing meals	5	-1.4	46	4	1.0	55
Dressing	2	-1.5	47	6	1.0	55
Eating and drinking	1	-2.0	49	5	1.0	55
Shopping for groceries	3	-1.7	49	3	1.0	55
Walking / mobility	4	-1.3	48	3	2.0	55
Using hands to change body position	6	-1.5	45	4	1.0	55
General household chores	8	-1.1	44	3	1.0	55
Sleeping	1	-1.0	46	8	1.6	55
Thinking and memory	5	-1.6	47	2	1.5	54
Fine hand use / dexterity	3	-1.0	50	2	1.0	55
Hobbies and recreation	6	-1.2	46	3	1.0	55
Return to usual work / job	1	-2.0	49	5	1.0	55
Carrying or lifting heavy objects	7	-1.0	43	5	1.0	55
Being able to control a vehicle for transport	3	-1.3	49	3	1.0	55
Keyboard and mouse / touchscreen use	6	-1.0	40	9	1.0	55
Socialising / relationships with friends and family	8	-1.1	41	6	1.2	55
Fatigue / tiredness	2	-1.0	51	2	1.0	55
Looking after dependents within household	8	-1.3	43	4	1.0	55
Intimate / sexual activity	3	-1.3	51	0	0.0	54
Appearance	6	-1.0	48	1	5.0	55
Emotional / mood impact to self	6	-1.0	45	4	1.0	55
Emotional / mood impact to friends and family	4	-1.0	50	1	2.0	55
Overall patient satisfaction	2	-1.0	46	7	1.1	55
Added expenses due to injury	7	-1.0	44	4	1.5	55
Loss of income	5	-1.0	41	9	1.2	55
Healthcare resources used	4	-1.3	45	6	1.3	55
Treatment complications	5	-1.2	44	6	1.0	55
Treatment inconvenience / discomfort	4	-1.5	46	5	1.2	55
Speed of movement	4	-1.0	46	4	1.0	54
Pain / discomfort during activity	0	0.0	47	7	1.0	54
Pain / discomfort during rest	1	-1.0	46	7	1.0	54
Patient satisfaction with outcome / result	2	-1.0	47	5	1.4	54

Appendix Table 5 Number of therapists with change in rating between Rounds 1 and 2 along with mean magnitude of change, per outcome domain

Outcome domain	Number that decreased rating	Mean decrease	Number with no rating change	Number that increased rating	Mean increase	Total number providing rating
Change in sensation	2	-1.5	26	9	2.2	37
Pain / discomfort	2	-2.5	26	9	1.8	37
Strength or muscle power	7	-1.7	20	10	1.3	37
Range of movement	4	-2.3	24	9	1.6	37
Confirmation of bone healing & alignment (fractures only)	7	-1.9	25	5	1.8	37
Confirmation of soft tissue healing	5	-2.2	21	11	1.4	37
Stability of joints	1	-2.0	22	14	1.4	37
Grasping and moving light objects	5	-2.0	26	6	1.2	37
Self-hygiene / personal care	1	-4.0	26	10	1.4	37
Cooking / preparing meals	5	-1.6	24	8	1.3	37
Dressing	1	-4.0	29	7	1.4	37
Eating and drinking	4	-2.0	25	8	1.3	37
Shopping for groceries	13	-1.9	20	4	1.0	37
Walking / mobility	8	-2.4	24	5	1.8	37
Using hands to change body position	12	-1.7	17	7	1.3	36
General household chores	6	-2.3	24	7	1.1	37
Sleeping	2	-2.5	24	11	1.8	37
Thinking and memory	7	-1.7	26	4	1.3	37
Fine hand use / dexterity	2	-3.5	24	11	1.5	37
Hobbies and recreation	5	-2.2	25	7	1.1	37
Return to usual work / job	0	N/A	25	12	1.3	37
Carrying or lifting heavy objects	10	-1.8	25	2	1.0	37
Being able to control a vehicle for transport	5	-2.0	25	7	1.7	37
Keyboard and mouse / touchscreen use	7	-1.9	20	10	1.4	37
Socialising / relationships with friends and family	10	-1.9	21	6	1.8	37
Fatigue / tiredness	9	-2.2	22	6	1.2	37
Looking after dependents within household	1	-1.0	26	10	1.7	37
Intimate / sexual activity	7	-1.7	22	8	1.4	37
Appearance	7	-1.6	26	4	1.5	37
Emotional / mood impact to self	5	-2.0	25	7	1.0	37
Emotional / mood impact to friends and family	7	-1.4	23	7	1.6	37
Overall patient satisfaction	0	N/A	31	6	1.2	37
Added expenses due to injury	11	-1.4	24	1	1.0	36
Loss of income	4	-1.5	22	10	1.1	36
Healthcare resources used	5	-1.4	22	9	1.8	36
Treatment complications	5	-1.0	24	8	1.4	37
Treatment inconvenience / discomfort	5	-1.0	23	8	1.4	36

Appendix Table 6 Number of therapists with change in rating between Rounds 2 and 3 along with mean magnitude of change, per outcome domain

Outcome domain	Number that decreased rating	Mean decrease	Number with no rating change	Number that increased rating	Mean increase	Total number providing rating
Change in sensation	1	-1.0	28	6	1.2	35
Pain / discomfort	1	-1.0	28	6	1.2	35
Strength or muscle power	1	-1.0	30	4	1.3	35
Range of movement	1	-1.0	32	2	2.0	35
Confirmation of bone healing & alignment (fractures only)	1	-1.0	33	1	1.0	35
Confirmation of soft tissue healing	3	-1.3	27	5	1.0	35
Stability of joints	1	-1.0	31	3	1.0	35
Grasping and moving light objects	1	-1.0	29	5	1.2	35
Self-hygiene / personal care	0	0.0	29	6	1.0	35
Cooking / preparing meals	4	-1.0	30	1	1.0	35
Dressing	3	-1.7	29	3	2.0	35
Eating and drinking	2	-1.0	28	5	1.2	35
Shopping for groceries	3	-1.0	29	3	1.3	35
Walking / mobility	3	-1.0	29	3	1.3	35
Using hands to change body position	0	0.0	34	1	2.0	35
General household chores	7	-1.1	24	4	1.5	35
Sleeping	1	-1.0	31	3	1.3	35
Thinking and memory	4	-1.5	27	4	1.3	35
Fine hand use / dexterity	0	0.0	30	5	2.0	35
Hobbies and recreation	2	-1.0	28	5	1.8	35
Return to usual work / job	0	0.0	31	4	1.0	35
Carrying or lifting heavy objects	0	0.0	32	3	2.0	35
Being able to control a vehicle for transport	1	-1.0	29	5	1.2	35
Keyboard and mouse / touchscreen use	2	-1.0	29	4	2.5	35
Socialising / relationships with friends and family	2	-2.0	29	4	1.5	35
Fatigue / tiredness	2	-1.0	28	5	1.4	35
Looking after dependents within household	4	-1.3	27	4	1.0	35
Intimate / sexual activity	1	-1.0	33	1	2.0	35
Appearance	3	-1.3	30	2	1.5	35
Emotional / mood impact to self	4	-1.3	28	3	1.3	35
Emotional / mood impact to friends and family	4	-1.3	25	6	1.2	35
Overall patient satisfaction	0	0.0	31	4	1.5	35
Added expenses due to injury	1	-2.0	31	3	1.3	35
Loss of income	2	-1.0	32	1	1.0	35
Healthcare resources used	2	-1.0	31	2	1.0	35
Treatment complications	1	-1.0	32	2	1.0	35
Treatment inconvenience / discomfort	1	-2.0	33	1	1.0	35
Speed of movement	1	-1.0	27	4	1.5	32
Pain / discomfort during activity	0	0.0	28	5	1.6	33
Pain / discomfort during rest	0	0.0	29	4	2.3	33
Patient satisfaction with outcome / result	0	0.0	26	7	1.4	33

Appendix 11 Number of significant Delphi ratings changes

Appendix Table 7 Number of significant changes by patients (n=55) between Rounds 1 and 2

Outcome domain	Boundary change of increased rating	Boundary change of decreased rating	Change from ‘ unable to score’	Change to ‘ unable to score’
Change in sensation	6	5	2	0
Pain / discomfort	4	2	0	0
Strength or muscle power	4	3	0	0
Range of movement	3	3	0	0
Confirmation of bone healing & alignment (fractures only)	2	2	3	0
Confirmation of soft tissue healing	4	2	2	0
Stability of joints	2	2	3	0
Grasping and moving light objects	10	4	0	0
Self-hygiene / personal care	7	3	0	0
Cooking / preparing meals	6	2	0	0
Dressing	6	3	0	0
Eating and drinking	6	2	0	0
Shopping for groceries	6	7	0	0
Walking / mobility	5	7	2	0
Using hands to change body position	11	4	0	0
General household chores	4	2	0	1
Sleeping	8	2	0	0
Thinking and memory	6	8	0	0
Fine hand use / dexterity	8	1	0	0
Hobbies and recreation	10	4	0	0
Return to usual work / job	4	3	1	0
Carrying or lifting heavy objects	7	1	1	0
Being able to control a vehicle for transport	3	3	3	0
Keyboard and mouse / touchscreen use	6	6	1	0
Socialising / relationships with friends and family	9	4	1	0
Fatigue / tiredness	7	3	1	0
Looking after dependents within household	5	4	4	0
Intimate / sexual activity	4	4	2	0
Appearance	13	5	0	0
Emotional / mood impact to self	10	6	0	0
Emotional / mood impact to friends and family	10	7	0	0
Overall patient satisfaction	5	0	1	0
Added expenses due to injury	4	7	0	0
Loss of income	6	4	1	1
Healthcare resources used	8	3	0	0
Treatment complications	2	2	2	0
Treatment inconvenience / discomfort	9	4	0	0

Appendix Table 8 Number of significant changes by patients (n=54) between Rounds 2 and 3

Outcome domain	Boundary change of increased rating	Boundary change of decreased rating	Change from ‘ unable to score’	Change to ‘ unable to score’
Change in sensation	7	0	0	0
Pain / discomfort	3	0	0	0
Strength or muscle power	2	0	0	0
Range of movement	3	1	0	0
Confirmation of bone healing & alignment (fractures only)	3	2	0	0
Confirmation of soft tissue healing	3	1	0	0
Stability of joints	4	0	0	0
Grasping and moving light objects	3	0	0	0
Self-hygiene / personal care	2	0	0	0
Cooking / preparing meals	3	2	0	0
Dressing	2	0	0	0
Eating and drinking	3	0	0	0
Shopping for groceries	4	1	0	0
Walking / mobility	2	1	0	0
Using hands to change body position	4	0	0	0
General household chores	3	1	1	0
Sleeping	3	1	0	0
Thinking and memory	5	2	0	1
Fine hand use / dexterity	2	0	0	0
Hobbies and recreation	5	0	0	0
Return to usual work / job	3	1	0	0
Carrying or lifting heavy objects	2	2	0	0
Being able to control a vehicle for transport	2	2	0	0
Keyboard and mouse / touchscreen use	4	0	0	0
Socialising / relationships with friends and family	5	2	0	0
Fatigue / tiredness	3	2	0	0
Looking after dependents within household	1	2	1	1
Intimate / sexual activity	1	1	0	0
Appearance	3	1	0	0
Emotional / mood impact to self	4	0	0	0
Emotional / mood impact to friends and family	3	1	0	1
Overall patient satisfaction	0	1	0	0
Added expenses due to injury	5	2	0	0
Loss of income	4	0	0	0
Healthcare resources used	4	4	0	0
Treatment complications	3	1	0	0
Treatment inconvenience / discomfort	4	1	0	0
Speed of movement	4	0	1	0
Pain / discomfort during activity	6	1	0	0
Pain / discomfort during rest	6	0	0	0
Patient satisfaction with outcome / result	4	1	1	0

Appendix Table 9 Number of significant changes by surgeons (n=55) between Rounds 1 and 2

Outcome domain	Boundary change of increased rating	Boundary change of decreased rating	Change from 'unable to score'	Change to 'unable to score'
Change in sensation	10	0	0	0
Pain / discomfort	2	0	0	0
Strength or muscle power	4	0	0	0
Range of movement	3	1	0	0
Confirmation of bone healing & alignment (fractures only)	3	6	0	0
Confirmation of soft tissue healing	4	3	0	0
Stability of joints	5	0	0	0
Grasping and moving light objects	5	1	0	0
Self-hygiene / personal care	3	0	0	0
Cooking / preparing meals	5	1	0	0
Dressing	2	0	0	0
Eating and drinking	3	0	0	0
Shopping for groceries	3	4	0	0
Walking / mobility	5	5	0	0
Using hands to change body position	5	2	0	0
General household chores	2	4	0	0
Sleeping	4	1	0	0
Thinking and memory	3	3	0	0
Fine hand use / dexterity	6	0	0	0
Hobbies and recreation	4	0	0	0
Return to usual work / job	3	0	0	0
Carrying or lifting heavy objects	5	1	0	0
Being able to control a vehicle for transport	5	1	0	0
Keyboard and mouse / touchscreen use	4	2	0	0
Socialising / relationships with friends and family	5	2	0	0
Fatigue / tiredness	3	6	0	0
Looking after dependents within household	2	1	0	0
Intimate / sexual activity	2	1	0	0
Appearance	3	4	0	0
Emotional / mood impact to self	4	2	0	0
Emotional / mood impact to friends and family	4	6	0	0
Overall patient satisfaction	3	1	0	0
Added expenses due to injury	5	3	0	0
Loss of income	7	2	0	0
Healthcare resources used	3	4	0	0
Treatment complications	3	0	0	0
Treatment inconvenience / discomfort	5	1	0	0

Appendix Table 10 Number of significant changes by surgeons (n=55) between Rounds 2 and 3

Outcome domain	Boundary change of increased rating	Boundary change of decreased rating	Change from ‘ unable to score’	Change to ‘ unable to score’
Change in sensation	2	0	0	0
Pain / discomfort	0	0	0	0
Strength or muscle power	3	1	0	0
Range of movement	1	0	0	0
Confirmation of bone healing & alignment (fractures only)	2	2	0	0
Confirmation of soft tissue healing	2	1	0	0
Stability of joints	1	0	0	0
Grasping and moving light objects	3	0	0	0
Self-hygiene / personal care	0	1	0	0
Cooking / preparing meals	2	1	0	0
Dressing	2	2	0	0
Eating and drinking	2	1	0	0
Shopping for groceries	0	2	0	0
Walking / mobility	1	1	0	0
Using hands to change body position	3	2	0	0
General household chores	1	2	0	0
Sleeping	3	0	0	0
Thinking and memory	1	1	0	0
Fine hand use / dexterity	0	0	0	0
Hobbies and recreation	2	0	0	0
Return to usual work / job	1	0	0	0
Carrying or lifting heavy objects	2	1	0	0
Being able to control a vehicle for transport	2	0	0	0
Keyboard and mouse / touchscreen use	3	0	0	0
Socialising / relationships with friends and family	1	3	0	0
Fatigue / tiredness	0	0	0	0
Looking after dependents within household	0	2	0	0
Intimate / sexual activity	0	1	0	0
Appearance	1	2	0	0
Emotional / mood impact to self	2	0	0	0
Emotional / mood impact to friends and family	1	1	0	0
Overall patient satisfaction	1	0	0	0
Added expenses due to injury	3	4	0	0
Loss of income	2	1	0	0
Healthcare resources used	2	1	0	0
Treatment complications	1	0	0	0
Treatment inconvenience / discomfort	1	0	0	0
Speed of movement	2	4	1	0
Pain / discomfort during activity	1	0	1	0
Pain / discomfort during rest	0	0	1	0
Patient satisfaction with outcome / result	0	0	1	0

Appendix Table 11 Number of significant changes by therapists (n=37) between Rounds 1 and 2

Outcome domain	Boundary change of increased rating	Boundary change of decreased rating	Change from ‘unable to score’	Change to ‘unable to score’
Change in sensation	3	1	0	0
Pain / discomfort	2	1	0	0
Strength or muscle power	4	3	0	0
Range of movement	3	1	0	0
Confirmation of bone healing & alignment (fractures only)	1	2	0	0
Confirmation of soft tissue healing	4	2	0	0
Stability of joints	5	0	0	0
Grasping and moving light objects	1	2	0	0
Self-hygiene / personal care	4	1	0	0
Cooking / preparing meals	4	1	0	0
Dressing	3	1	0	0
Eating and drinking	2	2	0	0
Shopping for groceries	2	5	0	0
Walking / mobility	2	4	0	0
Using hands to change body position	3	3	1	0
General household chores	5	2	0	0
Sleeping	8	2	0	0
Thinking and memory	2	5	0	0
Fine hand use / dexterity	4	1	0	0
Hobbies and recreation	2	1	0	0
Return to usual work / job	1	0	0	0
Carrying or lifting heavy objects	1	4	0	0
Being able to control a vehicle for transport	4	2	0	0
Keyboard and mouse / touchscreen use	4	2	0	0
Socialising / relationships with friends and family	2	4	0	0
Fatigue / tiredness	1	6	0	0
Looking after dependents within household	7	0	0	0
Intimate / sexual activity	2	2	0	0
Appearance	1	4	0	0
Emotional / mood impact to self	3	2	0	0
Emotional / mood impact to friends and family	4	2	0	0
Overall patient satisfaction	0	0	0	0
Added expenses due to injury	0	4	1	0
Loss of income	2	0	1	0
Healthcare resources used	4	1	1	0
Treatment complications	3	0	0	0
Treatment inconvenience / discomfort	3	1	1	0

Appendix Table 12 Number of significant changes by therapists (n=35) between Rounds 2 and 3

Outcome domain	Boundary change of increased rating	Boundary change of decreased rating	Change from ' unable to score'	Change to ' unable to score'
Change in sensation	1	0	0	0
Pain / discomfort	0	0	0	0
Strength or muscle power	1	0	0	0
Range of movement	1	0	0	0
Confirmation of bone healing & alignment (fractures only)	1	1	0	0
Confirmation of soft tissue healing	1	1	0	0
Stability of joints	2	0	0	0
Grasping and moving light objects	2	0	0	0
Self-hygiene / personal care	1	0	0	0
Cooking / preparing meals	0	0	0	0
Dressing	2	1	0	0
Eating and drinking	2	0	0	0
Shopping for groceries	2	2	0	0
Walking / mobility	0	1	0	0
Using hands to change body position	0	0	0	0
General household chores	2	2	0	0
Sleeping	1	0	0	0
Thinking and memory	1	1	0	0
Fine hand use / dexterity	1	0	0	0
Hobbies and recreation	4	0	0	0
Return to usual work / job	0	0	0	0
Carrying or lifting heavy objects	1	0	0	0
Being able to control a vehicle for transport	2	0	0	0
Keyboard and mouse / touchscreen use	3	0	0	0
Socialising / relationships with friends and family	2	0	0	0
Fatigue / tiredness	2	1	0	0
Looking after dependents within household	1	1	0	0
Intimate / sexual activity	0	0	0	0
Appearance	1	1	0	0
Emotional / mood impact to self	3	1	0	0
Emotional / mood impact to friends and family	2	1	0	0
Overall patient satisfaction	1	0	0	0
Added expenses due to injury	1	0	0	0
Loss of income	0	0	0	0
Healthcare resources used	1	0	0	0
Treatment complications	1	0	0	0
Treatment inconvenience / discomfort	0	1	0	0
Speed of movement	1	1	3	0
Pain / discomfort during activity	3	0	2	0
Pain / discomfort during rest	4	0	2	0
Patient satisfaction with outcome / result	1	0	2	0

Appendix 12 Tables showing categories for Delphi participants' reasons for significant score change between rounds

Appendix Table 13 Reason categories for patients' significant score changes between Rounds 1 & 2

Reason category	**Number and % of significant changes in ratings for which reason was provided**	**Number of patients**
Perspective changed due to experience of symptoms/problems	108/171 (63%)	27
Perspective changed due to increased activity as recovery progressed	14/171 (8%)	7
Domain not as important as others	14/171 (8%)	7
Domain was misperceived/not fully understood previously	8/171 (5%)	4
Reconsidered importance to subset of patients	7/171 (4%)	3
Issues relating to domain can be bypassed	6/171 (4%)	3
Value as marker of effectiveness of treatment/recovery	2/171 (1%)	2
COVID-19 impact influenced change in importance	2/171 (1%)	2
Conforming to general Delphi panel	1/171 (0.6%)	1
Boundary change – no clear reason for change	9/171 (5%)	7

Appendix Table 14 Reason categories for patients' significant score changes between Rounds 2 & 3

Reason category	**Number and % of significant changes in ratings for which reason was provided**	**Number of patients**
Perspective changed due to experience of symptoms/problems	32/68 (47%)	9
Perspective changed due to increased activity as recovery progressed	9/68 (13%)	6
Domain was misperceived/not fully understood previously	3/68 (4%)	3
Conforming to general Delphi panel	3/68 (4%)	2
Value as marker of effectiveness of treatment/recovery	2/68 (3%)	2
Loss of employment as consequence of injury	1/68 (1%)	1
Reconsidered importance of measuring domain to improve future treatments	1/68 (1%)	1
COVID-19 impact influenced change in importance	1/68 (1%)	1
Previously unrated item – no clear reason for change	2/68 (3%)	2
Boundary change – no clear reason for change	14/68 (21%)	6

Appendix Table 15 Reason categories for surgeons' significant score changes between Rounds 1 & 2

Reason category	**Number and % of significant changes in ratings for which reason was provided**	**Number of surgeons**
Conforming to patient stakeholder group	38/141 (27%)	13
Domain was misperceived/not fully understood previously	12/141 (9%)	8
Domain not as important as others	9/141 (6%)	7
Conforming to general Delphi panel	15/141 (11%)	5
Reconsidered as domain can be influenced by non-injury factors	8/141 (6%)	5
Reconsidered importance to subset of patients	7/141 (5%)	3
Value as marker of effectiveness of treatment/recovery	4/141 (3%)	3
Linked to other domains which are important	3/141 (2%)	3
Reconsidered as domain cannot be affected by treatment	3/141 (2%)	2
Answered without looking at previous score	7/141 (5%)	1
Reconsidered importance due to ease/difficulty in measuring domain	6/141 (4%)	1
Issues relating to domain can be bypassed	1/141 (1%)	1
Reconsidered importance in relation to research	1/141 (1%)	1
Boundary change – no clear reason for change	27/141 (19%)	13

Appendix Table 16 Reason categories for surgeons' significant score changes between Rounds 2 & 3

Reason category	**Number and % of significant changes in ratings for which reason was provided**	**Number of surgeons**
Conforming to patient stakeholder group	13/69 (19%)	5
Conforming to general Delphi panel	6/69 (9%)	4
Value as marker of effectiveness of treatment/recovery	5/69 (7%)	3
Issues relating to domain can be bypassed	3/69 (4%)	2
Perspective changed due to recent patient encounters	2/69 (3%)	1
Reconsidered importance to subset of patients	2/69 (3%)	2
Reconsidered importance due to ease/difficulty in measuring domain	2/69 (3%)	1
Reconsidered importance in relation to research	2/69 (3%)	1
Domain not as important as others	2/69 (3%)	2
Domain was misperceived/not fully understood previously	2/69 (3%)	2
Conforming to surgeon stakeholder group	1/69 (1%)	1
Boundary change – no clear reason for change	29/69 (42%)	11

Appendix Table 17 Reason categories for therapists' significant score changes between Rounds 1 & 2

Reason category	**Number and % of significant changes in ratings for which reason was provided**	**Number of therapists**
Domain was misperceived/not fully understood previously	13/106 (12%)	11
Conforming to patient stakeholder group	21/106 (20%)	8
Domain not as important as others	8/106 (8%)	6
Reconsidered importance to subset of patients	6/106 (6%)	5
Conforming to general Delphi panel	19/106 (18%)	4
Linked to other domains which are important	7/106 (7%)	4
Issues relating to domain can be bypassed	5/106 (5%)	4
Perspective changed due to recent patient encounters	4/106 (4%)	4
COVID-19 impact influenced change in importance	4/106 (4%)	4
Reconsidered importance in relation to research	2/106 (2%)	1
Reconsidered as domain can be influenced by non-injury factors	1/106 (1%)	1
Value as marker of effectiveness of treatment/recovery	1/106 (1%)	1
Reconsidered importance as domain would enable specific research studies	1/106 (1%)	1
Boundary change – no clear reason for change	14/106 (13%)	9

Appendix Table 18 Reason categories for therapists' significant score changes between Rounds 2 & 3

Reason category	**Number and % of significant changes in ratings for which reason was provided**	**Number of therapists**
Conforming to patient stakeholder group	11/55 (20%)	7
Domain was misperceived/not fully understood previously	5/55 (9%)	5
Conforming to general Delphi panel	11/55 (20%)	3
Issues relating to domain can be bypassed	3/55 (5%)	3
Reconsidered importance to subset of patients	3/55 (5%)	3
Reconsidered importance in relation to research	3/55 (5%)	1
Linked to other domains which are important	1/55 (2%)	1
Domain not as important as others	1/55 (2%)	1
Value as marker of effectiveness of treatment/recovery	1/55 (2%)	1
Boundary change – no clear reason for change	16/55 (29%)	5

Appendix 13 Projected figures in hypotheticals of Delphi study attrition

Appendix Table 19 ‘Last known rating’ scenario percentages per rating category

Outcome domain	% Patients			% Surgeons			% Therapists		
	1-3	4-6	7-9	1-3	4-6	7-9	1-3	4-6	7-9
Change in sensation	0	16	84	0	13	87	0	10	90
Pain / discomfort	0	5	95	0	0	100	0	5	95
Strength or muscle power	0	2	98	0	16	84	0	8	93
Range of movement	0	11	89	0	27	73	0	0	100
Confirmation of bone healing & alignment (fractures only)	0	11	89	7	57	35	0	40	60
Confirmation of soft tissue healing	2	11	87	2	54	44	3	30	68
Stability of joints	0	2	98	0	17	83	0	0	100
Grasping and moving light objects	0	2	98	0	4	96	0	8	93
Self-hygiene / personal care	2	0	98	0	5	95	0	5	95
Cooking / preparing meals	0	18	82	0	29	71	0	15	85
Dressing	0	14	86	0	20	80	3	8	90
Eating and drinking	0	5	95	0	13	87	0	10	90
Shopping for groceries	5	33	61	5	65	29	5	63	33
Walking / mobility	13	28	58	22	45	33	50	25	25
Using hands to change body position	2	30	68	4	38	58	5	45	50
General household chores	5	33	61	2	49	49	3	35	63
Sleeping	7	11	82	4	9	87	0	23	78
Thinking and memory	22	27	51	35	33	31	43	35	23
Fine hand use / dexterity	0	2	98	0	0	100	0	3	98
Hobbies and recreation	0	19	81	2	33	65	0	15	85
Return to usual work / job	4	4	93	0	2	98	0	0	100
Carrying or lifting heavy objects	4	25	72	2	44	55	0	33	68
Being able to control a vehicle for transport	4	9	88	2	25	73	0	23	78
Keyboard and mouse / touchscreen use	5	18	76	0	22	78	0	25	75
Socialising / relationships with friends and family	7	42	51	11	45	44	8	40	53
Fatigue / tiredness	7	37	56	24	53	24	15	60	25
Looking after dependents within household	12	17	71	2	56	42	5	15	80
Intimate / sexual activity	22	42	36	13	48	39	18	60	23
Appearance	12	44	44	16	45	38	8	63	30
Emotional / mood impact to self	4	18	79	0	36	64	3	18	80
Emotional / mood impact to friends and family	13	44	44	22	62	16	8	68	25
Overall patient satisfaction	0	7	93	0	13	87	0	0	100
Added expenses due to injury	19	40	40	22	55	24	13	48	40
Loss of income	15	19	67	5	27	67	0	28	73
Healthcare resources used	5	26	68	11	62	27	0	35	65
Treatment complications	4	13	84	0	7	93	0	5	95
Treatment inconvenience / discomfort	7	21	72	11	38	51	5	33	63
Speed of movement	0	28	72	15	60	25	8	43	49
Pain / discomfort during activity	0	6	94	0	9	91	0	3	97
Pain / discomfort during rest	0	4	96	2	9	89	0	8	92
Patient satisfaction with outcome / result	0	4	96	2	9	89	0	0	100

Appendix Table 20 'Contrary to stakeholder group' scenario percentages per rating category

Outcome domain	% Patients			% Surgeons			% Therapists		
	1-3	4-6	7-9	1-3	4-6	7-9	1-3	4-6	7-9
Change in sensation	5	14	81	0	13	87	13	10	78
Pain / discomfort	5	5	89	0	0	100	13	5	83
Strength or muscle power	5	2	93	0	16	84	13	8	80
Range of movement	5	11	84	0	27	73	13	0	88
Confirmation of bone healing & alignment (fractures only)	6	11	83	7	57	35	0	35	65
Confirmation of soft tissue healing	7	11	82	2	54	44	3	25	73
Stability of joints	5	2	93	0	17	83	13	0	88
Grasping and moving light objects	5	2	93	0	4	96	13	5	83
Self-hygiene / personal care	7	0	93	0	5	95	13	5	83
Cooking / preparing meals	5	18	77	0	29	71	13	15	73
Dressing	5	14	81	0	20	80	15	8	78
Eating and drinking	5	5	89	0	13	87	13	10	78
Shopping for groceries	5	33	61	5	65	29	5	58	38
Walking / mobility	11	28	60	22	45	33	48	23	30
Using hands to change body position	2	29	70	4	38	58	5	43	53
General household chores	5	32	63	2	49	49	3	35	63
Sleeping	11	11	79	4	9	87	13	23	65
Thinking and memory	20	27	53	35	33	31	43	25	33
Fine hand use / dexterity	5	2	93	0	0	100	13	3	85
Hobbies and recreation	5	18	77	2	33	65	13	15	73
Return to usual work / job	9	2	89	0	2	98	13	0	88
Carrying or lifting heavy objects	9	23	68	2	44	55	0	30	70
Being able to control a vehicle for transport	9	7	84	2	25	73	13	18	70
Keyboard and mouse / touchscreen use	11	16	73	0	22	78	13	18	70
Socialising / relationships with friends and family	7	40	53	11	45	44	8	38	55
Fatigue / tiredness	7	35	57	24	53	24	15	55	30
Looking after dependents within household	12	17	71	2	56	42	18	13	70
Intimate / sexual activity	20	42	38	13	48	39	18	53	30
Appearance	12	42	46	16	45	38	5	60	35
Emotional / mood impact to self	9	18	74	0	36	64	15	15	70
Emotional / mood impact to friends and family	13	44	44	22	62	16	8	60	33
Overall patient satisfaction	5	7	88	0	13	87	13	0	88
Added expenses due to injury	19	39	42	22	55	24	13	40	48
Loss of income	15	17	69	5	27	67	13	25	63
Healthcare resources used	5	25	70	11	62	27	0	30	70
Treatment complications	7	13	80	0	7	93	13	3	85
Treatment inconvenience / discomfort	11	21	68	11	38	51	5	28	68
Speed of movement	5	26	68	15	60	25	8	38	55
Pain / discomfort during activity	5	5	89	0	9	91	13	3	85
Pain / discomfort during rest	5	4	91	2	9	89	13	8	80
Patient satisfaction with outcome / result	5	4	91	2	9	89	13	0	88

Appendix Table 21 'Contrary to overall Delphi panel' scenario percentages per rating category

Outcome domain	% Patients			% Surgeons			% Therapists		
	1-3	4-6	7-9	1-3	4-6	7-9	1-3	4-6	7-9
Change in sensation	5	14	81	0	13	87	13	10	78
Pain / discomfort	5	5	89	0	0	100	13	5	83
Strength or muscle power	5	2	93	0	16	84	13	8	80
Range of movement	5	11	84	0	27	73	13	0	88
Confirmation of bone healing & alignment (fractures only)	0	11	89	7	57	35	0	35	65
Confirmation of soft tissue healing	2	11	87	2	54	44	3	25	73
Stability of joints	5	2	93	0	17	83	13	0	88
Grasping and moving light objects	5	2	93	0	4	96	13	5	83
Self-hygiene / personal care	7	0	93	0	5	95	13	5	83
Cooking / preparing meals	5	18	77	0	29	71	13	15	73
Dressing	5	14	81	0	20	80	15	8	78
Eating and drinking	5	5	89	0	13	87	13	10	78
Shopping for groceries	5	33	61	5	65	29	5	58	38
Walking / mobility	11	28	60	22	45	33	48	23	30
Using hands to change body position	2	29	70	4	38	58	5	43	53
General household chores	5	32	63	2	49	49	3	35	63
Sleeping	11	11	79	4	9	87	13	23	65
Thinking and memory	20	27	53	35	33	31	43	25	33
Fine hand use / dexterity	5	2	93	0	0	100	13	3	85
Hobbies and recreation	0	18	82	2	33	65	0	15	85
Return to usual work / job	9	2	89	0	2	98	13	0	88
Carrying or lifting heavy objects	4	23	74	2	44	55	0	30	70
Being able to control a vehicle for transport	9	7	84	2	25	73	13	18	70
Keyboard and mouse / touchscreen use	11	16	73	0	22	78	13	18	70
Socialising / relationships with friends and family	7	40	53	11	45	44	8	38	55
Fatigue / tiredness	7	35	57	24	53	24	15	55	30
Looking after dependents within household	12	17	71	2	56	42	5	13	83
Intimate / sexual activity	20	42	38	13	48	39	18	53	30
Appearance	12	42	46	16	45	38	5	60	35
Emotional / mood impact to self	4	18	79	0	36	64	3	15	83
Emotional / mood impact to friends and family	13	44	44	22	62	16	8	60	33
Overall patient satisfaction	5	7	88	0	13	87	13	0	88
Added expenses due to injury	19	39	42	22	55	24	13	40	48
Loss of income	15	17	69	5	27	67	0	25	75
Healthcare resources used	5	25	70	11	62	27	0	30	70
Treatment complications	7	13	80	0	7	93	13	3	85
Treatment inconvenience / discomfort	5	21	74	11	38	51	5	28	68
Speed of movement	0	26	74	15	60	25	8	38	55
Pain / discomfort during activity	5	5	89	0	9	91	13	3	85
Pain / discomfort during rest	5	4	91	2	9	89	13	8	80
Patient satisfaction with outcome / result	5	4	91	2	9	89	13	0	88

Appendix 14 Consensus meeting preparatory information pack

SO-HANDI

Consensus Meeting

Information pack

Title of Study: **Standardised Outcomes for Hand fractures and joint Injuries (SO-HANDI study): Delphi Study & Consensus Meeting**
IRAS Project ID: 283554 Sponsor: University of Nottingham
Name of Chief Investigator: Dr Alexia Karantana
Local Researcher(s): Dr Sandeep Deshmukh, Dr Paul Leighton, Prof Alan Montgomery

Contents

1. Meeting objective

Thank you for agreeing to participate in the virtual SO-HANDI consensus meeting. This information pack is designed to summarise key results from the previous research we have done on this topic and to give you an idea of the sorts of discussions we will have at the meeting. We ask that you please take some time to read through the material here and think through some of the things we have marked with a 🗨 symbol. This will give you a chance to gather your thoughts ahead of time, which will really help out with discussions on the day of the meeting.

The overall objective of the meeting is to agree on what outcomes should be included in a core outcome set for all future clinical research on hand fractures and joint injuries in adults.

To do this, we plan to run the meeting using the following timetable on the day. There will be some short breaks between the sessions.

2. Provisional agenda

Time (BST)	Activity
12:50	Joining the online meeting
13:00	Introduction • Welcome and overview of core outcome sets in general • Scope of our core outcome set • Highlight outcomes deemed 'very important' from the Delphi study • Questions from participants
13:15	**Whole group session 1**: Decide whether any other outcome domains from the Delphi study should also be considered in the rest of the meeting
13:50	**Small group session 1**: Discussion on which outcomes are most important to include in a core outcome set
14:40	**Whole group session 2**: Discuss and finalise outcomes to include in core outcome set
15:45	**Small group session 2**: Work in small groups to consider grouping of outcomes
16:25	**Whole group session 3**: Discuss and finalise grouping of domains and naming of these groups
16:55	Close of meeting

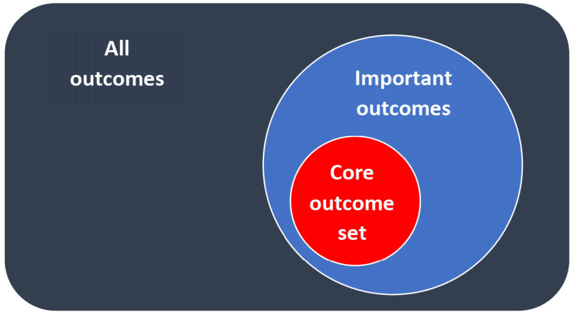
---

3. What is a core outcome set?

A core outcome set is an essential group of outcomes that should be measured, as a minimum, in **all clinical research** on a given health condition (we are focused on hand fractures and joint injuries for our meeting). Using a core outcome set would mean that it is easier to compare results from different research projects, because they will at least have the outcomes of the core outcome set in common. This helps us make best use of **all** the research that is being done.

A core outcome set does not restrict researchers from adding other outcomes that they consider important for their specific research projects. However, in order to be consistent across all our research, the core outcome set must be measured alongside whatever other outcomes are being used.

In this meeting we aim to produce a core outcome set for adults (age 18 or older) with a fracture (break) of a bone or any joint injury in the hand or wrist. As represented by the figure below, there are many different outcomes used in ongoing research on these injuries – several of these could be considered 'important' but the core outcome set will be what we decide are the **'critically important'** ones:



In this first stage of creating a core outcome set, we need to agree (reach consensus) on **'what'** to measure. We <u>will not</u> decide on 'how' to measure each outcome at this stage – this takes separate work to find out the best way to measure all the specific outcomes we select through the meeting. The meeting also <u>will not</u> involve deciding on 'when' each outcome should be measured, as that falls within the area of 'how' to measure. The output from this meeting is essential in helping to direct that future work, as we need to know 'what' to measure before we can think about 'how' to measure it.

4. Previous steps in this core outcome set project

There were a number of steps leading up to this point:

1. **A systematic review of treatment outcomes**

   This was a study in which we looked at 160 published studies to see what outcomes had been selected for previous clinical research. It gave us some idea of what outcomes have been used by researchers and healthcare professionals in the past.

2. **Interview and focus groups with patients**

   This study involved 25 interviews and 5 focus groups with patients. It helped us to learn what outcomes patients with hand fractures and joint injuries feel are relevant.

3. **Delphi consensus study**

   This three-round online questionnaire allowed us to learn more about the perspective of three key groups (patients, hand surgeons and hand therapists) in prioritising outcomes. 144 participants completed the study, rating each of the 41 outcomes 1-9 in terms of importance to measure in all future clinical research.

   Outcomes rated from 7-9 (very important) by 70% or more of participants in all three of the groups met the threshold for being automatically brought forward to this consensus meeting. This threshold had been set before the Delphi questionnaire study began, and although the results will guide our discussion it is through the consensus meeting that the final core outcome set will be decided.

The last part of the pack (Appendix – Page 12) shows more detailed information summarising what each outcome from the Delphi is meant to represent and the ratings it was given by patients, surgeons and therapists by the end of the study. This information is just meant to be a useful reference for you to use as you wish when thinking through some of the questions we raise in this pack.

For the meeting we want you to consider what should be in a core outcome set based on your own opinion but also bearing in mind the ratings of this large group of Delphi study participants.

---

5. Outcomes meeting Delphi consensus threshold

Several outcomes met the criteria we set to be counted as ‘very important’. This needed a rating of 7-9 by 70% or more of the participants in all three of the key groups – patients, surgeons and therapists. Listed in the table below are the outcomes that met this level of agreement as being ‘very important’. They are in descending order according to the average percentage scoring 7-9 across the three groups:

Outcome	**Percent scoring 7-9 in Delphi**		
	Patients (n=54)	**Surgeons** (n=55)	**Therapists** (n=35)
Fine hand use / dexterity	98	100	97
Return to usual work / job	94	98	100
Grasping and moving light objects	98	96	94
Pain / discomfort	94	100	94
Self-hygiene / personal care	98	95	94
Patient satisfaction with outcome / result	96	89	100
Pain / discomfort during activity	94	91	97
Stability of joints	98	83	100
Overall patient satisfaction	93	87	100
Pain / discomfort during rest	96	89	91
Treatment complications	85	93	97
Strength or muscle power	98	84	91
Eating and drinking	94	87	89
Range of movement	89	73	100
Change in sensation	85	87	89
Dressing	85	80	89
Sleeping	83	87	74
Being able to control a vehicle for transport	89	72	80
Cooking / preparing meals	81	71	83
Keyboard and mouse / touchscreen use	77	78	80

The ranked order above is just for information and to help prompt discussion at the meeting.

6. Outcomes not meeting Delphi consensus threshold

Here we summarise the outcomes which did not reach the threshold to automatically be entered for discussion in the meeting.

At the meeting, we will discuss and decide whether any of these outcomes should be included for further discussion in the meeting. Any outcomes not 'saved' at this stage will not be considered for the core outcome set.

Please look through the next few tables and decide whether any of the outcomes deserve to be considered further. Keep in mind that all these outcomes did not meet the threshold of importance set in the Delphi study. Any outcome from these tables should only be 'saved' if you feel it is critically important for the core outcome set.

Checking the detailed information (Appendix – Page 12) on any outcome you are thinking of 'saving' might help you with your decision.

The following outcomes **failed to meet the threshold** in **all** of the groups:

Outcome	Percent scoring 7-9 in Delphi		
	Patients (n=54)	**Surgeons** (n=55)	**Therapists** (n=35)
Intimate / sexual activity	35	39	20
Added expenses due to injury	39	24	40
Emotional / mood impact to friends and family	40	16	23
Appearance	43	38	26
Socialising / relationships with friends and family	50	44	49
Thinking and memory	50	31	23
Fatigue / tiredness	55	24	20
Walking / mobility	58	32	20
Shopping for groceries	59	29	29
General household chores	61	49	57
Using hands to change body position	68	58	46
Healthcare resources used	69	27	66

The outcomes below **met the threshold** in **only one** group:

Outcome	Percent scoring 7-9 in Delphi		
	Patients (n=54)	**Surgeons** (n=55)	**Therapists** (n=35)
Confirmation of bone healing and alignment (fractures only)	**88**	35	60
Confirmation of soft tissue healing	**87**	44	69
Carrying or lifting heavy objects	**72**	55	66
Speed of movement	**72**	25	49
Treatment inconvenience / discomfort	**72**	51	63
Looking after dependents within household	69	42	**80**
Loss of income	67	67	**71**

We have marked in bold red font where an outcome did reach 70% in a single stakeholder group.

Finally, the outcomes below **met the threshold** in **only two** groups:

Outcome	Percent scoring 7-9 in Delphi		
	Patients (n=54)	**Surgeons** (n=55)	**Therapists** (n=35)
Hobbies and recreation	**81**	65	**83**
Emotional / mood impact to self	**78**	64	**80**

We have marked in bold red font the two groups' scores where each outcome did reach 70%.

Please note: throughout this pack the numbers of participants for each stakeholder group shown near the top of any tables (i.e. n=…) were the total taking part in the study. For a few outcomes, a small number of participants chose the rating 'unable to score' – for these the percentages shown are based on the total number who did give a rating. If you are interested in the detailed number of participants who gave a rating for each outcome, we have provided this information in the Appendix (Page 12).

7. Which outcomes are critical for a core outcome set?

All the outcomes that have reached this stage are generally agreed to be very important, but **a core outcome set should include only the critically important outcomes**.

We need your help to decide which outcomes are **critically important** to include. The table below again shows all the outcomes that reached a threshold of 'very important' in the Delphi in descending order according to the average percentage scoring 7-9 across the three groups:

Outcome	Percent scoring 7-9 in Delphi		
	Patients (n=54)	**Surgeons** (n=55)	**Therapists** (n=35)
Fine hand use / dexterity	98	100	97
Return to usual work / job	94	98	100
Grasping and moving light objects	98	96	94
Pain / discomfort	94	100	94
Self-hygiene / personal care	98	95	94
Patient satisfaction with outcome / result	96	89	100
Pain / discomfort during activity	94	91	97
Stability of joints	98	83	100
Overall patient satisfaction	93	87	100
Pain / discomfort during rest	96	89	91
Treatment complications	85	93	97
Strength or muscle power	98	84	91
Eating and drinking	94	87	89
Range of movement	89	73	100
Change in sensation	85	87	89
Dressing	85	80	89
Sleeping	83	87	74
Being able to control a vehicle for transport	89	72	80
Cooking / preparing meals	81	71	83
Keyboard and mouse / touchscreen use	77	78	80

Please think about how you would sort these outcomes into three groups:

1. **Outcomes which are critical to include in the core outcome set**
2. **Outcomes which would be good to include in the core outcome set but aren't important enough to be considered critical**
3. **Outcomes which definitely don't need to be in a core outcome set**

You can base this on your own experience and also consider some of the information in the Appendix (Page 12).

---

8. Grouping of outcomes

Looking through the table on Page 5 you might have felt that some of the outcomes seem to link to others quite naturally. At the meeting we're going to decide whether some outcomes should be grouped together.

For example, 'pain/discomfort' could link to 'pain/discomfort during activity' and 'pain/discomfort during rest' as shown in red below. Or certain tasks might be linked through being everyday essential activities done to look after yourself as shown in blue below:

Fine hand use / dexterity

Return to usual work / job

Grasping and moving light objects

Pain / discomfort

Self-hygiene / personal care

Patient satisfaction with outcome / result

Pain / discomfort during activity

Stability of joints

Overall patient satisfaction

Pain / discomfort during rest

Treatment complications

Strength or muscle power

Eating and drinking

Range of movement

Change in sensation

Dressing

Sleeping

Being able to control a vehicle for transport

Cooking / preparing meals

Keyboard and mouse / touchscreen use

To decide how the outcomes should be grouped up, you might find it helpful to check the last part of this pack where we show the descriptions and examples for all the outcome domains (Appendix – Page 12). These are the same as we used in the Delphi study to help people understand what each outcome is meant to represent.

One way we have considered that the outcomes might be connected is shown below. This carries on from the examples shown on the previous page. Those outcomes which don't seem to fit all that clearly with any others are kept as individual outcomes instead of being placed in a group.

<u>Pain or discomfort</u>

Pain / discomfort

Pain / discomfort during activity

Pain / discomfort during rest

<u>Patient satisfaction</u>

Overall patient satisfaction

Patient satisfaction with outcome / result

Range of movement

Stability of joints

Change in sensation

Strength or muscle power

Treatment complications

<u>Activities of daily living</u>

Self-hygiene / personal care

Eating and drinking

Dressing

Sleeping

Cooking / preparing meals

<u>Higher level functions</u>

Grasping and moving light objects

Fine hand use / dexterity

Return to usual work / job

Keyboard and mouse / touchscreen use

Being able to control a vehicle for transport

What we have shown above is just a starting point for discussion Please think about any changes you would make to improve the grouping arrangement shown. You are free to make whatever changes you wish, including to the group labels (marked in bold and underlined).

For this part we do not need to worry about the importance of the outcomes towards a core outcome set. We are just focusing on whether or not it makes sense to group up each outcome with any others or to keep it separate.

Thank you for taking the time to read through this information pack

The rest of this pack is the Appendix, with explanations and examples to help make clear what each of the outcomes represents – this is the same wording that participants of the Delphi study saw. We have also provided the final Delphi study scores for each outcome

We look forward to seeing you at the online meeting

Appendix: Outcome descriptions, examples and final Delphi results

This part of the pack summarises key information about all the outcome domains. We have included the descriptions and examples as were used in the Delphi, which should help to make the meaning of each of the outcomes clear. The outcomes are split across two sections:

- Section 1 (starting on Page 13) – all outcomes which reached agreement of 'very important' in the Delphi
- Section 2 (starting on Page 23) – all the remaining outcomes

Within each section outcomes are listed in alphabetical order. The charts show the final result of the ratings for each one. The blue bars represent patients' scores, green ones the surgeons' scores and orange ones the therapists' scores.

We have included some anonymised quotes from the Delphi study in which participants summarised the reason behind the score they gave. We've also indicated which stakeholder group the quotes have come from. The quotes have generally been selected to reflect the typical score for the stakeholder group being quoted, but at times there were quite different views even within a single stakeholder group.

Please note many participants did not opt to comment on the outcomes regardless of the rating they gave, so the quotes are just a small window into some of the decision-making of those taking part. In some cases they have been slightly modified to improve readability but the underlying message is not altered.

The remainder of the pack comprised of pages summarising key Delphi survey data. A representative example page is shown on the right.

SECTION ONE

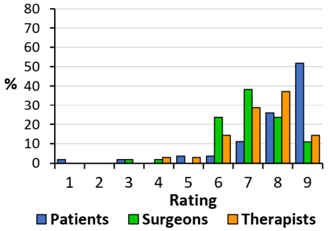
Outcomes which met threshold for 'very important' in Delphi study

Being able to control a vehicle for transport

Being able to drive a motor vehicle or ride a bicycle, whether for work, social or recreation reasons

Rated very important (7-9) by:

- 89% (48/54) of patients
- 73% (40/55) of surgeons
- 80% (28/35) of therapists



- "I'm worried about driving limitations" (Patient)
- "Many times patients have asked me about return to driving in the clinic" (Surgeon)
- "This is important to be independent" (Therapist)

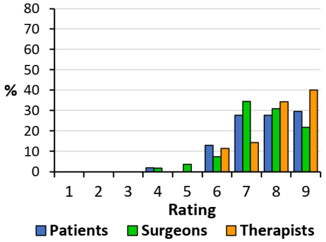
Change in sensation

Change from normal feeling in the hand / wrist (either sensations in the hand or in how things feel to touch)

e.g. numbness, pins and needles, tingling, being able to sense hot / cold through touch, not being able to properly sense how one's fingers or wrist are positioned, dysesthesia – in which the things that one touches feel different to normal

Rated very important (7-9) by:

- 85% (46/54) of patients
- 87% (48/55) of surgeons
- 89% (31/35) of therapists



- "I cannot distinguish clearly between round or flat objects with my injury, which is important in picking up the object I intended to" (Patient)
- "I think it's important for clinicians to know whether there is a change in sensation following treatment to help assess the effectiveness of the treatment / recovery" (Patient)
- "Important in the wider context of maintaining independence" (Surgeon)
- "Important/key for good hand function and safety with activities" (Therapist)

Appendix 15 Stage at which outcome domains were dropped from COS

Stage when domain dropped from COS	**Outcome domains**
Outcomes reaching consensus of 'very important' through Delphi survey but not reaching consensus for inclusion in COS at the meeting (presented in descending order according to final 'yes/no' vote results at meeting)	Eating and drinking (Reached threshold amongst professionals – 86% of professionals – but only 67% of patients) Treatment complications (Almost reached threshold amongst patients – 75% of patients – but only 67% of professionals) Strength or muscle power (Almost reached threshold amongst patients – 75% of patients – but only 53% of professionals) Dressing Sleeping Cooking / preparing meals Change in sensation Grasping and moving light objects Stability of joints Pain / discomfort Being able to control a vehicle for transport Keyboard and mouse / touchscreen use Overall patient satisfaction
Outcomes reaching threshold of 'very important' for patient and therapist groups in Delphi survey (but not surgeon group)	Hobbies and recreation Emotional / mood impact to self
Outcomes reaching threshold of 'very important' for patient group in Delphi survey (but not therapist or surgeon groups)	Confirmation of bone healing and alignment (fractures only) Confirmation of soft tissue healing Carrying or lifting heavy objects Speed of movement Treatment inconvenience / discomfort
Outcomes reaching threshold of 'very important' for therapist group in Delphi survey (but not patient or surgeon groups)	Looking after dependents within household Loss of income