

BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

International multi-phase mixed methods study protocol to develop a cross-cultural patient-reported outcome and experience measure for hand conditions (HAND-Q)

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2018-025822
Article Type:	Protocol
Date Submitted by the Author:	03-Aug-2018
Complete List of Authors:	Sierakowski, Kyra; Flinders University, College of Medicine and Public Health ; Flinders Medical Centre, Plastic & Reconstructive Surgery Dean, Nicola; Flinders University, College of Medicine & Public Health; Flinders Medical Centre, Plastic & Reconstructive Surgery Pusic, Andrea; Brigham and Women's Hospital Biomedical Research Institute Cano, Stefan; Modus Outcomes Griffin, Philip; Flinders Medical Centre, Plastic & Reconstructive Surgery Bain, Gregory; Flinders University, College of Medicine and Public Health ; Flinders Medical Centre, Orthopaedic Surgery Lalonde, Donald; Dalhousie University, Division of Plastic Surgery Klassen, Anne; McMaster University, Pediatrics
Keywords:	patient reported outcome measures, Hand & wrist < ORTHOPAEDIC & TRAUMA SURGERY, psychometrics, QUALITATIVE RESEARCH, PLASTIC & RECONSTRUCTIVE SURGERY

SCHOLARONE™
Manuscripts

International multi-phase mixed methods study to develop a cross-cultural patient-reported outcome and experience measure for hand conditions (HAND-Q)

Authors

1. Kyra L Sierakowski MD MS, College of Medicine & Public Health, Flinders University, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, kyra.sierakowski@flinders.edu.au
2. Nicola R Dean FRACS (Plas) PhD, Department of Plastic & Reconstructive Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, Nicola.dean@sa.gov.au
3. Andrea L Pusic FRCSC (Plas) MD MSc, Brigham and Women's Hospital, 75 Francis S, Boston, MA, USA, 02116, apusic@bwh.harvard.edu
4. Stefan Cano PhD CPsychol, Modus Outcomes, Suite 210b, Spirella Building, Letchworth Garden City, SG6 4ET, UK, stefan.cano@modusoutcomes.com
5. Philip A Griffin FRACS (Plas) MBBS, Department of Plastic & Reconstructive Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia philip.griffin@sa.gov.au
6. Gregory I Bain FRACS (Ortho) PhD, Department of Orthopaedic Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, greg@gregbain.com.au
7. Don Lalonde, FRCSC (Plas) MD, Division of Plastic Surgery, Dalhousie University, Saint John, Canada, drdonlalonde@nb.aibn.com
8. Anne F Klassen DPhil, McMaster University, 3N27, 1280 Main Street W, Hamilton, ON, Canada, L8N 3Z5, aklass@mcmaster.ca

Corresponding Author

Kyra L Sierakowski MD MS, College of Medicine & Public Health, Flinders University, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, kyra.sierakowski@flinders.edu.au +61431306997

Keywords

Psychometrics, hand, patient reported outcome measures, qualitative research, plastic and reconstructive surgery

Funding Statement

This work has received funding from the Australasian Foundation for Plastic Surgery Plastic & Reconstructive Surgery (PRS) Research Scholarship and the Royal Australasian College of Surgeons, Foundation for Surgery Small Project Grant. The authors have no financial interest to declare in relation to the content of this article. The article processing charge was paid from the PRS Research Scholarship.

Competing Interests Statement

The authors have no competing interests to report.

Word count = 3807

Author Contributions

- KS prepared the manuscript with the oversight of ND, GB, PG, DL, AK.

For peer review only - <http://bmjopen.bmj.com/site/about/guidelines.xhtml>

- AK, AP and SC developed the methodology utilised for the development of the HAND-Q and guided the content of the protocol.
- All authors reviewed the protocol manuscript and gave valuable feedback to the manuscript.

Acknowledgements

The authors acknowledge our partners in hand care centres around the world for participating in data collection in various phases of the study.

For peer review only

Abstract

Introduction

Patient reported outcome measures (PROMs) are instruments used to measure outcomes and experiences of healthcare from the patient perspective. The specific methodology used to develop a PROM should be communicated to establish the quality of the instrument. This mixed methods protocol describes the development of a cross-cultural, internationally applicable PROM for hand conditions, the HAND-Q.

Methods and analysis

The multi-phase approach used for this study has been previously used with the development of other PROMs by our team (e.g., BODY-Q, BREAST-Q, CLEFT-Q, FACE-Q). In Phase I, we establish what important concepts matter to patients with hand conditions. A conceptual framework is developed from a systematic review of existing PROMs in the field and an extensive international qualitative study. Interpretive description is the qualitative approach used. Item generation is based on the qualitative data. The preliminary scales will be created for each theme identified in the conceptual framework. These scales will be refined by cognitive debriefing interviews with participants and expert input. Phase II involves a large international sample of patients with varied hand conditions completing the field-test version of the HAND-Q. The scales will be refined using the modern psychometric approach of Rasch Measurement Theory (RMT). Analysis will result in a shortened set of clinically meaningful and scientifically robust HAND-Q scales.

Ethics and dissemination

This study is coordinated at Flinders University (Adelaide, Australia) where it has ethics board approval for Phase I and Phase II. Findings will be published in peer-reviewed journals and presented at local, national and international conferences.

Article Summary

Strengths and Limitations of this study

- The inclusion of a diverse international sample of participants ensures that the HAND-Q will serve the global community.
- Including the full spectrum of hand conditions and injuries with minimal participant exclusion criteria will ensure that the HAND-Q is broadly applicable to any form of congenital, traumatic or degenerative hand condition.
- Independently functioning scales will allow tailoring of scales to the patient, study or clinical setting, which will reduce patient and administrative burden.
- The comprehensiveness of the HAND-Q will provide a range of scales for use in clinical, research and quality assurance applications.
- Use of a modern psychometric approach will produce HAND-Q scales that are both scientifically sound and clinically meaningful.

Introduction

Patient-reported outcome (PRO) data is increasingly collected and used in the assessment of treatment effectiveness and healthcare funding(1, 2). It is important that PRO data is captured with appropriately robust instruments that satisfy best practice guidelines(3, 4). The methodology used to develop and validate a PROM can be complex; however, it is important for researchers and clinicians to understand the approach used in order to judge the appropriateness of an instrument for their specific application(5, 6).

There are a wide variety of conditions and injuries that affect the hand, and each has the potential to have an enormous impact on an individual's daily function, their livelihood and their quality of life. Conditions include chronic degenerative bony diseases, such as osteoarthritis and rheumatoid arthritis, chronic progressive soft tissue disorders, such as Dupuytren's contracture, and acute injuries of the hand. Hand injuries occur commonly within the workplace and in the home environment and can range from a simple skin laceration to complex tendon and bone damage. This heterogeneity of conditions and causes can make the measurement of treatment outcomes extremely challenging and can limit the clinical applicability of condition-specific PROMs.

There is growing awareness of the economic and environmental impacts of medical care, particularly in the surgical sphere where access to the operating theatre is competitive and expensive(1, 7, 8). These forces have led to an increasing amount of hand surgery performed outside of operating theatres with a fully awake patient and the use of only local anaesthetic(9). This wide-awake approach minimises the economic burden of hand surgery(10, 11). It is important to be able to measure patient satisfaction with their experience of hand surgery anaesthesia using an appropriately designed PROM to allow for comparisons to be made of the different anaesthetics approaches to hand surgery. The HAND-Q will explore experience of care concepts that have not been addressed by existing PROMs.

PROMs developed using a modern psychometric approach have the potential to function consistently regardless of the population that is being measured. This is possible because the psychometric properties of the PROM are that of the instrument itself, not a reflection of the population that is being tested. This approach to measurement allows the same PROM to be used in different populations with results comparable on the same metric. Legacy instruments developed using classical test theory methodology cannot be used in this manner. A further benefit of the modern psychometric approach is that measurement is given in interval format, which permits accurate and meaningful measurement of change, which is particularly important in the surgical field.

The vast majority of PROMs are developed in a single language and usually within a single country. As hand conditions are prevalent around the world, this approach means that there is limited content validation within different cultures and economic environments. Many PROMs used in hand surgery have been developed in English and then translated into multiple other languages. Outcome of Surgery-Hand/Arm (POS-

1 Hand/Arm)(12) and the Patient Evaluation Measure (PEM)(13) were both developed in
2 the United Kingdom. The Disabilities of the Arm, Shoulder and Hand (DASH)(14) and
3 the Michigan Hand Questionnaire (MHQ)(15) were developed in North America and
4 subsequently translated into many languages(16-19). Although guidelines exist to
5 ensure appropriate cross-cultural adaption of PROMs(20, 21), a more rigorous
6 approach is to involve people from different countries of varied economic status in the
7 development and validation of a new instrument, rather than performing translation
8 retrospectively(5). An instrument developed in a cross-cultural manner could
9 encourage and facilitate multinational research in the field of hand surgery.
10
11
12

13
14 The process to develop a cross-cultural PROM that is both scientifically sound and
15 clinically relevant is a complex undertaking. The following protocol details the
16 methodology that will be used for the international study to develop the HAND-Q.
17

18 **Methods and Analysis**

19 **Overview of PROM development**

20
21 We use a modern psychometric approach and engage patients and experts in all phases
22 of the development to create clinically meaningful and scientifically robust PROM
23 scales. The development of the HAND-Q follows the international best practice
24 guidelines as determined by the Scientific Advisory Committee of the Medical Outcomes
25 Trust(4), the USA Food and Drug Administration(3) and the International Society for
26 Pharmacoeconomics and Outcomes Research (ISPOR)(22, 23). We have previously
27 published(5) the 3-phase mixed methods approach that we employ to develop a PROM,
28 a summary is reproduced in Figure 1. This approach establishes a PROM that satisfies
29 the minimum requirements of reliability and validity as set forth by the International
30 Society for Quality of Life Research (ISOQOL)(24) and the Consensus-based Standards
31 for the Selection of Health Status Measurement Instruments (COSMIN)(6, 25). The
32 application of this 3-phase approach to the development of the HAND-Q will be
33 expanded upon in this paper.
34
35
36
37
38

39 **Phase I – What should we measure?**

40 The objective of Phase I is to develop preliminary scales founded on excellent content
41 validity. A systematic literature review was performed to identify the currently
42 available PROMs relevant to hand conditions and hand surgery(26). This
43 comprehensive literature search was carried out according to PRISMA guidelines with a
44 total of 3039 citations retrieved. A total of 24 instruments were reviewed; 10 regional
45 upper limb, 6-hand/wrist specific and 8 hand condition-specific instruments. Review of
46 the developmental methodology of those existing instruments revealed areas of
47 significant weakness(26). The review also found profound content limitations with no
48 hand-specific instrument found that covers the full spectrum of outcomes relevant to
49 hand surgery. Based on this review it was determined that there was need for a
50 comprehensive PROM for hand surgery patients and we proceeded with this study.
51 Concepts from other PROMs were used to form a preliminary interview guide.
52
53
54
55
56
57
58
59
60

Conceptual Framework

The starting point in any Phase I PROM development project is the development of a conceptual framework(4), which includes a description of the concepts of interest and the relationship between these concepts within the population that the instrument is designed for. Based on the findings of the systematic review, concepts from existing instruments are mapped along with other concepts of interest to create a preliminary conceptual framework.

Qualitative Study

The preliminary conceptual framework will be used as a basis for an interview guide to establish which of the concepts are meaningful to patients. We use an approach from applied health services research known as Interpretive Description(27, 28). This approach acknowledges that there is pre-existing theoretical and clinical knowledge informing the study. In addition, this approach aims to produce knowledge relevant to the clinical context with the proviso that the individual's understanding of a concept is of the greatest importance, regardless of the clinical or theoretical explanation(29).

Participants, setting and recruitment

Phase 1 qualitative study will recruit participants from centres in Adelaide, Australia and Saint John, New Brunswick, Canada. Recruitment will be by direct referral from clinicians or administrative staff. Participants will be eligible for inclusion if they had experienced surgery on either or both of their hands over the preceding twelve months, with a minimum of 4 weeks since their surgery. Exclusion criteria for this phase of the study will be the inability to speak English or cognitive delay that prevents participation in a semi-structured interview.

Sampling

Participants will be purposively sampled to include a heterogeneous population with respect to age, gender, hand condition, setting of surgery (hospital operating theatre vs private rooms), funding for surgery (public vs private) and whether surgery was performed with general anaesthesia, sedation or local anaesthetic. Sampling will continue until saturation is achieved, i.e. no further important concepts are identified in three consecutive interviews(30).

Data collection

Written consent will be obtained from all participants before commencing with individual, semi-structured interviews. Participant age, gender, hand condition and date of surgery will be collected. Interviews will be audio-recorded and transcribed verbatim with all identifiers removed. The interview guide (Table 1) used is based on concepts identified from the literature review [26]; this guide includes a list of open-ended questions to encourage discussion. The interviewer will probe for new concepts during the interviews. An iterative approach will be used, whereby data collection and analysis will take place concurrently in order to ask about newly identified relevant concepts.

Data analysis

A line-by-line approach will then used to code the data, with a process of constant comparison utilised to identify common concepts of interest(31). A codebook will be developed that outlines the codes for evolving domains and themes. The information

1 from these interviews will be used to shape and reconfigure the conceptual framework
2 to represent outcomes and experiences important to participants.
3

4 *Rigor*

5 Several strategies have been put in place to ensure rigor of this qualitative study. A
6 single interviewer will perform all of the qualitative interviews. One team member will
7 perform all of the coding of the transcripts, which will then corroborated by a second
8 team member. The iterative process allows for member-checking to confirm if concepts
9 are deemed to be valuable to subsequent participants(32). The members of the study
10 team will discuss data analysis and this peer-debriefing will provide consistency(32).
11
12
13

14 *Item generation*

15 The evolved conceptual framework will be used to identify which scales should be
16 generated. A comprehensive list of potential items will be generated from the coding
17 process. Each of the domains identified will have multiple themes; each important
18 theme will be developed into a preliminary scale. This approach will ensure that the
19 suite of scales that comprise the HAND-Q will cover all of the important concepts of
20 interest identified by patients with hand conditions. The scales are constructed from the
21 items identified in the coding process; participants' language is kept intact as much as
22 possible.
23
24
25

26 *Refining the preliminary scales*

27 Incorporating feedback from both patients and experts in hand conditions will enhance
28 the quality of the preliminary HAND-Q scales. Participants who have previously
29 participated in qualitative interviews as well as new participants will be invited to
30 participate in cognitive debriefing interviews. Input will also be sought from a
31 multidisciplinary international pool of experts who will be invited to provide feedback
32 via an online survey using Research Electronic Data Capture (REDCap).
33
34
35

36 *Cognitive debriefing interviews*

37 After the preliminary scales are constructed, they will be shown to participants with
38 hand conditions in the form of semi-structured cognitive interviews. These interviews
39 will be conducted by telephone, audio-recorded, transcribed and analysed. The purpose
40 of these interviews is to ensure that participants have understood the instructions,
41 items and the response options. The participant will be asked to go through each of the
42 scales using the "think aloud" technique, while the interviewer notes the specific
43 content that causes problems(33, 34). Feedback will be sought on the relevance of the
44 items to the individual. If an item is unclear to a participant, they will be asked for
45 suggestions on how we can improve item wording. Further probing will be used to
46 enquire about any new content that may have been overlooked. Cognitive interviews
47 will be performed in three rounds to provide the opportunity to make changes between
48 rounds. The first round will include participants from the qualitative interviews from
49 both the Australian and Canadian cohorts. Subsequent rounds will also include new
50 participants from the USA to ensure content validity for patients in that country.
51
52
53
54
55

56 *Expert clinician input*

57 It is useful to seek feedback from clinicians who specialise in the management of hand
58 conditions to ensure that the scales capture all clinically relevant concepts. Experts in
59

1 the management of patients with hand conditions from the fields of Plastic &
2 Reconstructive Surgery, Orthopaedic Surgery, Hand therapy and Physiotherapy will be
3 invited to review the HAND-Q in the form of an online survey. We will administer this
4 survey using REDCap, which is a secure web-based data collection system(35).
5 Researchers and academics in the field will also be invited to participate. An
6 international sample of experts will be included to ensure that we engage with
7 professionals working in diverse healthcare systems with varied treatment strategies.
8 Experts will be asked to provide feedback on all aspects of each scale (instructions,
9 response options, items) and to nominate missing items. Feedback from experts will be
10 used to refine the scales. Expert input will be obtained before the final round of
11 participant cognitive interviews in order to show any changes made from expert input
12 to participants.

17 **Translation**

18 To facilitate an international field-test, cultural and linguistic validation of the HAND-Q
19 into multiple languages will take place. The field-test scales will be translated into the
20 required field-test languages according to the guidelines set forth by the International
21 Society for Pharmacoeconomics and Outcomes Research (ISPOR)(20) and the World
22 Health Organization (WHO)(21) recommendations for linguistic validations. In
23 summary, this approach requires the forward translation (English to target language) to
24 be performed by two independent bilingual individuals, with back translation (target
25 language to English) to be performed by a third individual. Any discrepancies are
26 resolved at each step and the resultant version is then shown to a small group of
27 patients to ensure that the translation is valid and ready for use.

32 **Phase II – How should we measure the concepts identified in Phase I?**

33 A large heterogeneous sample of people aged 18 years and older with any type of hand
34 condition will be recruited. Analysis of the field-test data allows the scales to be refined
35 to include the subset of items that are the most effective in measuring the concepts of
36 interest. We will follow the modern psychometric method of Rasch Measurement
37 Theory (RMT)(36). This approach dictates that in order to achieve accurate
38 measurement the data must comply with the Rasch measurement model(37).
39 Performing RMT analysis will make it possible to identify poorly functioning items
40 and to only retain the best subset of items that together map out a clinical hierarchy for
41 each scale. The details of the psychometric parameters used in RMT analysis are
42 described in detail elsewhere(38). Scales refined using RMT analysis can provide
43 measurement that is suitable for application at an individual patient level, unlike
44 instruments developed using classical test theory, which are designed for analysis of
45 groups of individuals. RMT produces interval level measurement where the intervals
46 between adjacent scores are equal throughout the breadth of the possible scores.
47 Subsequently measuring change over time (such as comparing pre-operative and post-
48 operative scores which is a common PROM application in the surgical field) with RMT
49 designed scales is mathematically sound(39). Another benefit of scales derived using
50 the RMT approach is that the scales function consistently, independent of the
51 population that is being studied, which allows comparison of scores between different
52 populations.

International field-test and RMT analysis

The aim of the international field-test study is to collect scale responses from a large cohort of patients with hand conditions to allow the refinement of the scales using RMT and examine the psychometric properties of the resulting scales.

Study participants

The international field-test will include participants from multiple English and non-English speaking countries. Field-test sites will be included based on their interest in participating and the feasibility of recruiting an adequate sample size over the anticipated timeframe. We aim to include sites in five countries and with 200 participants recruited per country (n=1000). Participants will include anyone aged 18 years or older with a hand condition that is able to comprehend the study and legally provide consent. Participants will be recruited in clinic settings and asked to complete the scales while waiting for their appointment.

Data collection

Demographic information will be collected to allow for sub-cohort analysis. Participants will be asked to complete the HAND-Q scales either using a paper booklet or electronic device (smartphone or tablet) using a REDCap survey.

Data analysis

If participants complete the scales using a paper booklet, members of the research team will enter these data into the REDCap database. Completed data files will be downloaded into IBM SPSS Statistics Version 25. The SPSS file will be formatted suitably and imported into RUMM2030 for RMT analysis. Each of the scales will be analysed separately to determine its psychometric properties according to the following criteria. The thresholds for item response options must be appropriately ordered. This means that a response of '1' sits lower on a continuum than '2' etc. The hierarchy of items on the scale is then determined, from the 'easiest' to endorse to the 'hardest' to endorse. Several item fit statistics are used to determine if the scale works adequately; (1) log residuals, which reflect the item-person interaction, (2) chi squared statistic, which reflect the interaction between the item and the concept being measured and (3) item characteristic curves which demonstrate graphically the relationship between groups of responders and the trait of interest(40). Each item will be assessed using these parameters and excluded from the scales if they do not perform well. Lastly the targeting of the scale within the patient population is reviewed to ensure that the scale is able to measure the full breadth of the construct within the population of interest. The next requirement in the assessment of the scales is to ensure that the items display internal consistency; that is that they are appropriately inter-related. The first step is testing for unidimensionality; which is checking if all of the items that contribute to the scale are measuring the same underlying construct. The scale is then evaluated using the Person Separation Index, which is a similar measure to Cronbach alpha in classical test theory. This process of analysis is iterative; scales that are not functioning optimally are improved by excluding items with poor performance. This process is completed when each of the scales show the aforementioned statistics within the acceptable range.

Differential item functioning

As RMT produces a scale that behaves in the same manner regardless of the population for which it is used, differences between sub-groups of the population can be identified. When one sub-population answers an item differently to another sub-population, that item is said to display differential item functioning (DIF). Inclusion of multiple countries in the international field-test will allow for any differences based on country of origin to be considered. The software RUMM2030 can identify DIF and any items that display DIF can be excluded during the item reduction process, or other adjustments made to account for any differences.

Item reduction

The location of each item on the difficulty spectrum of each scale helps to determine if there is any redundant items; i.e. two or more items that measure the same point on the difficulty spectrum. Items will be reduced if they are identified as poorly functioning by the previously mentioned statistics or if they are redundant. Items will be reduced until the optimal number of items remains. This approach will be determined by a combination of the distribution of item locations as well as the clinical requirement for the degree of accuracy. Once this process is complete each of the scales is finalised. A scoring table is then produced for each scale using the RMT analysis; the cumulative score of a scale is more complex than the simple addition of individual items.

Construct validity

Following the finalisation of scales the logits will be used to transform scale scores for each participant from 0 (worst) to 100 (best). Basic relationships between scores and demographic characteristics can then be calculated using independent samples t-tests and analysis of variance (ANOVA) or the equivalent nonparametric tests as appropriate in SPSS.

There are various components of construct validity, such as structural validity, hypothesis testing and cross-cultural validity. RMT analysis of unidimensionality addresses the structural validity. Analysis for DIF addresses the cross-cultural validity of the scales. A priori hypothesis testing of known groups is a technique to assess whether the responses to the scales correlate or differ between groups in a way that would be expected(41). In the HAND-Q, we intend to test the following hypotheses using ANOVA in SPSS:

1. Those patients with hand conditions that affect the appearance of the hand (such as rheumatoid arthritis and Dupuytren's contracture) will be more dissatisfied with their hand appearance than those with carpal tunnel syndrome (which rarely has a significant affect on hand appearance).
2. Those patients requiring further intervention for their hand condition would have lower quality of life scores than those who do not require further intervention.
3. Those patients who describe their hand condition as "severe" will have lower quality of life scores than those who describe their conditions as "mild".

Phase III: how does the instrument work?

In developing the HAND-Q we will comply with the COSMIN checklist(42). Many of the required components will be addressed in preceding phases. The third and final phase includes further psychometric testing to establish the reliability, validity and responsiveness of the HAND-Q. All testing within this phase will utilise the final versions of the HAND-Q scales.

Patient and public involvement

Our patient-oriented approach engages patients and healthcare providers in all stages of our research as experts and research team members. Their input is fundamental to the design of the study and development of content for HAND-Q scales. All participants in the initial qualitative interviews will be invited to continue to collaborate in our study by taking part in cognitive interviews where they can provide feedback on our findings and help to refine the final set of scales. We will disseminate updates on the HAND-Q development to the patient community and healthcare providers via local health newsletters with links and information about presentations and publications.

Ethics

This international study is coordinated from Flinders University where the Southern Adelaide Clinical Human Research Ethics Committee has approved Phase I and II. Ethics board approval will be obtained prior to the study commencing at a collaborating site. Participants will provide written or oral consent prior to participating. As participants will be asked to discuss their experiences, they could potentially experience some distress. Participants will be advised that should this occur that the interviewer can put them in touch with an appropriate skilled clinician. Participants will be assured that any personal information will be kept confidential. In the qualitative phase all identifying information will be removed during the process of transcription. Any personal information required for follow up of individual participants will be kept secure and confidential following institution rules for data storage.

Dissemination

To ensure wide uptake of the HAND-Q, once developed it will be made available free of charge for non-commercial use. Our team will actively promote the HAND-Q at local, national and international conferences. We will publish findings for the HAND-Q in journals known to be valued and read by our target audiences. This study has no intention to directly compare outcomes from different centres. Any published work or public presentations resulting from this study will not identify specific centres. Collaborations with multiple sites internationally is hoped to increase the application of the HAND-Q in the future. The individual scored HAND-Q data from Phase II and Phase III will be returned to the providing sites for their own use.

Figure 1 Flow diagram illustrating the multiphase mixed methods approach to the development of the HAND-Q.

QUAN, quantitative study component; QUAL, qualitative study component.
Image reproduced from Wong Riff et al.(5)

For peer review only

Table 1 Interview guide for qualitative interviews to be performed in Phase 1.**Experience of Care:**

1. What treatments have you had for your condition?
2. What was good or bad about the treatment?
3. If the participant has had surgery –
 - a. What was your experience of the anaesthetic used? Probe: general anaesthetic, block, local
 - b. Would you have considered having treatment under local anaesthetic? Probe: why, why not
4. Who do you see when you come to the hospital clinic? Probe: receptionist, nurse, doctor, occupational therapist
5. What are the people like who care for you? Probe: friendly, made you feel comfortable, easy to talk to, listened to you
6. What kind of verbal and written information did they give you? Probe: gave enough information, let you ask questions, answered your questions, provided information about recovery

Physical Function:

7. Does your condition create any functional problems? Probe: work, personal care, hobbies
8. What specific things do you have difficulty with due to your hand problem? Probe: getting dressed, cooking, typing, sport
9. Do you experience any symptoms related to your functional problem? Probe: pain, discomfort, embarrassment, mood disturbance

Psychological Wellbeing:

10. How does your hand problem make you feel? Probe: frustrated, angry, upset, worried, stressed
11. How does your hand problem make you feel about yourself? Probe: self esteem, body image, confidence, self-conscious, different from others

Appearance:

12. How would you describe the appearance of your hand/s? Probe: from close up, from far away, symmetry, texture, attractiveness
13. How has your hand appearance changed since your treatment? Probe: scarring, descriptive detail
14. What do you like or dislike about your hand appearance?
15. Is there anything about your hand appearance that you would like to change? Probe: for details
16. Do you ever hide your hands? How do you do this?
17. How important is the appearance of your hands to you?

Other:

18. Is there anything I have not asked you that you think it is important for me to know?
19. Would you like to receive a copy of the transcript from today's discussion?
20. Would you be interested in participating in Cognitive Interview?

References

1. Porter ME. What is value in health care? *New England Journal of Medicine*. 2010;363(26):2477-81.
2. Coronini-Cronberg S, Appleby J, Thompson J. Application of patient-reported outcome measures (PROMs) data to estimate cost-effectiveness of hernia surgery in England. *Journal of the Royal Society of Medicine*. 2013;106(7):278-87.
3. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. p. 65132-3.
4. Lohr KN. Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research*. 2002;11(3):193-205.
5. Wong Riff KW, Tsangaris E, Goodacre T, Forrest CR, Pusic AL, Cano SJ, et al. International multiphase mixed methods study protocol to develop a cross-cultural patient-reported outcome instrument for children and young adults with cleft lip and/or palate (CLEFT-Q). *BMJ open*. 2017;7(1):e015467.
6. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*. 2010;19(4):539-49.
7. Leblanc MR, Lalonde J, Lalonde DH. A detailed cost and efficiency analysis of performing carpal tunnel surgery in the main operating room versus the ambulatory setting in Canada. *Hand*. 2007;2(4):173-8.
8. Thoma A, Strumas N, Rockwell G, McKnight L. The use of cost-effectiveness analysis in plastic surgery clinical research. *Clinics in plastic surgery*. 2008;35(2):285-96.
9. Lalonde DH. Wide-awake flexor tendon repair. *Plast Reconstr Surg*. 2009;123(2):623-5.
10. Van Demark RE, Smith VJ, Fiegen A. Lean and Green Hand Surgery. *Journal of Hand Surgery*. 2018;43(2):179-81.
11. Rhee PC, Fischer MM, Rhee LS, McMillan H, Johnson AE. Cost Savings and Patient Experiences of a Clinic-Based, Wide-Awake Hand Surgery Program at a Military Medical Center: A Critical Analysis of the First 100 Procedures. *The Journal of Hand Surgery*. 2016.
12. Cano SJ, Browne JP, Lamping DL, Roberts AH, McGrouther DA, Black NA. The Patient Outcomes of Surgery-Hand/Arm (POS-Hand/Arm): a new patient-based outcome measure. *Journal of Hand Surgery - British Volume*. 2004;29(5):477-85.
13. Macey AC, Burke FD. Outcomes of hand surgery. *Journal of Hand Surgery*. 1995;20 B(6):841-55.
14. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. *The Upper Extremity Collaborative Group (UECG)*. [Erratum appears in *Am J Ind Med* 1996 Sep;30(3):372]. *American Journal of Industrial Medicine*. 1996;29(6):602-8.
15. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *Journal of Hand Surgery - American Volume*. 1998;23(4):575-87.
16. Roh YH, Yang BK, Noh JH, Baek GH, Song CH, Gong HS. Cross-cultural adaptation and validation of the Korean version of the Michigan hand questionnaire. *Journal of Hand Surgery - American Volume*. 2011;36(9):1497-503.
17. Oksuz C, Akel BS, Oskay D, Leblebicioglu G, Hayran KM. Cross-cultural adaptation, validation, and reliability process of the Michigan Hand Outcomes Questionnaire in a Turkish population. *Journal of Hand Surgery - American Volume*. 2011;36(3):486-92.
18. Schonemann JO, Larsen K, Hansen TB, Soballe K. Reliability and validity of the Danish version of the disabilities of arm, shoulder, and hand questionnaire in patients with fractured wrists. *Journal of Plastic Surgery and Hand Surgery*. 2011;45(1):35-9.
19. Germann G, Harth A, Wind G, Demir E. [Standardisation and validation of the German version 2.0 of the Disability of Arm, Shoulder, Hand (DASH) questionnaire]. *Unfallchirurg*. 2003;106(1):13-9.
20. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee - Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient - reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in health*. 2005;8(2):94-104.
21. WHO (World Health Organization). Process of translation and adaptation of instruments. http://www.who.int/substance_abuse/research_tools/translation/en/
22. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. *Value in Health*. 2011;14(8):967-77.
23. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. *Value in Health*. 2011;14(8):978-88.

- 1 24. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends
2 minimum standards for patient-reported outcome measures used in patient-centered outcomes and
3 comparative effectiveness research. *Quality of Life Research*. 2013;22:1889-905.
- 4 25. Zywił MG, Mahomed A, Gandhi R, Perruccio AV, Mahomed NN. Measuring Expectations in
5 Orthopaedic Surgery: A Systematic Review. *Clinical Orthopaedics and Related Research*. 2013;471(11):3446-
6 56.
- 7 26. Sierakowski K, Evans Sanchez, KA, Damarell, RA, Dean, NR, Griffin, PA, Bain, GI. Measuring Quality of
8 Life & Patient Satisfaction in Hand Conditions: A systematic review of currently available patient reported
9 outcome instruments. *Australasian Journal of Plastic Surgery*. 2018;(in press).
- 10 27. Thorne S, Kirkham SR, MacDonald-Emes J. Focus on qualitative methods. Interpretive description: a
11 noncategorical qualitative alternative for developing nursing knowledge. *Research in nursing & health*.
12 1997;20(2):169-77.
- 13 28. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative
14 alternative for developing nursing knowledge. *Res Nurs Health*. 1997;20.
- 15 29. Claire Welford R, Dip N, Kathy Murphy R, Dympna Casey R. Demystifying nursing research
16 terminology. Part 1. *Nurse Researcher (through 2013)*. 2011;18(4):38.
- 17 30. Sandelowski M. Theoretical Saturation. In: Given LM, editor. *The SAGE Encyclopedia of Qualitative
18 Research Methods*. Thousand Oaks: SAGE Publications, Inc.; 2008.
- 19 31. Pope C, Ziebland S, Mays N. Analysing qualitative data. *British medical journal*. 2000;320(7227):114.
- 20 32. Cohen DJ, Crabtree BF. Evaluative criteria for qualitative research in health care: controversies and
21 recommendations. *The Annals of Family Medicine*. 2008;6(4):331-9.
- 22 33. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Quality of life research*.
23 2003;12(3):229-38.
- 24 34. Van Someren M, Barnard Y, Sandberg J. The think aloud method: a practical approach to modelling
25 cognitive. 1994.
- 26 35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture
27 (REDCap)—a metadata-driven methodology and workflow process for providing translational research
28 informatics support. *Journal of biomedical informatics*. 2009;42(2):377-81.
- 29 36. G R. *Studies in mathematical psychology: 1*. In: Institut. Dp, editor. 1960.
- 30 37. Cano SJ, Hobart JC. The problem with health measurement. *Patient Prefer Adherence*. 2011;5:279-90.
- 31 38. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role
32 of new psychometric methods: Prepress Projects; 2009.
- 33 39. Cano SJ, Klassen A, Pusic AL. The science behind quality-of-life measurement: a primer for plastic
34 surgeons. *Plastic and reconstructive surgery*. 2009;123(3):98e-106e.
- 35 40. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the
36 Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*. 2007;46(1):1-18.
- 37 41. Strauss ME, Smith GT. Construct validity: Advances in theory and methodology. *Annu Rev Clin
38 Psychol*. 2009;5:1-25.
- 39 42. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for
40 evaluating the methodological quality of studies on measurement properties: a clarification of its content.
41 *BMC medical research methodology*. 2010;10(1):22.

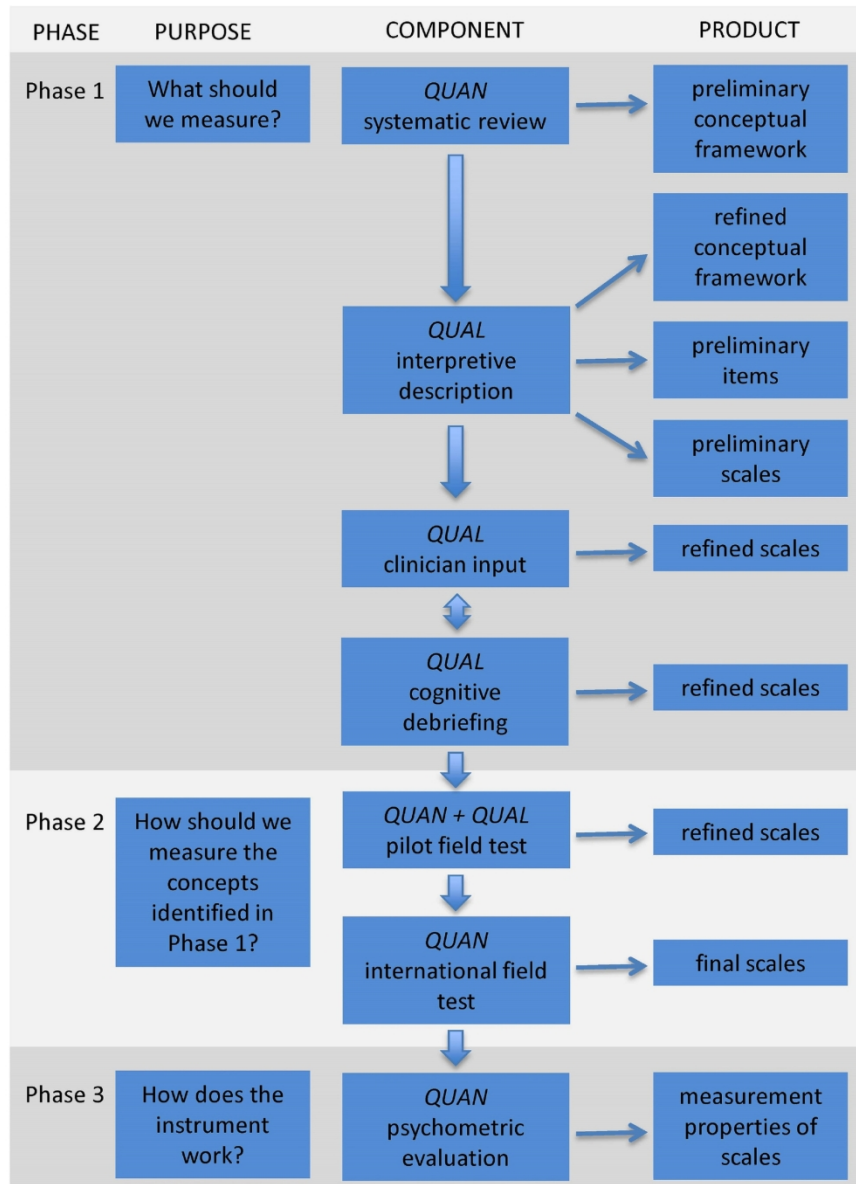


Figure 1 Flow diagram illustrating the multiphase mixed methods approach to the development of the HAND-Q.
 QUAN, quantitative study component; QUAL, qualitative study component.
 Image reproduced from Wong Riff et al.(5)

175x241mm (300 x 300 DPI)

BMJ Open

International multi-phase mixed methods study to develop a cross-cultural patient-reported outcome and experience measure for hand conditions (HAND-Q)

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2018-025822.R1
Article Type:	Protocol
Date Submitted by the Author:	15-Dec-2018
Complete List of Authors:	Sierakowski, Kyra; Flinders University, College of Medicine and Public Health ; Flinders Medical Centre, Plastic & Reconstructive Surgery Dean, Nicola; Flinders University, College of Medicine & Public Health; Flinders Medical Centre, Plastic & Reconstructive Surgery Pusic, Andrea; Brigham and Women's Hospital Biomedical Research Institute Cano, Stefan; Modus Outcomes Griffin, Philip; Flinders Medical Centre, Plastic & Reconstructive Surgery Bain, Gregory; Flinders University, College of Medicine and Public Health ; Flinders Medical Centre, Orthopaedic Surgery Lalonde, Donald; Dalhousie University, Division of Plastic Surgery Klassen, Anne; McMaster University, Pediatrics
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Surgery, Health services research, Patient-centred medicine, Qualitative research
Keywords:	patient reported outcome measures, Hand & wrist < ORTHOPAEDIC & TRAUMA SURGERY, psychometrics, QUALITATIVE RESEARCH, PLASTIC & RECONSTRUCTIVE SURGERY

SCHOLARONE™
Manuscripts

1 International multi-phase mixed methods study to develop a cross- 2 cultural patient-reported outcome and experience measure for 3 hand conditions (HAND-Q)

4 Authors

- 5 1. Kyra L Sierakowski MD MS, College of Medicine & Public Health, Flinders
6 University, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia,
7 kyra.sierakowski@flinders.edu.au
- 8 2. Nicola R Dean FRACS (Plas) PhD, Department of Plastic & Reconstructive
9 Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide,
10 SA, Australia, Nicola.dean@sa.gov.au
- 11 3. Andrea L Pusic FRCSC (Plas) MD MSc, Brigham and Women's Hospital, 75
12 Francis S, Boston, MA, USA, 02116, apusic@bwh.harvard.edu
- 13 4. Stefan Cano PhD CPsychol, Modus Outcomes, Suite 210b, Spirella Building,
14 Letchworth Garden City, SG6 4ET, UK, stefan.cano@modusoutcomes.com
- 15 5. Philip A Griffin FRACS (Plas) MBBS, Department of Plastic & Reconstructive
16 Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide,
17 SA, Australia philip.griffin@sa.gov.au
- 18 6. Gregory I Bain FRACS (Ortho) PhD, Department of Orthopaedic Surgery, Flinders
19 Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia,
20 greg@gregbain.com.au
- 21 7. Don Lalonde, FRCSC (Plas) MD, Division of Plastic Surgery, Dalhousie University,
22 Saint John, Canada, drdonlalonde@nb.aibn.com
- 23 8. Anne F Klassen DPhil, McMaster University, 3N27, 1280 Main Street W,
24 Hamilton, ON, Canada, L8N 3Z5, aklass@mcmaster.ca

26 Corresponding Author

27 Kyra L Sierakowski MD MS, College of Medicine & Public Health, Flinders
28 University, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia,
29 kyra.sierakowski@flinders.edu.au +61431306997

31 Keywords

32 Psychometrics, hand, patient-reported outcome measures, qualitative research, plastic
33 and reconstructive surgery

35 Funding Statement

36 This work has received funding from the Australasian Foundation for Plastic Surgery
37 Plastic & Reconstructive Surgery (PRS) Research Scholarship and the Royal
38 Australasian College of Surgeons, Foundation for Surgery Small Project Grant. The
39 authors have no financial interest to declare about the content of this article. The article
40 processing charge was paid from the PRS Research Scholarship.

42 Competing Interests Statement

43 The authors have no competing interests to report.

44 **Word count** = 4099

46 Author Contributions

- 47 • KS prepared the manuscript with the oversight of ND, GB, PG, DL, AK.

- 1
2
3
4
5
6
- 48 • AK, AP and SC developed the methodology utilised for the development of the
49 HAND-Q and guided the content of the protocol.
 - 50 • All authors reviewed the protocol manuscript and gave valuable feedback to the
51 paper.

7
8

52 **Acknowledgements**

53 The authors acknowledge our partners in hand care centres around the world for
54 participating in data collection in various phases of the study.

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Introduction

Patient-reported outcome measures (PROMs) are instruments used to measure outcomes and experiences of healthcare from the patient perspective. The specific methodology used to develop a PROM should be communicated to establish the quality of the instrument. This mixed methods protocol describes the development of a cross-cultural, internationally applicable PROM for hand conditions, the HAND-Q.

Methods and analysis

The multi-phase approach used for this study has been previously successful with the development of other PROMs by our team (e.g., BODY-Q, BREAST-Q, CLEFT-Q, FACE-Q). In Phase I, we established what important concepts matter to patients with hand conditions. A conceptual framework was developed from a systematic review of existing PROMs in the field and an extensive international qualitative study. An 'interpretive description' qualitative approach was used for in-depth interviews. Interviews were audio recorded, transcribed and analysed using a line-by-line method. Item generation was based on the qualitative data. Preliminary HAND-Q scales were created for each theme identified in the conceptual framework. These scales were refined by cognitive debriefing interviews with participants and input from clinical and research experts. Phase II will involve a large heterogeneous international sample of patients completing the HAND-Q field-test scales. The HAND-Q will then be refined using the modern psychometric approach of Rasch Measurement Theory (RMT). This analysis will result in a shortened set of clinically meaningful and scientifically robust HAND-Q scales.

Ethics and dissemination

This study is coordinated at Flinders University (Adelaide, Australia) where it has ethics board approval for Phase I and Phase II. Findings will be published in peer-reviewed journals and presented at local, national and international conferences.

Article Summary

Strengths and Limitations of this study

- The inclusion of a diverse international sample of participants ensures that the HAND-Q will serve the global community.
- Including participants with any hand condition or injury in the field-testing of the HAND-Q (Phase II), will ensure that the resulting instrument is broadly applicable to any form of congenital, traumatic or degenerative hand condition.
- The HAND-Q consists of independently functioning scales which allow for only the concepts of interest to be measured for a specific application. The ability to tailor the choice of scales used reduces patient and administrative burden.
- The comprehensiveness of the HAND-Q provides a range of scales for use in clinical, research and quality assurance applications.
- Use of Rasch Measurement Theory will produce HAND-Q scales that are scientifically sound. Scales will be clinically meaningful; each scale will measure a clinically relevant hierarchy.

Introduction

Increasingly, patient-reported outcome (PRO) data is used in the assessment of treatment effectiveness and healthcare funding(1, 2). It is essential that PRO data is captured with appropriately robust instruments that satisfy best practice guidelines(3, 4). The methodology used to develop and validate a PROM can be complicated; however, it is crucial for researchers and clinicians to understand the approach used to judge the appropriateness of an instrument for their specific application(5, 6).

There are a wide variety of conditions and injuries that affect the hand, and all have the potential to have an enormous impact on an individual's daily function, their livelihood and their quality of life. Conditions include chronic degenerative bony diseases, such as osteoarthritis and rheumatoid arthritis, chronic progressive soft tissue disorders, such as Dupuytren's contracture, and acute injuries of the hand. Hand injuries occur within the workplace and in the home environment and vary from a simple skin laceration to complex tendon and bone damage. This heterogeneity of conditions can make the measurement of treatment outcomes extraordinarily challenging and can limit the clinical applicability of condition-specific PROMs.

There is a substantial burden of illness as a result of hand conditions internationally. Injuries to the hand and wrist are very common, reported to account for 20% of all emergency department visits(7, 8). Non-traumatic conditions of this region are also prevalent; carpal tunnel release was one of the most commonly performed surgical procedures in Ontario, Canada between 2002-2011(9); and Osteoarthritis of the hand has been reported to have a lifetime risk of almost 40%(10). Due to the high health burden of hand conditions, there is a growing awareness of the economic and environmental impacts of medical care, particularly in the surgical sphere where access to the operating theatre is competitive and expensive(1, 11, 12). These factors have led to an increasing amount of hand surgery being performed outside of operating theatres with a fully awake patient and the use of only local anaesthesia(13). This wide-awake approach minimises the economic burden of hand surgery(14, 15) and may provide other benefits to patients by avoiding general anaesthesia(16). It is essential to be able to measure patient satisfaction with their experience of hand surgery and the associated anaesthesia, using an appropriately designed PROM. The HAND-Q will explore the experience of care concepts that have not been addressed by existing PROMs.

There are broadly two approaches taken to establish validity and reliability of PROMs: traditional psychometric methods, which include classical test theory (CTT); and modern psychometric methods, which encompass Rasch Measurement Theory (RMT) and Item Response Theory (IRT). The mathematical basis of these varied techniques is beyond the scope of this protocol paper and is discussed at length elsewhere (17, 18). Instruments developed using the latter approaches have the potential to function consistently regardless of the population that is being measured. This characteristic is due to the psychometric properties of the PROM being that of the instrument itself, not a reflection of the population being measured (18). This approach to measurement allows the same PROM to be used in different patient cohorts with comparable results. Legacy instruments developed using CTT cannot be used in this manner(17). A further

benefit of the RMT and IRT approaches is that measurement is generated as interval format data, which means that scores are consistent in their spacing throughout the measurement metric (17). This format of data enables accurate and meaningful measures of change(19). Interval format data is particularly important in the surgical field, where pre-operative scores are compared with postoperative scores to determine overall improvement and treatment effectiveness(17).

The vast majority of PROMs are developed in a single language and usually within a single country. As hand conditions are prevalent around the world(20-22), this approach means that there is limited content validation within different cultures and economic environments. Many PROMs used in hand surgery have been developed in English and then translated into multiple other languages. The Patient Outcomes of Surgery-Hand/Arm (POS-Hand/Arm)(23) and the Patient Evaluation Measure (PEM)(24) were both developed in the United Kingdom. The Disabilities of the Arm, Shoulder and Hand (DASH)(25) and the Michigan Hand Questionnaire (MHQ)(26) were developed in North America and subsequently translated into many languages(27-30). Although guidelines exist to ensure appropriate cross-cultural adaptation of PROMs(31, 32), a more rigorous approach is to involve people from different countries of varied economic status in the development and validation of a new instrument, rather than performing translation retrospectively(5). A tool developed in a cross-cultural manner could encourage and facilitate multinational research in the field of hand surgery.

The process to develop a cross-cultural PROM that is both scientifically sound and clinically relevant is a complex undertaking. The following protocol details the methodology that will be used for the international study to develop the HAND-Q.

Methods and Analysis

Overview of PROM development

We have engaged patients and clinical experts in all phases of development to create scales that are clinically meaningful. We will use RMT to ensure the HAND-Q scales are scientifically robust. The development of the HAND-Q has followed the international best practice guidelines as determined by the Scientific Advisory Committee of the Medical Outcomes Trust(4), the USA Food and Drug Administration(3) and the International Society for Pharmacoeconomics and Outcomes Research (ISPOR)(33, 34). We have previously published(5) the 3-phase mixed methods approach that we employ to develop a PROM; a summary is reproduced in Figure 1. This approach produces a PROM that satisfies the minimum requirements of reliability and validity as set forth by the International Society for Quality of Life Research (ISOQOL)(35) and the Consensus-based Standards for the Selection of Health Status Measurement Instruments (COSMIN)(6, 36). Our approach to the development of the HAND-Q will be expanded upon in this paper.

Phase I – What should we measure?

The objective of Phase I was to establish the concepts that are important to patients with hand conditions; these concepts would form the basis of the preliminary HAND-Q scales. A systematic literature review was performed to identify the currently available

PROMs relevant to hand conditions and hand surgery(37). This comprehensive literature search was carried out according to PRISMA guidelines. A total of 3039 citations were retrieved, 24 instruments reviewed, including ten regional upper limb, 6-hand/wrist specific and eight hand condition-specific instruments. Review of the developmental methodology of existing PROMs revealed areas of significant weakness(37). The review also found content limitations; no hand-specific instrument was found that covers the full spectrum of outcomes relevant to hand surgery. Based on this review we determined that there is a clinical need for a comprehensive PROM for hand surgery patients.

To develop this instrument, we choose to use a multi-phase mixed methods approach. The design typology is in keeping with that described by Cresswell & Plano's definition of exploratory sequential design, instrument development variant (38). In this approach, the findings from a qualitative study with a small cohort (Phase I) informs the development of a questionnaire. The generalisability of the questionnaire is then tested by its application in a large cohort using a quantitative study (Phase II). The starting point of any PROM development project is a conceptual framework(4), which includes a description of the concepts of interest and the relationship between these concepts within the population of interest. Based on the findings of the systematic review, concepts from existing instruments were mapped along with other new concepts of interest to create a preliminary conceptual framework.

This framework was used as a basis for an interview guide to establish which concepts are meaningful to patients. This document was used to lead the discussions during the qualitative interviews which were conducted using an approach from applied health services research known as Interpretive Description(39, 40). This approach acknowledges that there is pre-existing theoretical and clinical knowledge informing the study. Also, it aims to produce knowledge relevant to the clinical context with the proviso that the individual's understanding of a concept is of the highest importance, regardless of the clinical or theoretical explanation(41).

Participants, setting and recruitment

The Phase I qualitative study recruited participants from centres in Adelaide, Australia and Saint John, New Brunswick, Canada. Recruitment was by direct referral from clinicians or administrative staff. Participants were eligible for inclusion if they had experienced surgery on either or both of their hands over the preceding twelve months, with a minimum of 4 weeks since their most recent hand surgery. Exclusion criteria was the inability to speak English or cognitive problem that prevented participation in a semi-structured interview.

Participants were purposively sampled to include a heterogeneous population with respect to age, gender, hand condition, the setting of surgery (hospital operating theatre vs private rooms), funding for surgery (public vs private) and whether surgery was performed with general anaesthesia, sedation or local anaesthesia. Sampling continued until saturation was achieved, i.e. no further important concepts were identified in three consecutive interviews(42). Based on our previous experience in PROM

1 development, we anticipated that approximately 50 qualitative interviews would be
2 required to reach saturation (43-46).
3

4 *Data collection & analysis*

5
6 Written consent was obtained from all participants before commencing with the
7 individual, semi-structured interviews. Participant age, gender, hand condition and date
8 of surgery was collected. Interviews were audio-recorded and transcribed verbatim
9 with all identifiers removed. The initial interview guide (Table 1) included a list of
10 open-ended questions to encourage discussion. The interviewer probed for new
11 concepts during the interviews. An iterative approach was used, whereby data
12 collection and analysis took place concurrently. This technique resulted in multiple
13 adjustments to the interview guide throughout the study.
14
15
16
17

18 A line-by-line approach was used to code the data. Constant comparison was utilised to
19 identify common concepts of interest between participants(47). A codebook was
20 developed that outlines the codes for evolving domains and themes. The information
21 from these interviews was used to shape and reconfigure the conceptual framework to
22 represent outcomes and experiences relevant to participants.
23
24
25

26 Several strategies have been used to ensure the rigour of this qualitative study. A single
27 interviewer performed all of the qualitative interviews. Iterative techniques were used,
28 which allowed for member-checking to confirm if concepts identified by participants
29 resonated with other participants in subsequent interviews(48). One team member
30 performed the coding of the transcripts, which was then corroborated by a second team
31 member. The members of the study team discussed the data analysis throughout the
32 study and thus used peer-debriefing to provide consistency(48).
33
34
35

36 *Item generation & refinement*

37
38 The evolved conceptual framework was used to identify which scales were generated. A
39 comprehensive item pool was developed from the coding process. Each of the domains
40 identified had multiple themes; each theme was developed into a preliminary scale. This
41 approach ensured that the scales of the HAND-Q cover all of the important concepts of
42 interest identified by patients with hand conditions. The scales were constructed from
43 the items identified in the coding process; participants' language was kept intact as
44 much as possible.
45
46
47
48

49 Incorporating feedback from both patients and clinical experts enhanced the quality of
50 the preliminary HAND-Q scales. Participants who had previously taken part in
51 qualitative interviews, as well as new participants, were invited to give feedback during
52 cognitive debriefing interviews. Input was also sought from a multidisciplinary
53 international team of clinical and research experts who were asked to provide feedback
54 via an online survey using Research Electronic Data Capture (REDCap).
55
56
57

58 After the preliminary scales were constructed, they were then shown to participants
59 during semi-structured cognitive interviews. These interviews were conducted by
60 telephone, audio-recorded, transcribed and analysed. The purpose of these interviews
was to ensure that participants understood the instructions, items and the response

options. The participants were asked to go through each of the scales using the “think aloud” technique, while the interviewer noted any content that caused problems(49, 50). Feedback was also sought on the relevance of the items to the individual. If a phrase was unclear to a participant, it was discussed with them how best to improve it.

Further probing was used to inquire about any new content that may have been overlooked. Cognitive interviews were performed in three rounds to provide the opportunity to make changes to the HAND-Q. The first round included participants from the qualitative interview cohorts in Australia and Canada. Subsequent rounds included new participants from the United States (US) that were recruited using the same inclusion criteria as for the qualitative study. The addition of a US cohort was to ensure content validity within this country.

To make certain that the HAND-Q scales captured all clinically relevant concepts we sought feedback from clinicians who specialise in the management of hand conditions. Clinical experts from the fields of Plastic & Reconstructive Surgery, Orthopaedic Surgery, Hand therapy and Physiotherapy were invited to review the HAND-Q via an online survey. Researchers and academics in the field were also asked to participate. This survey was administered using REDCap, a secure web-based data collection system(51). An international sample was included to ensure that we engaged with professionals working in diverse healthcare systems with different treatment strategies. Both clinical and academic experts were asked to provide feedback on all aspects of each scale (e.g. instructions, response options, items) and to nominate any missing items. This feedback was used to refine the scales further. Participants in the cognitive interviews were able to review the changes made as a result of expert feedback to ensure clarity was maintained.

Translation

To facilitate international field-testing of the HAND-Q, cultural and linguistic validation into multiple languages is currently underway. The field-test scales will be translated into the required languages according to the guidelines set forth by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR)(31) and the World Health Organization (WHO)(32) recommendations for linguistic validations. In summary, this approach requires the forward translation (English to target language) to be performed by two independent bilingual individuals, with back translation (target language to English) to be completed by a third individual. Any discrepancies between independent translations will be resolved before advancing with the next step of the process. The resulting translated, and culturally adapted versions of the HAND-Q will then be shown to a small group of patients who speak the relevant language to ensure that the translation is valid and ready for use.

Phase II – How should we measure the concepts identified in Phase I?

The next phase of developing the HAND-Q involves field-testing the scales. A large heterogeneous sample of people with any hand condition will be asked to complete the HAND-Q. Analysis of the field-test data will allow the scales to be refined to include only the subset of items that are the most effective in measuring the concepts of interest. We use Rasch Measurement Theory (RMT) which dictates that the data must comply with

1 the Rasch measurement model to attain accurate measurement(52). Performing RMT
2 analysis will make it is possible to identify poorly functioning items. Only the best
3 subset of items that map out a clinical hierarchy for each scale will be retained. The
4 details of the psychometric parameters used in RMT analysis are described in detail
5 elsewhere(53). Scales refined using RMT analysis can provide a measurement that is
6 suitable for application at an individual patient level, unlike instruments developed
7 using CTT, which are designed for analysis of groups of individuals(17).
8
9
10

11
12 The international field-test study aims to collect scale responses from a sizeable
13 heterogeneous cohort of patients with hand conditions. These responses from patients
14 will form the basis of scale refinement using RMT. The psychometric properties of the
15 resulting scales can then be established.
16
17

18 *Participants*

19 The international field-test will include participants from multiple English and non-
20 English speaking countries. Field-test sites will be included based on their interest in
21 participating and the feasibility of recruiting an adequate sample size over the
22 anticipated timeframe. We aim to include sites in five countries and with 200
23 participants recruited per country (n=1000). This sample size estimation is based on a
24 minimum requirement of 50 participants in each of four class intervals to allow for
25 analysis by country (differential item functioning or DIF) and will provide a robust
26 sample to examine DIF for other patient characteristics (e.g., gender, hand condition)
27 (54). Participants will include anyone aged 18 years or older with a hand condition that
28 can understand the study and legally provide consent. Participants will be recruited in
29 clinic settings and asked to complete the scales while waiting for their appointment.
30
31
32
33
34

35 *Data collection & analysis*

36 Demographic information will be collected to allow for sub-cohort analysis. Participants
37 will be asked to complete the HAND-Q scales either using a paper booklet or electronic
38 device (smartphone or tablet) using a REDCap survey.
39
40
41

42 If participants complete the scales using a paper booklet, members of the research team
43 will manually enter these data into the REDCap database. Completed data files will be
44 downloaded into IBM SPSS Statistics Version 25. The SPSS file will be formatted suitably
45 and imported into RUMM2030 for RMT analysis. Each of the scales will be analysed
46 separately to determine its psychometric properties according to the following criteria:
47
48

- 49 • The thresholds for item response options must be appropriately ordered.
50 That is, a response of '1' sits lower on a continuum than '2' etc. The
51 hierarchy of items on the scale is then determined, from the 'easiest' to
52 endorse to the 'hardest' to endorse.
- 53 • Several item fit statistics are used to determine if the scale works
54 adequately: (1) log residuals, which reflect the item-person interaction,
55 (2) chi-squared statistic, which indicate the interaction between the item
56 and the concept being measured and (3) item characteristic curves which
57 demonstrate graphically the relationship between groups of responders
58 and the trait of interest(55). Each item will be assessed using these
59 parameters and excluded from the scales if it does not perform well.
60

- The targeting of the scale within the patient population will be reviewed to ensure that the scale can measure the full breadth of the construct within the population of interest.
- The next requirement in the assessment of the scales is to ensure that the items display internal consistency; that is that they are appropriately inter-related. The first step will be testing for unidimensionality; which is checking that all items contributing to the scale are measuring the same underlying construct. The scale will then be evaluated using the Person Separation Index (PSI), which is a similar measure to Cronbach alpha in CTT. The process of analysis will be iterative; scales that are not functioning will be improved by excluding items with inferior performance. This process will be completed when each of the scales shows statistics within the acceptable range.

As RMT produces a scale that behaves in the same manner regardless of the population for which it is used, differences between sub-groups of the population can be found. When one sub-population answers an item differently to another sub-population, that item displays DIF. The inclusion of multiple countries in the international field-test will allow for any differences based on country of origin to be considered. The software RUMM2030 can identify DIF and any items that display DIF can be excluded during the item reduction process, or other adjustments made to account for any differences.

The location of each item on the difficulty spectrum of each scale helps to determine if there are any redundant items; i.e. two or more items that measure the same point on the difficulty spectrum. Items will be reduced if they are identified as poorly functioning by the previously mentioned statistics or if they are redundant. Item reduction will continue until the optimal number of items remain. This number is determined by consideration of the distribution of item locations as well as the clinical requirement for accuracy. Once this process is completed each of the scales will be finalised. A scoring table will be produced for each scale using the RMT analysis.

Following the finalisation of scales, the logits will be used to transform scale scores for each participant from 0 (worst) to 100 (best). Basic relationships between scores and demographic characteristics can then be calculated using independent samples t-tests and analysis of variance (ANOVA) or the equivalent nonparametric tests as appropriate.

There are various components of construct validity, such as structural validity, hypothesis testing and cross-cultural validity. Rasch analysis of unidimensionality addresses the structural validity. Analysis of DIF addresses the cross-cultural validity of the scales. A priori hypothesis testing of known groups is a technique to assess whether the responses to the scales correlate or differ between groups in a way that would be expected(56). In the HAND-Q, we intend to test the following hypotheses:

1. The associations between the scale scores will show the extent to which each scale measures a separate but related construct. *We hypothesise that these intercorrelations would range between $r = 0.30$ and $r = 0.70$ as the scales are being developed to measure distinct concepts but with clear clinical relations.*

2. The associations between the scale scores and the patient characteristics of gender, age and nationality; to determine the extent that a scale may be vulnerable to bias based on these variables. *We predict that these correlations will be low (<0.30).*
3. The association between appearance scores of patients with hand conditions that have a pronounced effect on appearance (Dupuytren's and Rheumatoid Arthritis) and those where the aesthetic impact is minimal (Carpal Tunnel Syndrome). *We hypothesise that the former group would have lower appearance scores than the latter group.*
4. The relationship between patients who describe their hand condition as "severe", with those who describe their condition as "moderate" or "mild" will be explored. *We hypothesise that the scores will be incrementally lower for increasing severity of hand condition.*
5. The comparison of scores between patients who anticipate needing further surgery for their hand condition compared to patients who did not. *We hypothesise that scores will be lower in the group that does not anticipate needing future surgery.*

Phase III: how does the instrument work?

In developing the HAND-Q, we are complying with the COSMIN checklist(57). Many of the required components will be addressed in the preceding phases. The third and final phase includes extended psychometric testing to further establish the reliability, validity and to examine the responsiveness of the HAND-Q. All testing within this phase will utilise the final versions of the HAND-Q scales.

Patient and public involvement

Our patient-oriented approach engages patients and healthcare providers in all stages of our research. Their input is fundamental to the design of the study and development of content for HAND-Q scales. All participants in the initial qualitative interviews will be invited to continue to collaborate in our research by taking part in cognitive interviews where they can provide feedback on our findings and help to refine the final set of scales. We will disseminate updates on the HAND-Q development to the patient community and healthcare providers via local health newsletters with links and information about presentations and publications.

Ethics

This international study is coordinated from Flinders University where the Southern Adelaide Clinical Human Research Ethics Committee has approved Phase I and II. Ethical approval will be obtained before the study commences at a collaborating site. Participant confidentiality will be maintained as a priority. The coordinating centre will collect no identifying participant information; only de-identified data will be used for RMT analysis. Participants will provide written or oral consent before their contribution. As participants will be asked to answer the HAND-Q based on their experiences, it is possible they could potentially experience some distress. Participants will be advised that should this occur, the researcher who recruited them can put them in touch with an appropriately skilled clinician. Participants have and will be assured that any personal information will be kept confidential. In the qualitative phase, all identifying information was removed during the process of transcription. Any personal information required for follow up of individual participants will be kept secure and

1 confidential at each collaborating site, with local rules for data storage followed at each
2 institution.
3

4 **Dissemination**

5 To ensure broad uptake of the HAND-Q, once finalised it will be made available free of
6 charge for non-commercial use. Our team will actively promote the HAND-Q at local,
7 national and international conferences. We will publish findings for the HAND-Q in
8 journals known to be valued and read by our target audiences. This study has no
9 intention to compare outcomes from different centres. Any published work or public
10 presentations resulting from this study will not identify specific centres. Our
11 collaborations with multiple sites internationally are hoped to increase the application
12 of the HAND-Q in the future. The individual scored HAND-Q data from Phase II and
13 Phase III will be returned to the collaborating sites for their use.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 **Figure 1** Flow diagram illustrating the multiphase mixed methods approach to the
2 **development of the HAND-Q.**
3

4 QUAN, quantitative study component; QUAL, qualitative study component.

5 Image reproduced from Wong Riff et al. (5)
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

Table 1 Phase I preliminary interview guide for qualitative interviews.**Experience of Care:**

1. What treatments have you had for your condition?
2. What was good or bad about the treatment?
3. If the participant has had surgery –
 - a. What was your experience of the anaesthetic used? Probe: general anaesthetic, block, local
 - b. Would you have considered having treatment under local anaesthetic? Probe: why, why not
4. Who do you see when you come to the hospital clinic? Probe: receptionist, nurse, doctor, occupational therapist
5. What are the people like who care for you? Probe: friendly, made you feel comfortable, easy to talk to, listened to you
6. What kind of verbal and written information did they give you? Probe: gave enough information, let you ask questions, answered your questions, provided information about recovery

Physical Function:

7. Does your condition create any functional problems? Probe: work, personal care, hobbies
8. What specific things do you have difficulty with due to your hand problem? Probe: getting dressed, cooking, typing, sport
9. Do you experience any symptoms related to your functional problem? Probe: pain, discomfort, embarrassment, mood disturbance

Psychological Wellbeing:

10. How does your hand problem make you feel? Probe: frustrated, angry, upset, worried, stressed
11. How does your hand problem make you feel about yourself? Probe: self esteem, body image, confidence, self-conscious, different from others

Appearance:

12. How would you describe the appearance of your hand/s? Probe: from close up, from far away, symmetry, texture, attractiveness
13. How has your hand appearance changed since your treatment? Probe: scarring, descriptive detail
14. What do you like or dislike about your hand appearance?
15. Is there anything about your hand appearance that you would like to change? Probe: for details
16. Do you ever hide your hands? How do you do this?
17. How important is the appearance of your hands to you?

Other:

18. Is there anything I have not asked you that you think it is important for me to know?
19. Would you like to receive a copy of the transcript from today's discussion?
20. Would you be interested in participating in Cognitive Interview?

References

1. Porter ME. What is value in health care? *New Engl J Med*. 2010;363(26):2477-81.
2. Coronini-Cronberg S, Appleby J, Thompson J. Application of patient-reported outcome measures (PROMs) data to estimate cost-effectiveness of hernia surgery in England. *J R Soc Med*. 2013;106(7):278-87.
3. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. p. 65132-3.
4. Lohr KN. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res*. 2002;11(3):193-205.
5. Wong Riff KW, Tsangaris E, Goodacre T, Forrest CR, Pusic AL, Cano SJ, et al. International multiphase mixed methods study protocol to develop a cross-cultural patient-reported outcome instrument for children and young adults with cleft lip and/or palate (CLEFT-Q). *BMJ open*. 2017;7(1):e015467.
6. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-49.
7. Larsen CF, Mulder S, Johansen AMT, Stam C. The epidemiology of hand injuries in The Netherlands and Denmark. *Eur J Epidemiol*. 2004;19(4):323-7.
8. Angermann P, Lohmann M. Injuries to the hand and wrist. A study of 50,272 injuries. *J Hand Surg Am*. 1993;18(5):642-4.
9. Feinberg AE, Porter J, Saskin R, Rangrej J, Urbach DR. Regional variation in the use of surgery in Ontario. *CMAJ open*. 2015;3(3):E310.
10. Qin J, Barbour KE, Murphy LB, Nelson AE, Schwartz TA, Helmick CG, et al. Lifetime risk of symptomatic hand osteoarthritis: the Johnston County Osteoarthritis Project. *Arthritis & Rheumatology*. 2017;69(6):1204-12.
11. Leblanc MR, Lalonde J, Lalonde DH. A detailed cost and efficiency analysis of performing carpal tunnel surgery in the main operating room versus the ambulatory setting in Canada. *Hand*. 2007;2(4):173-8.
12. Thoma A, Strumas N, Rockwell G, McKnight L. The use of cost-effectiveness analysis in plastic surgery clinical research. *Clin Plast Surg*. 2008;35(2):285-96.
13. Lalonde DH. Wide-awake flexor tendon repair. *Plast Reconstr Surg*. 2009;123(2):623-5.
14. Van Demark RE, Smith VJ, Fiegen A. Lean and Green Hand Surgery. *J Hand Surg Am*. 2018;43(2):179-81.
15. Rhee PC, Fischer MM, Rhee LS, McMillan H, Johnson AE. Cost Savings and Patient Experiences of a Clinic-Based, Wide-Awake Hand Surgery Program at a Military Medical Center: A Critical Analysis of the First 100 Procedures. *The Journal of Hand Surgery*. 2016.
16. Lalonde D. Minimally invasive anesthesia in wide awake hand surgery. *Hand Clin*. 2014;30(1):1-6.
17. Cano SJ, Hobart JC. The problem with health measurement. *Patient Prefer Adherence*. 2011;5:279-90.
18. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health*. 2015;18(1):25-34.
19. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70(12):857.
20. Van Onselen E, Karim R, Hage JJ, Ritt M. Prevalence and distribution of hand fractures. *J Hand Surg Am*. 2003;28(5):491-5.
21. Zhang Y, Niu J, Kelly-Hayes M, Chaisson CE, Aliabadi P, Felson DT. Prevalence of symptomatic hand osteoarthritis and its impact on functional status among the elderly: The Framingham Study. *Am J Epidemiol*. 2002;156(11):1021-7.
22. Chung KC, Spilson SV. The frequency and epidemiology of hand and forearm fractures in the United States. *The Journal of hand surgery*. 2001;26(5):908-15.
23. Cano SJ, Browne JP, Lamping DL, Roberts AH, McGrouther DA, Black NA. The Patient Outcomes of Surgery-Hand/Arm (POS-Hand/Arm): a new patient-based outcome measure. *Journal of Hand Surgery - British Volume*. 2004;29(5):477-85.
24. Macey AC, Burke FD. Outcomes of hand surgery. *J Hand Surg Am*. 1995;20 B(6):841-55.
25. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). [Erratum appears in *Am J Ind Med* 1996 Sep;30(3):372]. *Am J Ind Med*. 1996;29(6):602-8.
26. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *Journal of Hand Surgery - American Volume*. 1998;23(4):575-87.
27. Roh YH, Yang BK, Noh JH, Baek GH, Song CH, Gong HS. Cross-cultural adaptation and validation of the Korean version of the Michigan hand questionnaire. *Journal of Hand Surgery - American Volume*. 2011;36(9):1497-503.

- 1 28. Oksuz C, Akel BS, Oskay D, Leblebicioglu G, Hayran KM. Cross-cultural adaptation, validation, and
2 reliability process of the Michigan Hand Outcomes Questionnaire in a Turkish population. *Journal of Hand
3 Surgery - American Volume*. 2011;36(3):486-92.
- 4 29. Schonemann JO, Larsen K, Hansen TB, Soballe K. Reliability and validity of the Danish version of the
5 disabilities of arm, shoulder, and hand questionnaire in patients with fractured wrists. *J Plast Surg Hand Surg*.
6 2011;45(1):35-9.
- 7 30. Germann G, Harth A, Wind G, Demir E. [Standardisation and validation of the German version 2.0 of
8 the Disability of Arm, Shoulder, Hand (DASH) questionnaire]. *Unfallchirurg*. 2003;106(1):13-9.
- 9 31. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee - Lorenz A, et al. Principles of good
10 practice for the translation and cultural adaptation process for patient - reported outcomes (PRO) measures:
11 report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health*. 2005;8(2):94-104.
- 12 32. Lewis JJ, Leung D, Casper ES, Woodruff J, Hajdu SI, Brennan MF. Multifactorial analysis of long-term
13 follow-up (more than 5 years) of primary extremity sarcoma. *Arch Surg*. 1999;134(2):190-4.
- 14 33. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing
15 and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical
16 product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new
17 PRO instrument. *Value Health*. 2011;14(8):967-77.
- 18 34. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing
19 and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical
20 product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent
21 understanding. *Value Health*. 2011;14(8):978-88.
- 22 35. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends
23 minimum standards for patient-reported outcome measures used in patient-centered outcomes and
24 comparative effectiveness research. *Qual Life Res*. 2013;22:1889-905.
- 25 36. Zywiell MG, Mahomed A, Gandhi R, Perruccio AV, Mahomed NN. Measuring Expectations in
26 Orthopaedic Surgery: A Systematic Review. *Clin Orthop Relat Res*. 2013;471(11):3446-56.
- 27 37. Sierakowski K, Evans Sanchez, KA, Damarell, RA, Dean, NR, Griffin, PA, Bain, GI. Measuring Quality of
28 Life & Patient Satisfaction in Hand Conditions: A systematic review of currently available patient reported
29 outcome instruments. *Australasian Journal of Plastic Surgery*. 2018;(in press).
- 30 38. Cresswell J, Plano Clark V. *Designing and conducting mixed method research*. 2nd Sage. Thousand
31 Oaks, CA. 2011.
- 32 39. Thorne S, Kirkham SR, MacDonald-Emes J. Focus on qualitative methods. Interpretive description: a
33 noncategorical qualitative alternative for developing nursing knowledge. *Res Nurs Health*. 1997;20(2):169-77.
- 34 40. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative
35 alternative for developing nursing knowledge. *Res Nurs Health*. 1997;20.
- 36 41. Claire Welford R, Dip N, Kathy Murphy R, Dymna Casey R. Demystifying nursing research
37 terminology. Part 1. *Nurse Researcher (through 2013)*. 2011;18(4):38.
- 38 42. Sandelowski M. Theoretical Saturation. In: Given LM, editor. *The SAGE Encyclopedia of Qualitative
39 Research Methods*. Thousand Oaks: SAGE Publications, Inc.; 2008.
- 40 43. Wong Riff KW, Tsangaris E, Goodacre TE, Forrest CR, Lawson J, Pusic AL, et al. What matters to
41 patients with cleft lip and/or palate: an international qualitative study informing the development of the
42 CLEFT-Q. *The Cleft Palate-Craniofacial Journal*. 2018;55(3):442-50.
- 43 44. Pusic AL, Klassen AF, Scott AM, Klok JA, Cordeiro PG, Cano SJ. Development of a new patient-reported
44 outcome measure for breast surgery: the BREAST-Q. *Plast Reconstr Surg*. 2009;124(2):345-53.
- 45 45. Klassen AF, Cano SJ, Alderman A, Soldin M, Thoma A, Robson S, et al. The BODY-Q: a patient-reported
46 outcome instrument for weight loss and body contouring treatments. *Plastic and Reconstructive Surgery
47 Global Open*. 2016;4(4).
- 48 46. Klassen AF, Cano SJ, Schwitzer JA, Scott AM, Pusic AL. FACE-Q scales for health-related quality of life,
49 early life impact, satisfaction with outcomes, and decision to have treatment: development and validation.
50 *Plast Reconstr Surg*. 2015;135(2):375-86.
- 51 47. Pope C, Ziebland S, Mays N. Analysing qualitative data. *Br Med J*. 2000;320(7227):114.
- 52 48. Cohen DJ, Crabtree BF. Evaluative criteria for qualitative research in health care: controversies and
53 recommendations. *The Annals of Family Medicine*. 2008;6(4):331-9.
- 54 49. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res*.
55 2003;12(3):229-38.
- 56 50. Van Someren M, Barnard Y, Sandberg J. *The think aloud method: a practical approach to modelling
57 cognitive*. 1994.
- 58 51. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture
59 (REDCap)—a metadata-driven methodology and workflow process for providing translational research
60 informatics support. *J Biomed Inform*. 2009;42(2):377-81.
- 61 52. G R. *Studies in mathematical psychology: 1*. In: Institut. Dp, editor. 1960.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
53. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods: Prepress Projects; 2009.
54. Linacre JM. Sample Size and Item Calibration (or Person Measure) Stability. Rasch Measurement Transactions. 1994;7.
55. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). Br J Clin Psychol. 2007;46(1):1-18.
56. Strauss ME, Smith GT. Construct validity: Advances in theory and methodology. Annu Rev Clin Psychol. 2009;5:1-25.
57. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010;10(1):22.

For peer review only

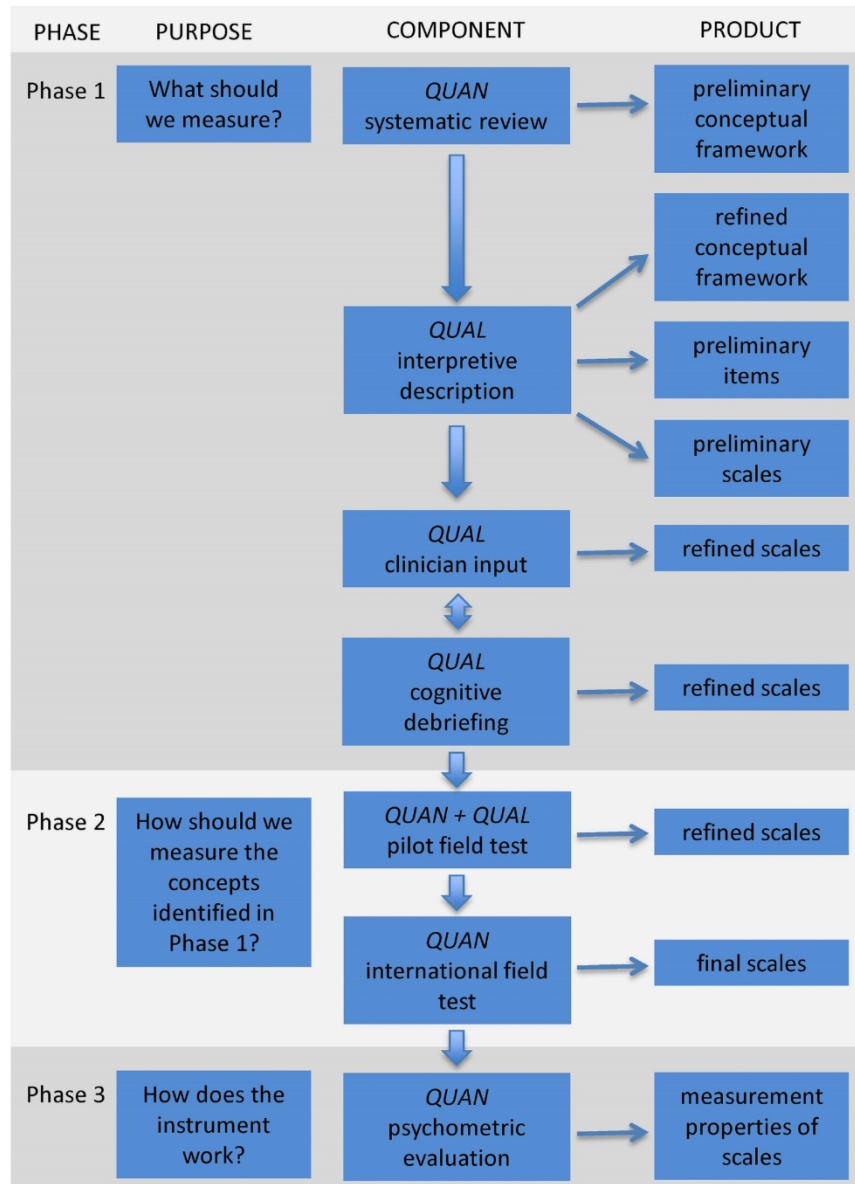


Figure 1 Flow diagram illustrating the multiphase mixed methods approach to the development of the HAND-Q.
 QUAN, quantitative study component; QUAL, qualitative study component.
 Image reproduced from Wong Riff et al.(5)

175x241mm (300 x 300 DPI)

BMJ Open

International multi-phase mixed methods study protocol to develop a cross-cultural patient-reported outcome and experience measure for hand conditions (HAND-Q)

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2018-025822.R2
Article Type:	Protocol
Date Submitted by the Author:	17-Jan-2019
Complete List of Authors:	Sierakowski, Kyra; Flinders University, College of Medicine and Public Health ; Flinders Medical Centre, Plastic & Reconstructive Surgery Dean, Nicola; Flinders University, College of Medicine & Public Health; Flinders Medical Centre, Plastic & Reconstructive Surgery Pusic, Andrea; Brigham and Women's Hospital Biomedical Research Institute Cano, Stefan; Modus Outcomes Griffin, Philip; Flinders Medical Centre, Plastic & Reconstructive Surgery Bain, Gregory; Flinders University, College of Medicine and Public Health ; Flinders Medical Centre, Orthopaedic Surgery Klassen, Anne; McMaster University, Pediatrics Lalonde, Donald; Dalhousie University, Division of Plastic Surgery
Primary Subject Heading:	Research methods
Secondary Subject Heading:	Surgery, Health services research, Patient-centred medicine, Qualitative research
Keywords:	patient reported outcome measures, Hand & wrist < ORTHOPAEDIC & TRAUMA SURGERY, psychometrics, QUALITATIVE RESEARCH, PLASTIC & RECONSTRUCTIVE SURGERY

SCHOLARONE™
Manuscripts

International multi-phase mixed methods study protocol to develop a cross-cultural patient-reported outcome and experience measure for hand conditions (HAND-Q)

Authors

1. Kyra L Sierakowski MD MS, College of Medicine & Public Health, Flinders University, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, kyra.sierakowski@flinders.edu.au
2. Nicola R Dean FRACS (Plas) PhD, Department of Plastic & Reconstructive Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, Nicola.dean@sa.gov.au
3. Andrea L Pusic FRCSC (Plas) MD MSc, Brigham and Women's Hospital, 75 Francis S, Boston, MA, USA, 02116, apusic@bwh.harvard.edu
4. Stefan Cano PhD CPsychol, Modus Outcomes, Suite 210b, Spirella Building, Letchworth Garden City, SG6 4ET, UK, stefan.cano@modusoutcomes.com
5. Philip A Griffin FRACS (Plas) MBBS, Department of Plastic & Reconstructive Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia philip.griffin@sa.gov.au
6. Gregory I Bain FRACS (Ortho) PhD, Department of Orthopaedic Surgery, Flinders Medical Centre, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, greg@gregbain.com.au
7. Anne F Klassen DPhil, McMaster University, 3N27, 1280 Main Street W, Hamilton, ON, Canada, L8N 3Z5, aklass@mcmaster.ca
8. Don Lalonde, FRCSC (Plas) MD, Division of Plastic Surgery, Dalhousie University, Saint John, Canada, drdonlalonde@nb.aibn.com

Corresponding Author

Kyra L Sierakowski MD MS, College of Medicine & Public Health, Flinders University, Flinders Drive, Bedford Park 5042, Adelaide, SA, Australia, kyra.sierakowski@flinders.edu.au +61431306997

Keywords

Psychometrics, hand, patient reported outcome measures, qualitative research, plastic and reconstructive surgery

Funding Statement

This work has received funding from the Australasian Foundation for Plastic Surgery Plastic & Reconstructive Surgery (PRS) Research Scholarship and the Royal Australasian College of Surgeons, Foundation for Surgery Small Project Grant. The authors have no financial interest to declare in relation to the content of this article. The article processing charge was paid from the PRS Research Scholarship.

Competing Interests Statement

The authors have no competing interests to report.

Word count = 3807

Author Contributions

- KS prepared the manuscript with the oversight of ND, GB, PG, DL, AK.

- AK, AP and SC developed the methodology utilised for the development of the HAND-Q and guided the content of the protocol.
- All authors reviewed the protocol manuscript and gave valuable feedback to the manuscript.

Acknowledgements

The authors acknowledge our partners in hand care centres around the world for participating in data collection in various phases of the study.

For peer review only

Abstract

Introduction

Patient reported outcome measures (PROMs) are instruments used to measure outcomes and experiences of healthcare from the patient perspective. The specific methodology used to develop a PROM should be communicated to establish the quality of the instrument. This mixed methods protocol describes the development of a cross-cultural, internationally applicable PROM for hand conditions, the HAND-Q.

Methods and analysis

The multi-phase approach used for this study has been previously used with the development of other PROMs by our team (e.g., BODY-Q, BREAST-Q, CLEFT-Q, FACE-Q). In Phase I, we establish what important concepts matter to patients with hand conditions. A conceptual framework is developed from a systematic review of existing PROMs in the field and an extensive international qualitative study. Interpretive description is the qualitative approach used. Item generation is based on the qualitative data. The preliminary scales will be created for each theme identified in the conceptual framework. These scales will be refined by cognitive debriefing interviews with participants and expert input. Phase II involves a large international sample of patients with varied hand conditions completing the field-test version of the HAND-Q. The scales will be refined using the modern psychometric approach of Rasch Measurement Theory (RMT). Analysis will result in a shortened set of clinically meaningful and scientifically robust HAND-Q scales.

Ethics and dissemination

This study is coordinated at Flinders University (Adelaide, Australia) where it has ethics board approval for Phase I and Phase II. Findings will be published in peer-reviewed journals and presented at local, national and international conferences.

Article Summary

Strengths and Limitations of this study

- The inclusion of a diverse international sample of participants in the field-test ensures that the HAND-Q will serve the global community.
- Qualitative interviews have only been performed with English speaking patients from Australia, the United States and Canada.
- Including the full spectrum of hand conditions and injuries with minimal participant exclusion criteria will ensure that the HAND-Q is broadly applicable to any form of congenital, traumatic or degenerative hand condition.
- Independently functioning scales will allow tailoring of scales to the patient, study or clinical setting, which will reduce patient and administrative burden.
- Use of a modern psychometric approach will produce HAND-Q scales that are both scientifically sound and clinically meaningful.

Introduction

Patient-reported outcome (PRO) data is increasingly collected and used in the assessment of treatment effectiveness and healthcare funding(1, 2). It is important that PRO data is captured with appropriately robust instruments that satisfy best practice guidelines(3, 4). The methodology used to develop and validate a PROM can be complex; however, it is important for researchers and clinicians to understand the approach used in order to judge the appropriateness of an instrument for their specific application(5, 6).

There are a wide variety of conditions and injuries that affect the hand, and each has the potential to have an enormous impact on an individual's daily function, their livelihood and their quality of life. Conditions include chronic degenerative bony diseases, such as osteoarthritis and rheumatoid arthritis, chronic progressive soft tissue disorders, such as Dupuytren's contracture, and acute injuries of the hand. Hand injuries occur commonly within the workplace and in the home environment and can range from a simple skin laceration to complex tendon and bone damage. This heterogeneity of conditions and causes can make the measurement of treatment outcomes extremely challenging and can limit the clinical applicability of condition-specific PROMs.

There is growing awareness of the economic and environmental impacts of medical care, particularly in the surgical sphere where access to the operating theatre is competitive and expensive(1, 7, 8). These forces have led to an increasing amount of hand surgery performed outside of operating theatres with a fully awake patient and the use of only local anaesthetic(9). This wide-awake approach minimises the economic burden of hand surgery(10, 11). It is important to be able to measure patient satisfaction with their experience of hand surgery anaesthesia using an appropriately designed PROM to allow for comparisons to be made of the different anaesthetics approaches to hand surgery. The HAND-Q will explore experience of care concepts that have not been addressed by existing PROMs.

PROMs developed using a modern psychometric approach have the potential to function consistently regardless of the population that is being measured. This is possible because the psychometric properties of the PROM are that of the instrument itself, not a reflection of the population that is being tested. This approach to measurement allows the same PROM to be used in different populations with results comparable on the same metric. Legacy instruments developed using classical test theory methodology cannot be used in this manner. A further benefit of the modern psychometric approach is that measurement is given in interval format, which permits accurate and meaningful measurement of change, which is particularly important in the surgical field.

The vast majority of PROMs are developed in a single language and usually within a single country. As hand conditions are prevalent around the world, this approach means that there is limited content validation within different cultures and economic environments. Many PROMs used in hand surgery have been developed in English and then translated into multiple other languages. Outcome of Surgery-Hand/Arm (POS-

1 Hand/Arm)(12) and the Patient Evaluation Measure (PEM)(13) were both developed in
2 the United Kingdom. The Disabilities of the Arm, Shoulder and Hand (DASH)(14) and
3 the Michigan Hand Questionnaire (MHQ)(15) were developed in North America and
4 subsequently translated into many languages(16-19). Although guidelines exist to
5 ensure appropriate cross-cultural adaption of PROMs(20, 21), a more rigorous
6 approach is to involve people from different countries of varied economic status in the
7 development and validation of a new instrument, rather than performing translation
8 retrospectively(5). An instrument developed in a cross-cultural manner could
9 encourage and facilitate multinational research in the field of hand surgery.
10
11
12
13
14

15 The process to develop a cross-cultural PROM that is both scientifically sound and
16 clinically relevant is a complex undertaking. The following protocol details the
17 methodology that will be used for the international study to develop the HAND-Q.
18
19

20 **Methods and Analysis**

21 **Overview of PROM development**

22 We use a modern psychometric approach and engage patients and experts in all phases
23 of the development to create clinically meaningful and scientifically robust PROM
24 scales. The development of the HAND-Q follows the international best practice
25 guidelines as determined by the Scientific Advisory Committee of the Medical Outcomes
26 Trust(4), the USA Food and Drug Administration(3) and the International Society for
27 Pharmacoeconomics and Outcomes Research (ISPOR)(22, 23). We have previously
28 published(5) the 3-phase mixed methods approach that we employ to develop a PROM,
29 a summary is reproduced in Figure 1. This approach establishes a PROM that satisfies
30 the minimum requirements of reliability and validity as set forth by the International
31 Society for Quality of Life Research (ISOQOL)(24) and the Consensus-based Standards
32 for the Selection of Health Status Measurement Instruments (COSMIN)(6, 25). The
33 application of this 3-phase approach to the development of the HAND-Q will be
34 expanded upon in this paper.
35
36
37
38
39
40
41

42 **Phase I – What should we measure?**

43 The objective of Phase I is to develop preliminary scales founded on excellent content
44 validity. A systematic literature review was performed to identify the currently
45 available PROMs relevant to hand conditions and hand surgery(26). This
46 comprehensive literature search was carried out according to PRISMA guidelines with a
47 total of 3039 citations retrieved. A total of 24 instruments were reviewed; 10 regional
48 upper limb, 6-hand/wrist specific and 8 hand condition-specific instruments. Review of
49 the developmental methodology of those existing instruments revealed areas of
50 significant weakness(26). The review also found profound content limitations with no
51 hand-specific instrument found that covers the full spectrum of outcomes relevant to
52 hand surgery. Based on this review it was determined that there was need for a
53 comprehensive PROM for hand surgery patients and we proceeded with this study.
54 Concepts from other PROMs were used to form a preliminary interview guide.
55
56
57
58
59
60

Conceptual Framework

The starting point in any Phase I PROM development project is the development of a conceptual framework(4), which includes a description of the concepts of interest and the relationship between these concepts within the population that the instrument is designed for. Based on the findings of the systematic review, concepts from existing instruments are mapped along with other concepts of interest to create a preliminary conceptual framework.

Qualitative Study

The preliminary conceptual framework will be used as a basis for an interview guide to establish which of the concepts are meaningful to patients. We use an approach from applied health services research known as Interpretive Description(27, 28). This approach acknowledges that there is pre-existing theoretical and clinical knowledge informing the study. In addition, this approach aims to produce knowledge relevant to the clinical context with the proviso that the individual's understanding of a concept is of the greatest importance, regardless of the clinical or theoretical explanation(29).

Participants, setting and recruitment

Phase 1 qualitative study will recruit participants from centres in Adelaide, Australia and Saint John, New Brunswick, Canada. Recruitment will be by direct referral from clinicians or administrative staff. Participants will be eligible for inclusion if they had experienced surgery on either or both of their hands over the preceding twelve months, with a minimum of 4 weeks since their surgery. Exclusion criteria for this phase of the study will be the inability to speak English or cognitive delay that prevents participation in a semi-structured interview.

Sampling

Participants will be purposively sampled to include a heterogeneous population with respect to age, gender, hand condition, setting of surgery (hospital operating theatre vs private rooms), funding for surgery (public vs private) and whether surgery was performed with general anaesthesia, sedation or local anaesthetic. Sampling will continue until saturation is achieved, i.e. no further important concepts are identified in three consecutive interviews(30).

Data collection

Written consent will be obtained from all participants before commencing with individual, semi-structured interviews. Participant age, gender, hand condition and date of surgery will be collected. Interviews will be audio-recorded and transcribed verbatim with all identifiers removed. The interview guide (Table 1) used is based on concepts identified from the literature review [26]; this guide includes a list of open-ended questions to encourage discussion. The interviewer will probe for new concepts during the interviews. An iterative approach will be used, whereby data collection and analysis will take place concurrently in order to ask about newly identified relevant concepts.

Data analysis

A line-by-line approach will then used to code the data, with a process of constant comparison utilised to identify common concepts of interest(31). A codebook will be developed that outlines the codes for evolving domains and themes. The information

1 from these interviews will be used to shape and reconfigure the conceptual framework
2 to represent outcomes and experiences important to participants.
3

4 *Rigor*

5 Several strategies have been put in place to ensure rigor of this qualitative study. A
6 single interviewer will perform all of the qualitative interviews. One team member will
7 perform all of the coding of the transcripts, which will then corroborated by a second
8 team member. The iterative process allows for member-checking to confirm if concepts
9 are deemed to be valuable to subsequent participants(32). The members of the study
10 team will discuss data analysis and this peer-debriefing will provide consistency(32).
11
12
13
14

15 *Item generation*

16 The evolved conceptual framework will be used to identify which scales should be
17 generated. A comprehensive list of potential items will be generated from the coding
18 process. Each of the domains identified will have multiple themes; each important
19 theme will be developed into a preliminary scale. This approach will ensure that the
20 suite of scales that comprise the HAND-Q will cover all of the important concepts of
21 interest identified by patients with hand conditions. The scales are constructed from the
22 items identified in the coding process; participants' language is kept intact as much as
23 possible.
24
25
26
27

28 *Refining the preliminary scales*

29 Incorporating feedback from both patients and experts in hand conditions will enhance
30 the quality of the preliminary HAND-Q scales. Participants who have previously
31 participated in qualitative interviews as well as new participants will be invited to
32 participate in cognitive debriefing interviews. Input will also be sought from a
33 multidisciplinary international pool of experts who will be invited to provide feedback
34 via an online survey using Research Electronic Data Capture (REDCap).
35
36
37
38

39 *Cognitive debriefing interviews*

40 After the preliminary scales are constructed, they will be shown to participants with
41 hand conditions in the form of semi-structured cognitive interviews. These interviews
42 will be conducted by telephone, audio-recorded, transcribed and analysed. The purpose
43 of these interviews is to ensure that participants have understood the instructions,
44 items and the response options. The participant will be asked to go through each of the
45 scales using the "think aloud" technique, while the interviewer notes the specific
46 content that causes problems(33, 34). Feedback will be sought on the relevance of the
47 items to the individual. If an item is unclear to a participant, they will be asked for
48 suggestions on how we can improve item wording. Further probing will be used to
49 enquire about any new content that may have been overlooked. Cognitive interviews
50 will be performed in three rounds to provide the opportunity to make changes between
51 rounds. The first round will include participants from the qualitative interviews from
52 both the Australian and Canadian cohorts. Subsequent rounds will also include new
53 participants from the USA to ensure content validity for patients in that country.
54
55
56
57
58
59

60 *Expert clinician input*

It is useful to seek feedback from clinicians who specialise in the management of hand
conditions to ensure that the scales capture all clinically relevant concepts. Experts in

1 the management of patients with hand conditions from the fields of Plastic &
2 Reconstructive Surgery, Orthopaedic Surgery, Hand therapy and Physiotherapy will be
3 invited to review the HAND-Q in the form of an online survey. We will administer this
4 survey using REDCap, which is a secure web-based data collection system(35).
5 Researchers and academics in the field will also be invited to participate. An
6 international sample of experts will be included to ensure that we engage with
7 professionals working in diverse healthcare systems with varied treatment strategies.
8 Experts will be asked to provide feedback on all aspects of each scale (instructions,
9 response options, items) and to nominate missing items. Feedback from experts will be
10 used to refine the scales. Expert input will be obtained before the final round of
11 participant cognitive interviews in order to show any changes made from expert input
12 to participants.

18 **Translation**

19 To facilitate an international field-test, cultural and linguistic validation of the HAND-Q
20 into multiple languages will take place. The field-test scales will be translated into the
21 required field-test languages according to the guidelines set forth by the International
22 Society for Pharmacoeconomics and Outcomes Research (ISPOR)(20) and the World
23 Health Organization (WHO)(21) recommendations for linguistic validations. In
24 summary, this approach requires the forward translation (English to target language) to
25 be performed by two independent bilingual individuals, with back translation (target
26 language to English) to be performed by a third individual. Any discrepancies are
27 resolved at each step and the resultant version is then shown to a small group of
28 patients to ensure that the translation is valid and ready for use.

34 **Phase II – How should we measure the concepts identified in Phase I?**

35 A large heterogeneous sample of people aged 18 years and older with any type of hand
36 condition will be recruited. Analysis of the field-test data allows the scales to be refined
37 to include the subset of items that are the most effective in measuring the concepts of
38 interest. We will follow the modern psychometric method of Rasch Measurement
39 Theory (RMT)(36). This approach dictates that in order to achieve accurate
40 measurement the data must comply with the Rasch measurement model(37).
41 Performing RMT analysis will make it is possible to identify poorly functioning items
42 and to only retain the best subset of items that together map out a clinical hierarchy for
43 each scale. The details of the psychometric parameters used in RMT analysis are
44 described in detail elsewhere(38). Scales refined using RMT analysis can provide
45 measurement that is suitable for application at an individual patient level, unlike
46 instruments developed using classical test theory, which are designed for analysis of
47 groups of individuals. RMT produces interval level measurement where the intervals
48 between adjacent scores are equal throughout the breadth of the possible scores.
49 Subsequently measuring change over time (such as comparing pre-operative and post-
50 operative scores which is a common PROM application in the surgical field) with RMT
51 designed scales is mathematically sound(39). Another benefit of scales derived using
52 the RMT approach is that the scales function consistently, independent of the
53 population that is being studied, which allows comparison of scores between different
54 populations.

International field-test and RMT analysis

The aim of the international field-test study is to collect scale responses from a large cohort of patients with hand conditions to allow the refinement of the scales using RMT and examine the psychometric properties of the resulting scales.

Study participants

The international field-test will include participants from multiple English and non-English speaking countries. Field-test sites will be included based on their interest in participating and the feasibility of recruiting an adequate sample size over the anticipated timeframe. We aim to include sites in five countries and with 200 participants recruited per country (n=1000). Participants will include anyone aged 18 years or older with a hand condition that is able to comprehend the study and legally provide consent. Participants will be recruited in clinic settings and asked to complete the scales while waiting for their appointment.

Data collection

Demographic information will be collected to allow for sub-cohort analysis. Participants will be asked to complete the HAND-Q scales either using a paper booklet or electronic device (smartphone or tablet) using a REDCap survey.

Data analysis

If participants complete the scales using a paper booklet, members of the research team will enter these data into the REDCap database. Completed data files will be downloaded into IBM SPSS Statistics Version 25. The SPSS file will be formatted suitably and imported into RUMM2030 for RMT analysis. Each of the scales will be analysed separately to determine its psychometric properties according to the following criteria. The thresholds for item response options must be appropriately ordered. This means that a response of '1' sits lower on a continuum than '2' etc. The hierarchy of items on the scale is then determined, from the 'easiest' to endorse to the 'hardest' to endorse. Several item fit statistics are used to determine if the scale works adequately; (1) log residuals, which reflect the item-person interaction, (2) chi squared statistic, which reflect the interaction between the item and the concept being measured and (3) item characteristic curves which demonstrate graphically the relationship between groups of responders and the trait of interest(40). Each item will be assessed using these parameters and excluded from the scales if they do not perform well. Lastly the targeting of the scale within the patient population is reviewed to ensure that the scale is able to measure the full breadth of the construct within the population of interest. The next requirement in the assessment of the scales is to ensure that the items display internal consistency; that is that they are appropriately inter-related. The first step is testing for unidimensionality; which is checking if all of the items that contribute to the scale are measuring the same underlying construct. The scale is then evaluated using the Person Separation Index, which is a similar measure to Cronbach alpha in classical test theory. This process of analysis is iterative; scales that are not functioning optimally are improved by excluding items with poor performance. This process is completed when each of the scales show the aforementioned statistics within the acceptable range.

Differential item functioning

As RMT produces a scale that behaves in the same manner regardless of the population for which it is used, differences between sub-groups of the population can be identified. When one sub-population answers an item differently to another sub-population, that item is said to display differential item functioning (DIF). Inclusion of multiple countries in the international field-test will allow for any differences based on country of origin to be considered. The software RUMM2030 can identify DIF and any items that display DIF can be excluded during the item reduction process, or other adjustments made to account for any differences.

Item reduction

The location of each item on the difficulty spectrum of each scale helps to determine if there is any redundant items; i.e. two or more items that measure the same point on the difficulty spectrum. Items will be reduced if they are identified as poorly functioning by the previously mentioned statistics or if they are redundant. Items will be reduced until the optimal number of items remains. This approach will be determined by a combination of the distribution of item locations as well as the clinical requirement for the degree of accuracy. Once this process is complete each of the scales is finalised. A scoring table is then produced for each scale using the RMT analysis; the cumulative score of a scale is more complex than the simple addition of individual items.

Construct validity

Following the finalisation of scales the logits will be used to transform scale scores for each participant from 0 (worst) to 100 (best). Basic relationships between scores and demographic characteristics can then be calculated using independent samples t-tests and analysis of variance (ANOVA) or the equivalent nonparametric tests as appropriate in SPSS.

There are various components of construct validity, such as structural validity, hypothesis testing and cross-cultural validity. RMT analysis of unidimensionality addresses the structural validity. Analysis for DIF addresses the cross-cultural validity of the scales. A priori hypothesis testing of known groups is a technique to assess whether the responses to the scales correlate or differ between groups in a way that would be expected(41). In the HAND-Q, we intend to test the following hypotheses using ANOVA in SPSS:

1. Those patients with hand conditions that affect the appearance of the hand (such as rheumatoid arthritis and Dupuytren's contracture) will be more dissatisfied with their hand appearance than those with carpal tunnel syndrome (which rarely has a significant affect on hand appearance).
2. Those patients requiring further intervention for their hand condition would have lower quality of life scores than those who do not require further intervention.
3. Those patients who describe their hand condition as "severe" will have lower quality of life scores than those who describe their conditions as "mild".

Phase III: how does the instrument work?

In developing the HAND-Q we will comply with the COSMIN checklist(42). Many of the required components will be addressed in preceding phases. The third and final phase includes further psychometric testing to establish the reliability, validity and responsiveness of the HAND-Q. All testing within this phase will utilise the final versions of the HAND-Q scales.

Patient and public involvement

Our patient-oriented approach engages patients and healthcare providers in all stages of our research as experts and research team members. Their input is fundamental to the design of the study and development of content for HAND-Q scales. All participants in the initial qualitative interviews will be invited to continue to collaborate in our study by taking part in cognitive interviews where they can provide feedback on our findings and help to refine the final set of scales. We will disseminate updates on the HAND-Q development to the patient community and healthcare providers via local health newsletters with links and information about presentations and publications.

Ethics

This international study is coordinated from Flinders University where the Southern Adelaide Clinical Human Research Ethics Committee has approved Phase I and II. Ethics board approval will be obtained prior to the study commencing at a collaborating site. Participants will provide written or oral consent prior to participating. As participants will be asked to discuss their experiences, they could potentially experience some distress. Participants will be advised that should this occur that the interviewer can put them in touch with an appropriate skilled clinician. Participants will be assured that any personal information will be kept confidential. In the qualitative phase all identifying information will be removed during the process of transcription. Any personal information required for follow up of individual participants will be kept secure and confidential following institution rules for data storage.

Dissemination

To ensure wide uptake of the HAND-Q, once developed it will be made available free of charge for non-commercial use. Our team will actively promote the HAND-Q at local, national and international conferences. We will publish findings for the HAND-Q in journals known to be valued and read by our target audiences. This study has no intention to directly compare outcomes from different centres. Any published work or public presentations resulting from this study will not identify specific centres. Collaborations with multiple sites internationally is hoped to increase the application of the HAND-Q in the future. The individual scored HAND-Q data from Phase II and Phase III will be returned to the providing sites for their own use.

1 **Figure 1** Flow diagram illustrating the multiphase mixed methods approach to the
2 **development of the HAND-Q.**
3

4 QUAN, quantitative study component; QUAL, qualitative study component.

5 Image reproduced from Wong Riff et al.(5)
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 Interview guide for qualitative interviews to be performed in Phase 1.

Experience of Care:

1. What treatments have you had for your condition?
2. What was good or bad about the treatment?
3. If the participant has had surgery –
 - a. What was your experience of the anaesthetic used? Probe: general anaesthetic, block, local
 - b. Would you have considered having treatment under local anaesthetic? Probe: why, why not
4. Who do you see when you come to the hospital clinic? Probe: receptionist, nurse, doctor, occupational therapist
5. What are the people like who care for you? Probe: friendly, made you feel comfortable, easy to talk to, listened to you
6. What kind of verbal and written information did they give you? Probe: gave enough information, let you ask questions, answered your questions, provided information about recovery

Physical Function:

7. Does your condition create any functional problems? Probe: work, personal care, hobbies
8. What specific things do you have difficulty with due to your hand problem? Probe: getting dressed, cooking, typing, sport
9. Do you experience any symptoms related to your functional problem? Probe: pain, discomfort, embarrassment, mood disturbance

Psychological Wellbeing:

10. How does your hand problem make you feel? Probe: frustrated, angry, upset, worried, stressed
11. How does your hand problem make you feel about yourself? Probe: self esteem, body image, confidence, self-conscious, different from others

Appearance:

12. How would you describe the appearance of your hand/s? Probe: from close up, from far away, symmetry, texture, attractiveness
13. How has your hand appearance changed since your treatment? Probe: scarring, descriptive detail
14. What do you like or dislike about your hand appearance?
15. Is there anything about your hand appearance that you would like to change? Probe: for details
16. Do you ever hide your hands? How do you do this?
17. How important is the appearance of your hands to you?

Other:

18. Is there anything I have not asked you that you think it is important for me to know?
19. Would you like to receive a copy of the transcript from today's discussion?
20. Would you be interested in participating in Cognitive Interview?

References

1. Porter ME. What is value in health care? *New England Journal of Medicine*. 2010;363(26):2477-81.
2. Coronini-Cronberg S, Appleby J, Thompson J. Application of patient-reported outcome measures (PROMs) data to estimate cost-effectiveness of hernia surgery in England. *Journal of the Royal Society of Medicine*. 2013;106(7):278-87.
3. U.S. Department of Health and Human Services Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009. p. 65132-3.
4. Lohr KN. Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of Life Research*. 2002;11(3):193-205.
5. Wong Riff KW, Tsangaris E, Goodacre T, Forrest CR, Pusic AL, Cano SJ, et al. International multiphase mixed methods study protocol to develop a cross-cultural patient-reported outcome instrument for children and young adults with cleft lip and/or palate (CLEFT-Q). *BMJ open*. 2017;7(1):e015467.
6. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*. 2010;19(4):539-49.
7. Leblanc MR, Lalonde J, Lalonde DH. A detailed cost and efficiency analysis of performing carpal tunnel surgery in the main operating room versus the ambulatory setting in Canada. *Hand*. 2007;2(4):173-8.
8. Thoma A, Strumas N, Rockwell G, McKnight L. The use of cost-effectiveness analysis in plastic surgery clinical research. *Clinics in plastic surgery*. 2008;35(2):285-96.
9. Lalonde DH. Wide-awake flexor tendon repair. *Plast Reconstr Surg*. 2009;123(2):623-5.
10. Van Demark RE, Smith VJ, Fiegen A. Lean and Green Hand Surgery. *Journal of Hand Surgery*. 2018;43(2):179-81.
11. Rhee PC, Fischer MM, Rhee LS, McMillan H, Johnson AE. Cost Savings and Patient Experiences of a Clinic-Based, Wide-Awake Hand Surgery Program at a Military Medical Center: A Critical Analysis of the First 100 Procedures. *The Journal of Hand Surgery*. 2016.
12. Cano SJ, Browne JP, Lamping DL, Roberts AH, McGrouther DA, Black NA. The Patient Outcomes of Surgery-Hand/Arm (POS-Hand/Arm): a new patient-based outcome measure. *Journal of Hand Surgery - British Volume*. 2004;29(5):477-85.
13. Macey AC, Burke FD. Outcomes of hand surgery. *Journal of Hand Surgery*. 1995;20 B(6):841-55.
14. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. *The Upper Extremity Collaborative Group (UECG)*. [Erratum appears in *Am J Ind Med* 1996 Sep;30(3):372]. *American Journal of Industrial Medicine*. 1996;29(6):602-8.
15. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *Journal of Hand Surgery - American Volume*. 1998;23(4):575-87.
16. Roh YH, Yang BK, Noh JH, Baek GH, Song CH, Gong HS. Cross-cultural adaptation and validation of the Korean version of the Michigan hand questionnaire. *Journal of Hand Surgery - American Volume*. 2011;36(9):1497-503.
17. Oksuz C, Akel BS, Oskay D, Leblebicioglu G, Hayran KM. Cross-cultural adaptation, validation, and reliability process of the Michigan Hand Outcomes Questionnaire in a Turkish population. *Journal of Hand Surgery - American Volume*. 2011;36(3):486-92.
18. Schonemann JO, Larsen K, Hansen TB, Soballe K. Reliability and validity of the Danish version of the disabilities of arm, shoulder, and hand questionnaire in patients with fractured wrists. *Journal of Plastic Surgery and Hand Surgery*. 2011;45(1):35-9.
19. Germann G, Harth A, Wind G, Demir E. [Standardisation and validation of the German version 2.0 of the Disability of Arm, Shoulder, Hand (DASH) questionnaire]. *Unfallchirurg*. 2003;106(1):13-9.
20. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee - Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient - reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in health*. 2005;8(2):94-104.
21. WHO (World Health Organization). Process of translation and adaptation of instruments. http://www.who.int/substance_abuse/research_tools/translation/en/
22. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. *Value in Health*. 2011;14(8):967-77.
23. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2—assessing respondent understanding. *Value in Health*. 2011;14(8):978-88.

- 1 24. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends
2 minimum standards for patient-reported outcome measures used in patient-centered outcomes and
3 comparative effectiveness research. *Quality of Life Research*. 2013;22:1889-905.
- 4 25. Zywił MG, Mahomed A, Gandhi R, Perruccio AV, Mahomed NN. Measuring Expectations in
5 Orthopaedic Surgery: A Systematic Review. *Clinical Orthopaedics and Related Research*. 2013;471(11):3446-
6 56.
- 7 26. Sierakowski K, Evans Sanchez, KA, Damarell, RA, Dean, NR, Griffin, PA, Bain, GI. Measuring Quality of
8 Life & Patient Satisfaction in Hand Conditions: A systematic review of currently available patient reported
9 outcome instruments. *Australasian Journal of Plastic Surgery*. 2018;(in press).
- 10 27. Thorne S, Kirkham SR, MacDonald-Emes J. Focus on qualitative methods. Interpretive description: a
11 noncategorical qualitative alternative for developing nursing knowledge. *Research in nursing & health*.
12 1997;20(2):169-77.
- 13 28. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative
14 alternative for developing nursing knowledge. *Res Nurs Health*. 1997;20.
- 15 29. Claire Welford R, Dip N, Kathy Murphy R, Dympna Casey R. Demystifying nursing research
16 terminology. Part 1. *Nurse Researcher (through 2013)*. 2011;18(4):38.
- 17 30. Sandelowski M. Theoretical Saturation. In: Given LM, editor. *The SAGE Encyclopedia of Qualitative
18 Research Methods*. Thousand Oaks: SAGE Publications, Inc.; 2008.
- 19 31. Pope C, Ziebland S, Mays N. Analysing qualitative data. *British medical journal*. 2000;320(7227):114.
- 20 32. Cohen DJ, Crabtree BF. Evaluative criteria for qualitative research in health care: controversies and
21 recommendations. *The Annals of Family Medicine*. 2008;6(4):331-9.
- 22 33. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Quality of life research*.
23 2003;12(3):229-38.
- 24 34. Van Someren M, Barnard Y, Sandberg J. The think aloud method: a practical approach to modelling
25 cognitive. 1994.
- 26 35. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture
27 (REDCap)—a metadata-driven methodology and workflow process for providing translational research
28 informatics support. *Journal of biomedical informatics*. 2009;42(2):377-81.
- 29 36. G R. *Studies in mathematical psychology: 1*. In: Institut. Dp, editor. 1960.
- 30 37. Cano SJ, Hobart JC. The problem with health measurement. *Patient Prefer Adherence*. 2011;5:279-90.
- 31 38. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role
32 of new psychometric methods: Prepress Projects; 2009.
- 33 39. Cano SJ, Klassen A, Pusic AL. The science behind quality-of-life measurement: a primer for plastic
34 surgeons. *Plastic and reconstructive surgery*. 2009;123(3):98e-106e.
- 35 40. Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the
36 Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology*. 2007;46(1):1-18.
- 37 41. Strauss ME, Smith GT. Construct validity: Advances in theory and methodology. *Annu Rev Clin
38 Psychol*. 2009;5:1-25.
- 39 42. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for
40 evaluating the methodological quality of studies on measurement properties: a clarification of its content.
41 *BMC medical research methodology*. 2010;10(1):22.
- 42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

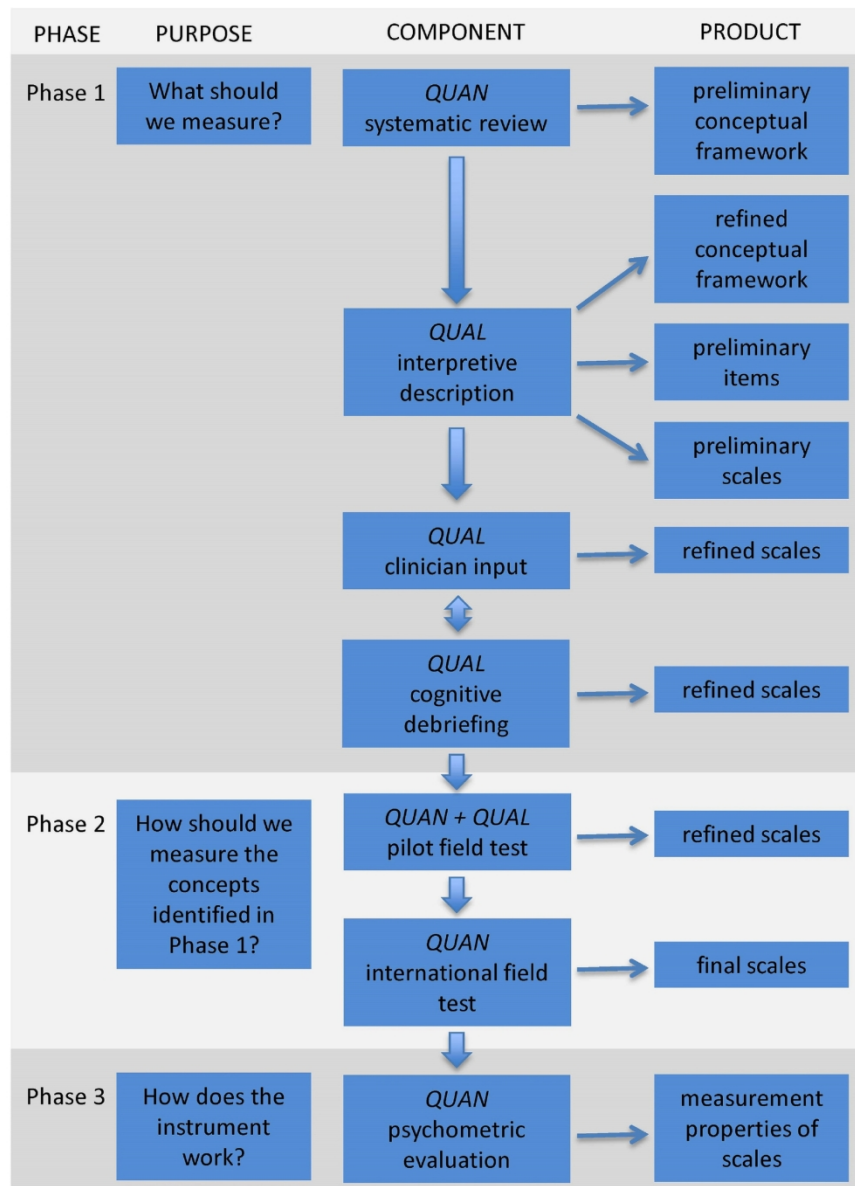


Figure 1 Flow diagram illustrating the multiphase mixed methods approach to the development of the HAND-Q. QUAN, quantitative study component; QUAL, qualitative study component. Image reproduced from Wong Riff et al.(5)

175x241mm (300 x 300 DPI)