

BMJ Open International multiphase mixed methods study protocol to develop a patient-reported outcome instrument for children and adolescents with lower limb deformities

Harpreet Chhina,^{1,2} Anne Klassen,³ Jacek A Kopec,⁴ John Oliffe,⁵ Anthony Cooper¹

To cite: Chhina H, Klassen A, Kopec JA, *et al.* International multiphase mixed methods study protocol to develop a patient-reported outcome instrument for children and adolescents with lower limb deformities. *BMJ Open* 2019;**9**:e027079. doi:10.1136/bmjopen-2018-027079

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-027079>).

Received 15 October 2018
Revised 5 February 2019
Accepted 21 March 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Anthony Cooper;
anthony.cooper@cw.bc.ca,
hchhina@cw.bc.ca

ABSTRACT

Introduction Our recent systematic review has indicated the lack of a patient-reported outcome (PRO) instrument to measure health-related quality of life (HRQOL) of children and adolescents with lower limb deformities. We are developing a PRO instrument which will be applicable internationally across various countries. This manuscript describes our approach to the development of a new PRO instrument for measuring HRQOL for children and adolescents with lower limb deformities.

Methods and analysis Three phases in the development of this PRO instrument are as described: (1) This phase involves the development of a conceptual framework of HRQOL and item pool that is used to inform a set of preliminary scales. We have developed a preliminary conceptual framework of HRQOL based on our systematic review. Qualitative interviews are being conducted at five sites in Canada, Ethiopia, India and the USA. An item pool will be generated from this qualitative phase. The preliminary items and scales will be sent out to children at the five participating centres. Cognitive debriefing interviews will gather detailed feedback on the items from the children. Expert opinion will be sought from clinicians from the participating centres. (2) During this phase, an international field-test study will be conducted to refine the scales and examine their psychometric properties. (3) During this phase, tests of reliability, validity and responsiveness will be conducted. Phase 1 will also involve translations and cultural adaptations. At the end of this study, we expect to produce an internationally applicable PRO instrument which is scientifically sound and clinically relevant to the lower limb deformity population.

Ethics and dissemination This study is approved by Research Ethics Boards for each of the participating sites. Results of this study will be published in peer-reviewed journals and presented at national and international conferences. An integrated knowledge translation approach is applied to engage patients, families and clinicians from the start of the study.

INTRODUCTION

The growing trend towards patient-centred care has led to increased use and development of patient reported outcome (PRO) instruments. PRO instruments measure concepts

Strengths and limitations of this study

- International, multicentre study including patients from high-income and lower-middle income countries. This will ensure international applicability of this patient-reported outcome (PRO) instrument for children and adolescents with lower limb deformities
- Rigorous qualitative component of this study will ensure content validity of this PRO instrument.
- Adherence to rigorous guidelines of PRO instrument development and the use of modern psychometric methods will make this PRO instrument as scientifically sound and clinically relevant as possible.
- A long time period to completion is needed since this study involves multiple iterative steps involving patients and clinicians from high-income and lower-middle income countries.

that matter to patients (eg, health-related quality of life (HRQOL), symptoms and function) from their perspective.¹ Guidelines exist to inform the development of a new PRO instrument, which can be a lengthy, rigorous and iterative process.²⁻⁶ Most PRO instruments are developed in one language and in a single country. Subsequent versions are then translated and adapted for use in other countries when needed.²⁻⁷ A methodologically stronger approach for internationally applicable measures is to develop a new PRO instrument internationally from the start. In this approach, both the qualitative data used to develop scales and subsequent analysis to refine the scales can be used to ensure that final content of a new PRO instrument covers the concerns of patients in different contexts and cultures.

Lower limb deformities describe a range of conditions that may result from congenital defects in fetal development, or are acquired during growth.⁸ They may also arise from

trauma, infection, benign or malignant tumours, or other medical conditions. Lower limb deformities primarily include leg length discrepancy, lower limb deficiency and associated angular and rotational deformities. Epidemiological data vary for each of these abovementioned conditions. Overall, the incidence of lower limb reduction defects is estimated to be 2 per 10 000 live births in the USA.⁹ Leg length discrepancy of greater than 2 cm was found in at least 1 in every 1000 people as shown in a retrospective study from France.¹⁰

Children with lower limb deformities often have physical limitations and gait problems and may also experience pain and discomfort. The abnormal appearance, function and hence the resulting dissatisfaction can discourage participation in recreational and leisure activities in these children, which can, consequently, be associated with behavioural, emotional, psychological and social adjustment problems.^{11–14} The problems listed above, along with the complex and long treatment procedures involved for lower limb deformities, can have a considerable effect on the HRQOL of these children.^{15–18}

HRQOL is a multidimensional concept that measures a subject's perception of the effects of illness and/or treatment(s) on their daily life, well-being and physical, psychological and social functioning.^{3 19–21}

For accurate measurement of concepts such as HRQOL, it is important to have psychometrically sound (both reliable and valid) measurement instruments.

A systematic review indicated that no PRO instrument exists to measure HRQOL for children and adolescents with lower limb deformities.²² Due to the lack of a condition-specific PRO instrument for lower limb deformities, the HRQOL of children with lower limb deformities has been measured by either using a generic or a parent-reported HRQOL instrument.²²

To address this gap, our aim is to develop an internationally applicable PRO instrument to measure HRQOL of children and adolescents with lower limb deformities. In this article, we describe our methodology and progress to date.

The HRQOL tool developed at the end of this project will help assess the baseline HRQOL of children with lower limb deformities. In the long term, this tool will be used by healthcare professionals such as orthopaedic surgeons, physiotherapists, occupational therapists, podiatrists, nurses and other healthcare professionals involved in the care of children with limb deformities to look at the effectiveness of relevant interventions for these children in terms of improving their HRQOL. The evidence generated by the long-term use of this HRQOL instrument about the effectiveness of various treatments will eventually help develop or modify interventions/treatment plans, healthcare services and target care to achieve a better HRQOL for children and adolescents with lower limb deformities. The development of a PRO instrument will benefit future patients with lower limb deformities by measuring their HRQOL and thus helping improve their health outcomes. This new PRO instrument will also

help healthcare professionals, patients and their families in making important decisions about treatment choices, such as amputation versus reconstruction.

METHODS

We adhere to best practice guidelines for development of PRO instruments.^{2–4 20 21 23} Our team follows a mixed methods iterative approach (see figure 1).²⁴ The qualitative phase will ensure that the items generated are grounded in the experiences of the children and adolescents with lower limb deformities and the psychometric analysis will determine how well the items measure the concept of interest (COI). Three iterative and interactive phases involved include the following: (1) development of a conceptual framework of HRQOL and item pool that is used to inform a set of preliminary scales; (2) an international field-test study to refine the scales and examine their psychometric properties; and (3) further tests of reliability, validity and responsiveness (figure 1).

Patient and public involvement

There was no direct patient and public involvement during the protocol development. However, this study itself involves direct patient engagement.

Phase 1: what should we measure

This phase involves the development of a conceptual framework and generation of an item pool to inform scale development. This phase will establish content validity of the new instrument.^{4 20} This phase has three steps as explained below.

Step 1 involves development of a conceptual framework of HRQOL for children and adolescents with lower limb deformities (in progress).

Systematic review

We performed a systematic review to identify PRO instruments used to measure the HRQOL of children and adolescents with lower limb deformities. Findings were used to develop a preliminary conceptual framework.²²

Qualitative interviews

Rigorous qualitative research is a crucial foundation in the development of PRO instruments.²⁵

Recruitment

Participants are being recruited from five orthopaedic centres in high-income (Canada—two centres and USA), lower-middle income (India) and low-income countries (Ethiopia). Our aim is to recruit a maximum variation sample that differs by type of lower limb deformities, type and stage of treatment, age, gender and country. All interviews at one Canadian and one US English speaking centre and one Indian Hindi/Punjabi speaking centre are conducted by one lead interviewer (fluent in Punjabi, Hindi and English) at the main study site (the site where the first and the senior author are located). Interviews at the French speaking Canadian centre and Amharic

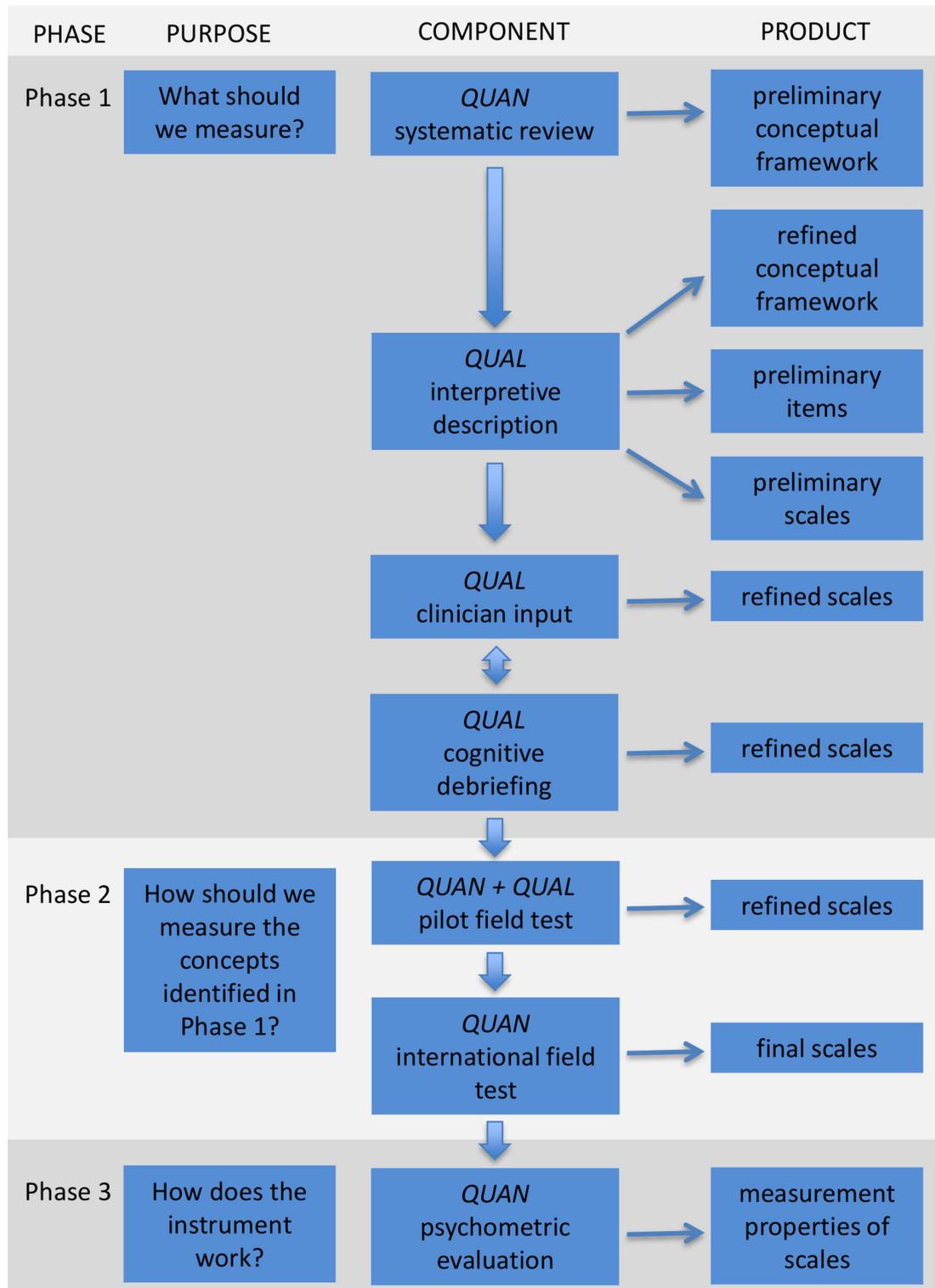


Figure 1 Flow diagram showing the multiphase mixed methods protocol to develop a patient-reported outcome instrument for children and adolescents with lower limb deformities. It is important to note that the process can be iterative and interactive as opposed to a strict linear process. This figure has been reproduced with permission from Wong Riff K W Y, Tsangaris E, Goodacre T, *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:e015467. QUAN, quantitative study component; QUAL, qualitative study component.

speaking Ethiopian centre are conducted by local interviewers at those respective sites, trained by the lead interviewer at the main study site to ensure uniformity during the interview process.

Inclusion and exclusion criteria

Participants are aged 8 to 18 years with any of the following congenital or acquired lower limb conditions as diagnosed by an orthopaedic surgeon: lower limb reduction defects,

leg length discrepancy and associated angular and rotational deformities. Isolated knee, foot and ankle conditions without any limb length discrepancy or limb deformity are not included. Children and adolescents with comorbidities that may have confounding effects on their HRQOL and/or cognitive or developmental delay that may affect their ability to communicate are excluded from this study.

An interview guide with open-ended questions and probes covering concepts from other PRO instruments identified in the systematic review was developed.²² The interview guide is modified after each interview as needed to include new themes that are then raised in subsequent interviews. The interview guide was prepared in English and translated into French, Amharic and Hindi/Punjabi. French translations were done by the native French speaker who also conducted the interviews at the French speaking site. Amharic translations were conducted by a native Amharic speaker who also conducted the interviews at the Amharic speaking site. Hindi/Punjabi translations were conducted by the lead interviewer who is a native Hindi/Punjabi speaker and also conducted interviews at the Hindi/Punjabi speaking site. Interviews will be conducted until data saturation.²⁶

Data collection

Parent consent and participant assent is obtained from each family. Demographic (age, gender, ethnicity) and clinical (diagnosis, type and stage of treatment) data are collected for each participant. Individual face-to-face interviews are conducted with the participant and parent(s) where possible. Interviews are audio-recorded and transcribed/translated into English. English interviews are transcribed by professional transcriptionist fluent in English. French interviews are translated and transcribed by a professional bilingual transcriptionist. Amharic translation and transcription are conducted by the same person conducting the interviews who is a native Amharic speaker. Hindi/Punjabi interviews are translated and transcribed by the lead interviewer who is native speaker of those languages.

Participants from the main study site in Canada are also invited to take part in photo elicitation interview (PEI) whereby they are invited to take pictures to show how their life is affected by their leg condition. These PEI using participant produced photographs will help us overcome any age-related linguistic and cognitive barriers in our young participants.²⁷

Data analysis

The qualitative methodology used is interpretive description (ID). ID is a 'non-categorical methodological approach' that acknowledges the theoretical and practical knowledge that the researchers bring to a project to generate relevant knowledge for a clinical context.^{28 29} The underlying qualitative paradigm is 'pragmatism' which means that an individual's (patients' and/or caregivers') understanding of a concept is of

greatest importance in understanding the impact of their condition on their HRQOL, regardless of the clinical explanations.³⁰

Interviews in English are transcribed verbatim. Interviews in other languages are transcribed and translated into English. All data are analysed by one researcher at the main study site. A line-by-line approach is used to code the data, with constant comparison methods to identify and categorise the concepts of interest (COI). The identified concepts are categorised into overarching top-level domains and subdomains. Data collection and analysis are being conducted concurrently and iteratively. The preliminary conceptual framework is revised based on the analyses.

Rigour

The truth value, applicability, consistency and neutrality will be assessed to establish rigour for the qualitative phase.^{31 32} We use member checking as a way of establishing the truth value and credibility of our results. The applicability of the results will be determined in terms of fittingness of the quotations from the study participants to the scales derived from this data.³¹ The scales will be taken back to the study participants for their feedback. The consistency of the qualitative findings will be evaluated by their auditability. Auditability will be maintained by keeping a clear decision trail during analysis. We will use confirmability as the criterion for neutrality, which means that we will be reflexive about the research process and the findings

Step 2 will involve item generation and scale development.

Item generation

Coded qualitative data will be used to generate a comprehensive item pool. In developing draft items, the wordings of patients will be used as much as possible to ensure that the items will resonate well with them. The item pool will be analysed to identify common/unique issues and how these vary by age (child vs adolescents), leg condition and country. We will follow guidelines in regard to appropriate health-related vocabulary, reading level, response scale, response options, recall period and length of instrument.^{23 33-36} We will use the item pool to create independently functioning scales that act like 'rulers' where the items map out a clinical hierarchy for each of the COI.

Step 3 will involve clinical input, cognitive debriefing interviews and refining the scales.

Clinician input

Clinicians (orthopaedic surgeons, nurses, occupational therapists, physiotherapists, podiatrists and psychologists) and other experts (experts in HRQOL studies and measurement) will be invited to provide feedback on the scales using a Research Electronic Data Capture (REDCap) survey.

Cognitive debriefing interviews

Cognitive debriefing interviews with participants will be conducted to determine if patients understand the instructions, response options and items of the new HRQOL instrument and to identify missing content. Participants will review the HRQOL instrument with the interviewer using the ‘Think Aloud’ technique with verbal probing.^{37 38} Interviews will take place in a series of rounds to allow time to make changes to the instrument and then obtain feedback on those changes. Interviews will start in English centres and continue until no further changes are needed. Translated versions (pragmatic translations) of items and scales will be created (one country at a time) and cognitive debriefing interviews will be conducted in turn in French, Amharic and Hindi/Punjabi.

Translation and cultural adaption

We will follow the principles of good practice for translation and cultural adaption (TCA) for PRO instruments.^{2 7 39} The first step will involve two independent forward translations from English to each target language (Canadian French, Amharic, Hindi and Punjabi). All translations will be conducted by the native speakers of that language who are also fluent in English. The two forward translations will be compared and reconciled. A back translation of the reconciled versions will be conducted by the native speakers of that language who are also fluent in English and reviewed by the research team (research team has members who are native speakers of English, French, Amharic, Hindi and Punjabi and are also fluent in English). Harmonisation of the multiple language versions will identify any discrepancies across language which will then be addressed. Cognitive debriefing interviews with a small group of children (at least five) with lower limb deformities in each country will be conducted to determine understandability, interpretation and cultural relevance of each translation. Translations will be revised and finalised based on the feedback obtained.

Phase 2: how should we measure the concepts identified in Phase 1 and which questions are effective?

We will field test the new PRO instrument in a large sample of children and adolescents with lower limb deformities in Canada, Ethiopia, India and the USA. The data collected will be analysed to examine scale performance and to develop their scoring algorithms. We will use Rasch measurement theory (RMT) analysis to examine item and scale performance against a set of psychometric criteria.^{40–42} In the RMT approach, scales are created along which an individual is placed based on the probability that the respondent answered the items in a certain way. RMT scales provide interval level measurement, which allows to accurately measure change over time.⁴³ A scale based on the RMT analysis will provide person estimates which are independent of the sampling distribution of the items. The psychometric properties of the scales will be defined using the RMT analysis.

Pilot field testing

A preliminary RMT analysis will be conducted with the pilot field test data from the Canadian and US centres to identify any changes needed prior to launching the international field test study. Participants will also be asked to provide qualitative feedback on the new instrument specifically about the length of the instrument, anything in the instrument that they do not like and any other comments on improving the instrument.

Study participants

Participant eligibility will be the same as in the qualitative phase.

Data collection

Participants will be invited to complete the new PRO instrument. Information on age at the time of data collection, sex, type of deformity and type and stage of treatment will also be collected.

Data analysis

The results from the RMT analyses and the qualitative feedback will be used to revise the scales.

International field test and RMT analysis

International field testing will be done at multiple centres to collect data from a large sample of children and adolescents with lower limb deformities.

Study participants

Participants from current collaborating centres and additional centres will be included. Eligibility criteria will be the same as in the qualitative interviews. Target sample size will be a minimum 108–200 participants per country in order to achieve item calibrations that are stable within 0.5 logits (person location estimates) with a 99% CI.⁴⁴

Data collection

Mode of data collection will be based at the discretion of each site and can include paper version or electronic with data entered directly into a REDCap.⁴⁵ Demographic data collected from each site will include age at the time of data collection, gender, ethnicity, country, language spoken, type of limb deformity, type and stage of treatment and any future treatment plans.

Data analysis

Data will be analysed using Rumm2030 software.⁴⁶ In RMT analysis, a number of tests and criteria are applied to examine item and scale performance. (1) Threshold for item response options must be ordered, that is, a ‘1’ on a 4-point scale must be lower on the continuum than a ‘2’. RMT analysis defines the hierarchy of items on the scale, from the easiest question for a participant to endorse to the hardest question to endorse. (2) Item fit statistics will be used to evaluate whether the data fits the Rasch model. Three indicators of fit to be used are log residuals (item person interaction), χ^2 values (item–trait interaction) and item characteristic curves.^{41 47} These

item fit statistics will be interpreted along with their clinical relevance. Items that do not fit these three criteria will be dropped unless they pertain to clinically important concepts. (3) All items of a scale should define a continuum. Inspecting where items are located on the continuum in respect to the other items will indicate how well the items map out a particular construct. The range of construct measured by a scale will be compared with the range of the construct experienced by the target study population (scale to sample targeting).

Unidimensionality

This will be checked to find out whether all the items on a scale measure only one common construct.⁴⁸ If the observed data (for each scale) from the field test fit the Rasch model, it will be assumed that the individual scales are unidimensional.

Internal consistency reliability

This will measure how items are related to each other on a scale. Person Separation Index will be used to measure the precision of a person estimate.⁴¹

Differential item functioning

This test identifies items that perform differently (bias) for subgroups within a sample. We will examine differential item functioning (DIF) by age-group, gender and country. When items are significant for DIF, the measurement ruler is not stable across the subgroups.⁴⁰

Item reduction

The graphical and statistical tests described above will be examined together to identify items with poor performance that could be dropped. Distribution of item locations and clinical relevance of items will also guide decisions about item reduction. RMT analysis will also provide a scoring system for the scale.

Normative data

Normative data scores for the field test participants will be calculated using the scoring system. Normative data and associations between the demographic variables and the test scores will be examined.

Construct validity

This scale property will be examined by assessing the internal structure (structural validity) of the new PRO instrument, internal relationships, hypothesis testing and cross-cultural validity.⁴⁹ The internal structure (dimensionality) and cross-cultural validity (DIF) will be tested during the RMT analysis. Specific a priori hypotheses include an indication of the expected direction and magnitude of correlations or differences in responses obtained from certain patient groups.^{50–52} Children and adolescents with untreated lower limb deformity will have a lower HRQOL as compared with the children who have received treatment.

Phase 3: how does the instrument work?

This phase will involve further evaluation of measurement properties of the final version of the HRQOL instrument including its reliability, validity and responsiveness.⁵²

Study participants

The inclusion criteria for participants in this phase of the study will be the same as the previous phases. For responsiveness, lower limb deformity participants that are undergoing limb lengthening/deformity correction using various methods (eg, external fixator device, internal fixation, epiphysiodesis or amputation) will be recruited from collaborating sites in Canada and USA.

Test–retest reliability

Test–retest reliability will be established by asking participants to complete the new PRO instrument 2 weeks after the initial assessment. Intraclass correlation coefficient of 0.70 will be considered acceptable for the scales being tested in participants whose leg condition is stable during the time interval.⁵³

Construct validity

Construct validity will be established by administering the new HRQOL instrument alongside the PedsQL, a generic HRQOL instrument measuring physical, social, emotional and school function.⁵⁴ Scores from the generic PRO instrument measuring similar domains as the new PRO instrument will be expected to correlate more strongly than domains measuring dissimilar constructs.

Criterion validity

Criterion validity is the degree to which the scores from the new PRO instrument adequately reflect a ‘gold standard’ instrument.⁵¹ The aim of this analysis is to compare the new PRO with another instrument used in the past in patients with lower limb conditions, that is, the Lower Limb Outcomes Questionnaire.⁵⁵ The Lower Limb Outcomes Questionnaire measures pain, stiffness, swelling and function. We expect scores from the physical health subscale of the new PRO instrument to correlate more with scores from the existing PRO instrument as compared with the correlation between other subscales such as psychological health and social health.

Responsiveness is the ability of the instrument to detect clinically meaningful change over time in the construct being measured.⁵¹ In order to assess the responsiveness of a PRO instrument, anchor-based and distribution-based approaches are recommended.^{56–59} RMT analysis allows for an increased detection of responsiveness.⁵⁹ The participants will be asked to complete the new PRO instrument both preoperatively and an appropriate time postoperatively to determine responsiveness.

For patient-based anchors, self-reported global assessment of change in their HRQOL will be used to categorise patients into groups that reflect different amounts of change in their HRQOL. Patients will be asked to classify them into four groups: no improvement, small improvement, medium improvement and large improvement

after removal of the external fixator. Minimally important difference will be determined as average change (post minus pre) observed in the score on the new HRQOL instrument in the small improvement group using the transformed Rasch scores. A distribution-based approach will be used. The Rasch scores before and after treatment will be compared using paired t-tests. An effect size and standardised response mean will be calculated to measure change.^{56 59} Finally, the significance of a person's own change will be calculated using individual person estimates, which are associated with standard errors.⁵⁹

ETHICS

All information collected is securely stored at institutional computers in password-protected files. Maintaining patient confidentiality presents challenges in any qualitative research. However, in this study, the authors will try to maintain patient confidentiality as much as possible. All qualitative data will be clearly examined to remove any personally identifying information and will be stored securely at each of the participating institutions.

Study participants may be asked to discuss about issues related to their limbs that are sensitive and may trigger distress. In order to address this issue, participants are made aware of this during the consenting process and are given an option to follow-up with the study team.

Dissemination

This study will take an integrated knowledge translation approach. Collaborations with multiple sites internationally will hopefully result in increased uptake and the use of this PRO instrument in the future. The results of this study will be published in peer-reviewed journals.

Expected outcome and significance

We expect to produce an international PRO instrument that is scientifically sound and clinically relevant for lower limb deformities. Once developed, we expect that this new PRO instrument will be used internationally by healthcare professionals and researchers to examine the effectiveness of interventions in terms of improving the patient HRQOL.

Author affiliations

¹Orthopaedics, BC Children's Hospital, Vancouver, British Columbia, Canada

²Experimental Medicine, University of British Columbia Faculty of Medicine, Vancouver, British Columbia, Canada

³Pediatrics, McMaster University, Hamilton, Ontario, Canada

⁴Arthritis Research Centre of Canada, Richmond, British Columbia, Canada

⁵School of Nursing, University of British Columbia, Vancouver, British Columbia, Canada

Acknowledgements The authors acknowledge their collaborating centres for contributing patients with limb deformities during various phases of data collection for this study.

Contributors HC, AK and AC designed the study and contributed towards the manuscript. JAK and JO contributed towards the study design and manuscript.

Funding This work is supported by Orthopaedics Research Excellence Funds, Department of Orthopaedics, University of British Columbia (Grant # 20R24312); Canadian Orthopaedics Association (FAS# F16- 05491).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study is approved by Research Ethics Boards for each of the participating sites (CW15- 0215/H15-00514, IRB #: 847125 Nemours/Alfred I. duPont Hospital for Children, 20160808-IRB, FWA 00004545, IEC-01/20 17-579) and by the main study centre and from each of the participating sites (The University of British Columbia/Children's and Women's Health Centre of British Columbia Research Ethics Board; Institutional review Board at Nemours/Alfred I. duPont Hospital for Children).

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

1. Weldring T, Smith MS. PROs and PROMs Health Services Insights Weldring. *Heal Serv Insights* 2013;6:61–8.
2. Wild D, Grove A, Martin M, et al. ISPOR Task Force for Translation and Cultural Adaptation. 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 Health* 2005;8:94–104.
3. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905.
4. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;4:79.
5. WHO. No Title [Internet]. <http://www.who.int/healthinfo/survey/whoqol-qualityoflife/en/> (cited 8 Nov 2016).
6. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials. *JAMA* 2013;309:814.
7. Wild D, Eremenco S, Mearl, et al. Multinational trials—recommendations on the translations required, approaches to using the same language in different c. 2009;12:430–40.
8. Wilson GN. Heritable limb deficiencies. In: Herring JA, Birch JG, eds. *The Child With a Limb Deficiency*. Rosemont, IL: American Academy of Orthopaedic Surgeons, 1998:39–51.
9. Canfield MA, Honein MA, Yuskiv N, et al. National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999–2001. *Birth Defects Res A Clin Mol Teratol* 2006;76:747–56.
10. Guichet JM, Spivak JM, Trouilloud P, et al. Lower limb-length discrepancy. An epidemiologic study. *Clin Orthop Relat Res* 1991;272:235–41.
11. Setoguchi Y, Rosenfelder R. *The limb deficient child*. Springfield, IL, 1982.
12. Cadman D, Boyle M, Szatmari P, et al. Chronic Illness, Disability, and Mental and Social Well-Being: Findings of the Ontario Child Health Study [Internet]. 1987;79:805–13 <http://pediatrics.aappublications.org/content/79/5/805.abstract>.
13. Michielsen A, van Wijk I, Ketelaar M. Participation and health-related quality of life of Dutch children and adolescents with congenital lower limb deficiencies. *J Rehabil Med* 2011;43:584–9.
14. Michielsen A, Van Wijk I, Ketelaar M. Participation and quality of life in children and adolescents with congenital limb deficiencies: A narrative review. *Prosthet Orthot Int* 2010;34:351–61.
15. Hrutkay JM, Eilert RE. Operative lengthening of the lower extremity and associated psychological aspects: the Children's Hospital experience. *J Pediatr Orthop* 1990;10:373–7.
16. Ghoneem HF, Wright JG, Cole WG, et al. The Ilizarov Method for Correction of Complex Deformities. *J Bone Jt Surgery, Am Vol* 1996;78:1480–5.
17. Ramaker RR, Lagro SW, van Roermund PM, et al. The psychological and social functioning of 14 children and 12 adolescents after Ilizarov leg lengthening. *Acta Orthop Scand* 2000;71:55–9.
18. Montpetit K, Hamdy RC, Dahan-Oliel N, et al. Measurement of health-related quality of life in children undergoing external

- fixator treatment for lower limb deformities. *J Pediatr Orthop* 2009;29:920–6.
19. Karimi M, Brazier J. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? *Pharmacoeconomics* 2016;34:645–9.
 20. Patrick DL, Burke LB, Gwaltney CJ, 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;14:967–77.
 21. Patrick DL, Burke LB, Gwaltney CJ, 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 Health* 2011;14:978–88.
 22. Chhina H, Klassen A, Kopec J, et al. Quality of life of children with lower limb deformities: A systematic review of patient-reported outcomes and development of a preliminary conceptual framework. *J Limb Lengthening Reconstr* 2017;3:19–29.
 23. Matza LS, Patrick DL, Riley AW, et al. Pediatric patient-reported outcome instruments for research to support medical product labeling: report of the ISPOR PRO good research practices for the assessment of children and adolescents task force. *Value Health* 2013;16:461–79.
 24. Creswell JW, Plano Clark VL, Gutmann M, et al. *Advanced mixed methods research designs Handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage, 2003.
 25. Lasch KE, Marquis P, Vigneux M, et al. PRO development: rigorous qualitative research as the crucial foundation. *Qual Life Res* 2010;19:1087–96 <http://www.jstor.org.ezproxy.library.ubc.ca/stable/40927744>.
 26. Bowen GA. Naturalistic inquiry and the saturation concept: a research note. *Qualitative Research* 2008;8:137–52.
 27. Clark CD. The Autodriven interview: A photographic viewfinder into children's experience. *Visual Sociology* 1999;14:39–50.
 28. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative alternative for developing nursing knowledge. *Res Nurs Health* 1997;20:169–77.
 29. Thorne S. *Interpretive Description*. 2nd ed, 2016.
 30. Welford C, Murphy K, Dymna C. Demystifying nursing research terminology: part 2. *Nurse Res* 2012;19:29–35.
 31. Guba EG, Lincoln YS. Effective Evaluation [Internet]. 1981 <http://books.google.co.jp/books?id=3fxGAAAAMAAJ>.
 32. Sandelowski M. The problem of rigor in qualitative research. *Advances in Nursing Science* 1986;8:27–37.
 33. Matza LS, Swensen AR, Flood EM, et al. Assessment of health-related life in children: a review of conceptual. *methodological and regulatory issues BT - Val Health* 2004.
 34. Clark LA, Watson D. *Constructing validity: Basic issues in objective scale development*: Psychol Assess, 1995.
 35. Borgers N, Hox J, Sikkil D. Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral mid-point. *Qual Quant* 2004.
 36. Weijters B, Cabooter E, Schillewaert N. The effect of rating scale format on response styles: The number of response categories and response category labels. *Int J Res Mark* 2010.
 37. Willis G. Cognitive interviewing: A tool for improving questionnaire design. *Sage* 2005.
 38. Willis GB. *Analysis of the Cognitive Interview in Questionnaire Design: Understanding Qualitative Research*: Oxford Univ Press, 2015:1–275.
 39. WHO (World Health Organization). Process of translation and adaptation of instruments [Internet]. http://www.who.int/substance_abuse/research_tools/translation/en/ (cited 2017 Jan 19).
 40. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess* 2009;13:iii, ix-x, 1-177.
 41. Andrich D. *Rasch models for measurement*. Newbury Park, Ca: SAGE Publications, Inc, 1988.
 42. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. *Inf. Control* 1961;4:382.
 43. Hobart JC. *The problem with health measurement*, 2011:279–90.
 44. Linacrej. Sample size and item calibration stability. *Rasch Meas Trans* 1994;7.
 45. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
 46. RUMM Lab. Welcome to RUMM. <http://www.rummlab.com.au/> (cited 1 Feb 2017).
 47. Wright B GM. *Rating scale analysis: Rasch measurement*. Chicago: MESA, 1982.
 48. Fayers PM, Hand DJ, Bjordal K, et al. Causal indicators in quality of life research. *Qual Life Res* 1997;6:393–406.
 49. Cronbach LJ, Meehl PE. *Psychological Bulletin*, 1955.
 50. Strauss ME, Smith GT. *Construct Validity : Advances in Theory and Methodology. Rev, Annu Psychol, Clin*, 2009.
 51. Mokkink LB, Terwee CB, Knol 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;10:22.
 52. Mokkink LB, Terwee CB, Patrick 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;19:539–49.
 53. Giraudeau B, Mary JY. Planning a reproducibility study: how many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. *Stat Med* 2001;20:3205–14.
 54. Varni JW, Burwinkle TM, Katz ER, et al. The PedsQL in pediatric cancer: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer* 2002;94:2090–106.
 55. Johanson NA, Liang MH, Daltroy L, et al. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am* 2004;86-A:902–9.
 56. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.
 57. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171–8.
 58. Sprangers MA, Moinpour CM, Moynihan TJ, et al. Assessing meaningful change in quality of life over time: a users' guide for clinicians. *Mayo Clin Proc* 2002;77:561–71.
 59. Hobart JC, Cano SJ, Thompson AJ. Effect sizes can be misleading: is it time to change the way we measure change? *J Neurol Neurosurg Psychiatry* 2010;81:1044–8.