Development of a new patient-reported outcome measure for Dupuytren disease: A study protocol

David Eckerdal 🆔
Department of Orthopedics, Hässleholm-Kristianstad Hospitals, Hässleholm, Sweden

Department of Clinical Sciences - Orthopedics, Lund University, Lund, Sweden

Per-Erik Lyrén
Department of Applied Educational Science, Umeå University, Umeå, Sweden

Jane McEachan
Fife Hand Clinic, Department of Orthopaedic Surgery, National Treatment Centre Fife, Victoria Hospital, Kirkcaldy, Scotland

Anna Lauritzson
Department of Orthopedics, Hässleholm-Kristianstad Hospitals, Hässleholm, Sweden

Jesper Nordenskjöld
Department of Clinical Sciences - Orthopedics, Lund University, Lund, Sweden

Department of Hand Surgery, Skåne University Hospital, Malmö, Sweden

Isam Atroshi
Department of Orthopedics, Hässleholm-Kristianstad Hospitals, Hässleholm, Sweden

Department of Clinical Sciences - Orthopedics, Lund University, Lund, Sweden

Abstract
Objectives: Dupuytren disease is a common condition that causes progressive finger contractures resulting in impaired hand function and difficulties in performing daily activities. Patient reported

Corresponding author:
David Eckerdal, Department of Orthopedics, Hässleholm Hospital, Hässleholm 28 125, Sweden.
Email: david.eckerdal@med.lu.se

outcome measures (PROMs) are commonly used in research and clinical practice to evaluate treatment outcomes. Both general upper extremity PROMs and Dupuytren-specific PROMs are available, typically developed using conventional methodology based on classical test theory. However, most current PROMs have been shown to have low responsiveness and the relevance of included items have been questioned. In this study we aim to develop a new Dupuytren-specific PROM using modern measurement methodology based on item response theory (IRT). **Methods:** The study will be performed in three phases. In Phase 1, (completed), an expert group developed a questionnaire with a large number of potentially relevant items derived from existing PROMs and patient collaboration. In Phase 2, the questionnaire will be administered to 300 patients with Dupuytren disease, and their responses will be analyzed with IRT methodology to identify the best performing items to be included in the new PROM. In Phase 3, the new PROM will be administered to 300 additional patients for validation. **Conclusion:** This new Dupuytren-specific patient-reported outcome measure will help advance clinical research on Dupuytren disease.

Introduction

Dupuytren disease is a common condition that causes progressive finger joint contractures resulting in impaired hand function and difficulties in performing daily activities. In a population-based study the prevalence of doctor-diagnosed Dupuytren disease in the general population of Skåne region in Southern Sweden was 1.4% in men and 0.5% in women.[1] Among men aged 50 years or older, 2.7% had been diagnosed with Dupuytren disease and it is estimated that annually 3 per 1000 seek healthcare for Dupuytren disease for the first time.[1] Patients with Dupuytren disease are commonly treated with surgery to improve hand function. Surgery is a major intervention that commonly requires prolonged rehabilitation and is associated with potential complications.[2] Minimal-invasive treatment methods, such as collagenase injection and needle fasciotomy, are also commonly used in clinical practice and new treatments are being researched.[3] Regardless of method, treatment of Dupuytren disease is common, with approximately 68,000 Dupuytren procedures performed in the United States during 2013.[4] In Sweden, surgical fasciectomy was performed on 36 per 100,000 persons aged 40 years or older in 2022; the number of other procedures such as needle fasciotomy is unknown.[5]

To compare the outcomes of different treatment methods it is essential to have reliable and valid outcome measures that can accurately detect benefits that are relevant to patients with Dupuytren disease.

In clinical practice and in research, range of motion of affected finger joints before and after treatment and rate of contracture recurrence have been the most commonly used measures of disease severity and treatment outcome.[6] While range of motion provides a measure of disease severity it has disadvantages, such as variability between examiners.[7] In addition, patients have different demands, and therefore patients with similar degrees of contracture may not have the same activity limitations. Consequently, it is important that patient-reported outcome measures (PROMs) are included when assessing treatment outcomes in clinical practice and in research.[8] In a systematic review, the 30-item disabilities of the arm, shoulder and hand (DASH) questionnaire was found to be the most commonly used PROM in assessing patients with Dupuytren disease.[6] The

QuickDASH, an 11-item form of the DASH[9] has been shown to be a valid PROM for patients with Dupuytren disease and has also been commonly used in Dupuytren research.,[6,10] The Unité Rhumatologique des Affections de la Main (URAM), developed as a disease-specific measure, is commonly used in Dupuytren research.[11] Another measure, the Southampton Dupuytren scoring scheme (SDSS) was also developed as a disease-specific measure but appears in only few studies.[12] The shortcoming common to all current PROMs is the relatively low responsiveness in patients treated for Dupuytren disease. High responsiveness is an essential property when using a PROM to compare treatment effectiveness. Almost all longitudinal studies that have used current PROMs to assess treatment effectiveness in Dupuytren disease have shown only small or moderate effect sizes, which is probably inadequate when comparing the effectiveness of different treatment methods.[13] Inadequate responsiveness may be related to the number of PROM items that have little relevance with respect to the activity limitations experienced by patients with Dupuytren disease. There is little use for an outcome measure on which a large proportion of patients score the best or almost the best on the scale, which is common for current PROMS used in Dupuytren disease.[6,13–15] Identifying the relevant items will be the first step to develop a PROM that has an adequate responsiveness to detect differences in the effectiveness of different treatment methods.

Patients with Dupuytren disease may present with multiple activity limitations not captured by currently available PROMs. Therefore, patient-specific measures, such as the Canadian Occupational Performance Measure (COPM), have been advocated as an alternative to fixed-items PROMs.[16,17] In a semi-structured interview the COPM allows patients to identify and rate their performance with up to 5 activities.[18] In a previous study, we have shown that the COPM is a highly responsive measure for evaluating treatment outcome in patients with Dupuytren disease.[17] However, use of an interview-based PROM may not be practical in large studies.

Most current PROMs used in Dupuytren research were developed using traditional methodology based on classical test theory (CTT). While CTT-based methods have been useful in the development of many PROMs, it is generally recognized that modern measurement methodology, based on item response theory (IRT), has advantages in developing new PROMs.[19,20] Methodology based on IRT provides additional benefits, such as person and item parameters that are on the same scale, use of item response models where the fit to the data can be examined, person parameter error estimates across the scale instead of just one overall reliability estimate, and ability to examine differential item functioning (DIF) that is integrated in the item response modeling.[20,21]

Different treatment methods that vary in complexity and cost are used for patients with Dupuytren disease. Assessment of treatment effectiveness requires accurate measurement of outcomes that are relevant and important to patients. This is not adequately obtained by currently available PROMs. Thus, there is a need for a new patient-reported measure of activity limitations in patients with Dupuytren disease that is reliable and valid and can accurately measure outcomes that are important to patients with Dupuytren disease.

The aim of this study is to develop a new disease-specific PROM for Dupuytren disease using modern measurement methodology.

Material and methods

Study setting

This is a multi-center study including two centers in Sweden and one in Scotland. The participating centers are the Department of Orthopedics, Hässleholm Hospital, Hässleholm, Sweden (primary center), the Department of Hand Surgery, Skåne University Hospital, Malmö, Sweden and the Fife

Hand Clinic, Department of Orthopaedic Surgery, National Treatment Centre Fife, Victoria Hospital, Kirkcaldy, Scotland.

Study phases

The study will be conducted in 3 phases (Figure 1).

Phase 1 (item selection). The first phase of the study was the development of the initial questionnaire. This began with an extensive literature search to identify PROMs related to upper-extremity conditions. The search yielded a total of 30 PROMs (supplemental material 1). These were screened for items/activities that may be relevant for patients with Dupuytren disease. The items were then supplemented with a number of additional activities identified in a previous study at the study center,[17] as well as items from a similar PROM development study from Canada.[22] In total approximately 900 items were used to create the initial list of items. During several meetings this list of items was reviewed by an expert group consisting of a hand surgeon, two orthopedic surgeons, and two occupational hand therapists. During the first meetings, duplicates were removed and when two or more items had similar content, the one deemed to fit best was selected. In subsequent meetings items deemed less relevant for Dupuytren disease were removed. The list of preliminary items was given to two patients with Dupuytren disease for review and their opinions and suggestions were taken into consideration through discussion within the group. The final product was the construction of a questionnaire in an English and a Swedish version, representing activities deemed relevant for patients with Dupuytren disease. Most of the final items used in the questionnaire were originally in English. For the Swedish version we used the wording from validated Swedish versions of the original items when available. If no validated Swedish version was available, we used wording from similar items in validated Swedish PROMs. The final questionnaire starts with two questions that inquire about hand dominance and which hand the patient seeks treatment for, followed by 85 items (supplemental material 2) (68 items inquiring about activity limitations, 11 items regarding symptoms in the affected hand and 6 items regarding satisfaction, self-confidence etc.) and finally 1 free-response question (similar to that in the COPM) asking patients to list up to 3 activities that they find most difficult to perform because of their Dupuytren disease. All items regarding activity limitations have 5 response options: "no difficulty", "mild difficulty", "moderate difficulty", "severe difficulty", "unable to perform", and the option of "not applicable". The same 5 response options are provided for the 3 most important activities that the patients list and rate. The other items have 5 response options ranging from "not at all" to "very often", or from "very satisfied" to "very dissatisfied".

Phase 2 (development sample). All patients referred to the study centers because of Dupuytren disease are screened for inclusion according to pre-defined eligibility criteria: Patient 18 years or older diagnosed with Dupuytren disease by an orthopedic or hand surgeon, able to independently understand and consent to participation and able to respond to questionnaires (Table 1). Eligible patients receive written information and are asked about participation in the study. Patients who accept to participate in the study receive the questionnaire either by mail, prior to their hospital visit, and are asked to return the completed questionnaire, or complete the questionnaire during their visit. Baseline data regarding gender, age, hand dominance, smoking status, diabetes, family history of Dupuytren disease, previous treatments, bilateral disease, and occupation are collected (Table 2). Passive extension deficit of the joints of the affected fingers is measured using a goniometer.
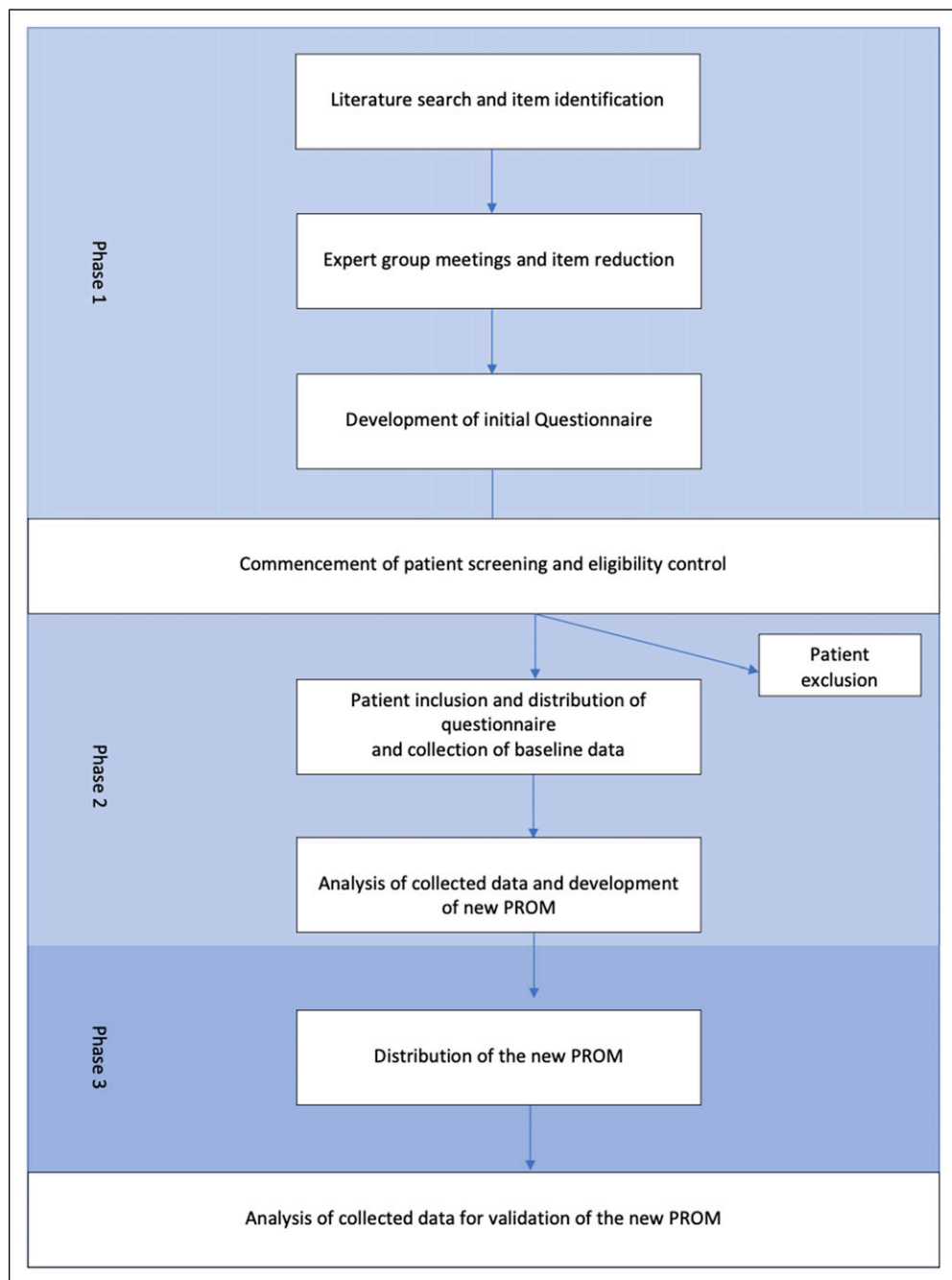
Figure 1. Flowchart of the study.

Table 1. Eligibility criteria.

Inclusion criteria
  - Patient diagnosed with Dupuytren disease by an orthopedic or hand surgeon
  - Age >18 years
  - Able to independently understand and consent to participation
Exclusion criteria
  - Cognitive impairment
  - Unable to complete the questionnaire independently
  - Patient unwilling to participate

Table 2. Collected baseline data.

Variable	Response options
- Age	[a]
- Occupation	[a]
- Gender	Male/Female
- Dominant hand	Right/Left
- Affected hand	Right/Left
- Previous treatment	Surgical fasciectomy/collagenase injection/PNF
- Bilateral disease	Yes/no
- Contracture	Digit 1, 2, 3, 4 or 5
- Smoking	Yes/no/previous
- Diabetes	Yes/no
- Heredity	Yes/no

PNF = percutaneous needle fasciotomy.
[a] = free-response question.

When the estimated sample size has been reached, responses will be entered into a database for analysis to develop a new PROM.

Phase 3 (validation sample). In the final phase of the study, the newly developed PROM will be administered to new patients in the validation sample. The same eligibility criteria as in the development sample will be used. Baseline data will be recorded and passive extension deficit of joints in the affected fingers will be measured as described in Phase 2. The collected data will be entered into a database for analysis. The newly developed PROM will be administered to patients both before and after treatment in this final phase.

Sample size

There is no universally accepted minimum sample size in PROMs development, but a sample of 300 respondents would be sufficient to meet the requirements for the types of statistical analyses that will be conducted.[21] This is the target sample size for both the development and validation samples.

Statistical analysis

Phase 1 (item selection). The goal is to develop a relatively short and unidimensional instrument. However, it is common practice to develop and trial a larger number of items than what is expected to be included in the final instrument. The recommendations regarding the "overdevelopment" of items range from twice to five times as many trialed items as final items.[23] We have previously successfully developed an instrument for measuring carpal tunnel syndrome (CTS) symptom severity with only 6 items,[24] but everyday activities are much less specific to Dupuytren disease than CTS symptoms are to CTS. Therefore, a larger instrument will probably be needed to capture the intended construct with high reliability and validity, and consequently we aim at a final instrument with around 20 items.

Phase 2 (development sample). A brief presentation of the statistical analyses and item reduction procedure for the development phase is presented in Figure 2. Initial item screening will be conducted. It is likely that some of the 85 listed issues are too rare to be included in the final instrument. When the number of observations in a category is low, parameter estimates will be imprecise and potentially unstable. Items that do not have at least 10 responses in each category of the rating scale will be removed[25] or, depending on the pattern of response frequencies, some categories may be collapsed.

Both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are common methods for analyzing the structure, or dimensionality, of a measurement instrument. The main conceptual difference between the two methods is whether there is a strong theory about how the items are related to each other. Common practice is to perform EFA on the try-out version of the instrument (in the development sample) and CFA on the final version of the instrument (in the validation sample).[23] The choice of methods for specific sub-procedures in EFA, such as factor extraction, factor rotation, and factor retention decisions, is to some extent dependent on the nature of the data, but common methods include principal components and principal axis factoring for factor extraction, varimax and promax for rotation, and parallel analysis for factor retention decisions. An evaluation criterion of 0.40 is used for the pattern coefficients to indicate association between the item and the factor. EFA and CFA analyses will be performed using Mplus).[26]

The items that are left after the EFA will be subject to IRT analyses. Polytomous Rasch models, specifically the partial credit model (PCM)[27] and the rating scale model (RSM)[28] will be used for estimating item and person parameters. The models are similar in that both include a "difficulty" parameter (which indicates whether respondents in general tend to endorse lower or higher response categories) and step parameters (the locations along the latent trait scale, relative to the item
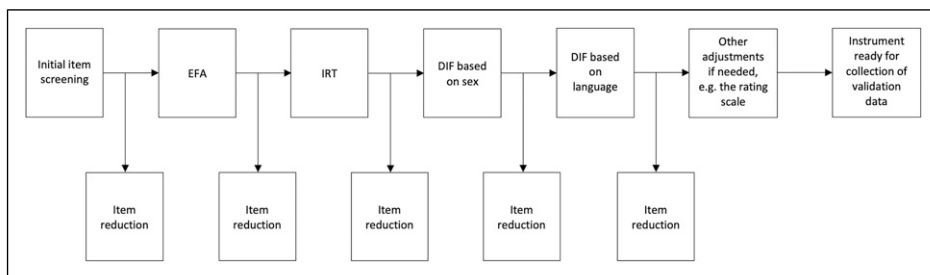

Figure 2. The planned statistical analyses and item reduction procedure for the development phase.

difficulty, at which categories k-1 and k are equally likely to be observed). In the PCM the step parameters are estimated for each item, whereas in the RSM they are constrained to be equal for all items. If all the items that are subjected to the IRT analyses have the same rating scale, the RSM will be the first-choice model. If this was not the case, the PCM will be the primary model. Should neither the RSM nor the PCM fit the data we will consider the generalized partial credit model (GPCM)[29] that in addition to the difficulty and step parameters also includes a discrimination parameter.

The IRT analyses will be executed using the standalone software ConQuest. Item fit will be examined by the infit and outfit measures, with values between 0.75 and 1.33 indicating acceptable fit.[30] The assumptions of local item independence and monotonicity underlying IRT will also be examined; local item independence by assessing item fit between each item pair using ConQuest's FIT procedure (with values exceeding 1.33 indicating local dependence), and monotonicity by visual inspection of the item observed scores across the latent trait (i.e., activity limitations). Respondents with more severe activity limitations having lower observed item scores than respondents with less severe activity limitations indicates a violation of the monotonicity assumption.

As part of the IRT analyses the instrument will be examined for DIF. An item exhibits DIF when respondents with the same level of the latent trait (e.g., functioning or symptoms) but from different groups (e.g., men/women) have different expected scores on that item. We will first investigate DIF with respect to sex, and then DIF with respect to language version (English/Swedish). The DIF values provided by ConQuest represent the differences between the item location estimates for the specific group and the item location estimates for all respondents. Values for DIF are considered both in relation to their standard errors (for statistical significance) and in relation to themselves (for practical significance). A ratio of 2 for the DIF value to its standard error indicates a significance level of approximately 0.05. Consequently, a ratio exceeding 2 indicates significant DIF. The magnitudes of DIF are evaluated according to recommended criteria,[31] where a logit difference (i.e., difference in item location) below 0.43 is considered "negligible", a value between 0.43 and 0.64 is considered "intermediate", and a value greater than 0.64 is considered "large".

Phase 3 (validation sample). The procedures and statistical analyses for instrument validation are essentially the same as for instrument development (Figure 3). The item screening follows the same procedure as in the analysis of the development sample.

The developed instrument is intended to be unidimensional. We will use CFA to test this by examining the fit between the data and a unidimensional measurement model. Fit in CFA is usually evaluated through several goodness-of-fit indices. We will use three of the most common indices,[32] namely the root mean squared error of approximation (RMSEA), the standardized residual mean
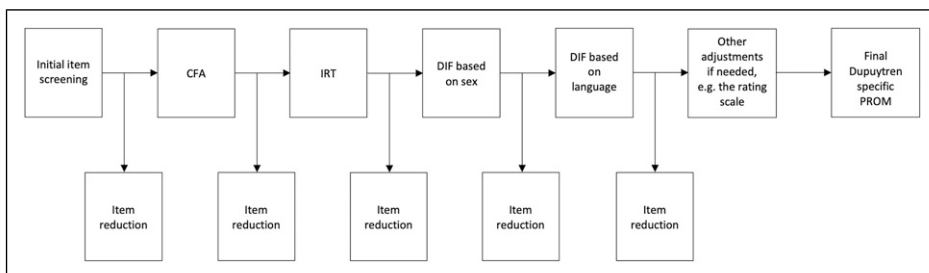

Figure 3. The planned procedures and statistical analyses for instrument validation.

square residual (SRMR), and the comparative fit index (CFI) with recommended cutoff values of 0.06, 0.08 and 0.95, respectively.[33] Within the CFA analysis we will also examine coefficient omega for scale reliability.[34]

The IRT analysis including DIF follows the same procedure as in the analysis of the development sample.

Measurement invariance (MI) analysis[35] is conducted to examine whether the instrument functions similarly across different groups of respondents. Although DIF and MI are related concepts, MI is a broader concept. Analysis of MI usually refers to a specific set of procedures within the framework of CFA and structural equation modelling (SEM). When an instrument exhibits MI this implies that it is possible to compare scores across groups. An analysis of MI is usually done in four steps, looking at a specific type of MI in each step: 1) configural invariance (invariance of model form), 2) metric invariance (equivalence of the item loadings on the factors), 3) scalar invariance (equivalence of item intercepts), and 4) residual invariance (equivalence of item residuals). Residual invariance is not a prerequisite for comparing scores between groups, and therefore this step is often omitted.[35]

Further analysis. Baseline data collected in the second and third phase will be analyzed to describe the included patients. Mean PED will be calculated for metacarpophalangeal (MCP) and proximal interphalangeal (PIP) joints respectively. The mean total (MCP + PIP) passive extension deficit (TPED) will also be calculated and used to analyze if certain items correlate more with increasing TPED. We will also perform further analysis regarding if there is a stronger correlation between specific items and MCP or PIP joint contractures respectively. The final version will be assessed for validity, including known-groups construct validity, responsiveness and determining the minimal clinically important change using appropriate methodology.[36] The TPED before and after treatment will also be used to analyze responsiveness of the final PROM.

Data management

The study will be led by a senior clinical researcher and an experienced research nurse at the principal study center (Department of Orthopedics, Hässleholm-Kristianstad Hospitals, Hässleholm, Sweden). All data will be stored at the research unit of the principal study center. All data will be coded and stored in a database accessible only to the research group.

Funding, ethics, and withdrawals

The study has been approved by the Swedish Ethical Review Authority (Ref: 2023-01,725-01 and 2023-05,885-02) and will be conducted in accordance with the Helsinki declaration (as revised in 2000). Before inclusion all patients will be given written and verbal information about the study by the examining surgeon or research nurse. Patients who accept participation will provide written consent. The study is supported by grants from Region Skåne and Guldbyxan Foundation.

Study status and time plan

Phase 1 of the study (creation of the initial PROM) began in January 2021 and has been completed. Phase 2 of the study commenced in June 2023 with ongoing patient recruitment (development sample) at all 3 participating centers. As of June 30, 2024, complete Phase 2 data have been collected from approximately 250 participants. It is expected that this process will be completed

during 2024. Data-analysis will then be conducted, and the final PROM will be constructed during early 2025. Phase 3 of the study will start with patient recruitment (validation sample) in 2025 and expected to be concluded within a year. The new Dupuytren-specific PROM will be ready for use in 2026.

Discussion

Dupuytren disease is a common disabling hand condition that often requires treatment. Different treatment methods ranging from major surgery to minimal-invasive procedures are currently used, even for the same level of disease severity. The lack of a reliable and valid measure that can accurately detect benefits that are relevant to patients with Dupuytren disease has substantially limited the ability to advise patients of the outcomes of different treatment methods allowing them to make well-informed treatment decisions. To fill this gap, we are developing a novel Dupuytren-specific PROM using modern measurement methodology, to be used in clinical setting as well as in research.

Currently, range of motion is the most common outcome used in Dupuytren disease.[37] It has previously been shown that the relationship between the degree of contracture and functional limitations vary between patients and some of the functional loss may depend on other factors than loss of finger extension.[6,8,16,38] To account for this, PROMs are important addition to clinical examination assuming the PROM is reliable, valid and responsive. However, development of a PROM for Dupuytren disease has proved to be challenging and no gold standard measure currently exists.[39]

The DASH and its shorter form the QuickDASH have been the most used PROMs in Dupuytren disease.[16] Although the DASH is a valid PROM for Dupuytren disease,[10] it includes items regarding pain, tingling and difficulties with sleep that are not typical for patients with Dupuytren disease and other hand, arm or shoulder conditions may influence the results. Previous research has shown that a larger finger joint contracture does not correlate with a worse DASH score.[40] Another perceived limitation of the DASH is that it is not possible to distinguish separate subscale scores.[39] The limitation of the QuickDASH can be seen in a recent randomized controlled trial that compared surgery, needle fasciotomy and collagenase injection and used the QuickDASH as one of the secondary outcomes.[41] In that trial, the baseline QuickDASH values were 18, 14, and 16, and the 3-months post-treatment values 13, 9 and 10, respectively. Considering that the authors used a value of 16 points as a minimal clinically important difference, it would obviously be almost impossible to find important between-group differences (unless an implausible large worsening occurs).[41]

High responsiveness is essential when using a PROM as the primary outcome measure. Two ongoing randomized controlled trials from the United Kingdom have used the patient evaluation measure (PEM) as a primary outcome.[42,43] Although a previous study found that severity of contracture correlates with the PEM score,[44] the responsiveness of the PEM in patients with Dupuytren disease is unclear and, as stated in the protocols for those two trials, the minimal clinical important difference has not been established. A recent systematic review of PROMs in Dupuytren disease did not include the PEM.[13]

The responsiveness of the URAM has varied between studies, with effect sizes ranging from 0.56 to 0.96.[13] In a systematic review, the authors reported general lack of what they described as sensitivity, with values of 52%–56%, suggesting that the URAM may not be able to detect all the disabilities experienced by Dupuytren patients.[13] A study showed that 55% of the activity limitations reported by patients were not included in the URAM, with the authors questioning the development process of the URAM and the choice to not include activities known to be experienced

by patients with Dupuytren disease, such as donning gloves.[16] Two studies comparing needle fasciotomy with collagenase in Sweden used the URAM as a secondary outcome,[45,46] but it is unclear what version was used because we could not find that a validated Swedish version of the URAM exists. In these studies, mean pre-treatment URAM scores (scale range 0 to 45) were 14 and 15 and median scores were 10 and 11 in the 2 groups, respectively, with score range of 0 to 33.[45,46] At 1 year, both studies reported mean URAM scores between 2 and 3. In a previous study, the minimal clinical important difference after surgical fasciectomy was calculated to be 10.5.[47] Thus, as also shown in other studies,[14,15] many patients with finger joint contractures have best possible scores on the URAM both before and after treatment.

Another issue common to all fixed-items PROMs is that the items may not identify individual patient's functional limitation. As an alternative to fixed-items PROMs individualized PROMs have been suggested to be used in Dupuytren disease.[16] One such individualized PROM is the Canadian Occupational Performance Measure (COPM) that has shown high responsiveness for Dupuytren disease.[17] However, its use is time and resource consuming and thus less usable in larger studies.

It has previously been established that patients with Dupuytren disease present with a variety of functional problems.[17,48] Thus, a strength in the design of this PROM is the choice to include approximately 900 items in the initial pool of items screened by the expert group. During the expert group sessions, we also chose to include a broad range of items in the trialed questionnaire to reduce the risk of bias. For instance, we chose to include questions regarding appearance, psychological aspects as well as pain. The items regarding pain in the DASH and QuickDASH as applied to Dupuytren disease have been debated and URAM does not include any items about pain.[39] While Dupuytren disease is not typically considered painful, pain may be present and in one previous study of 110 patients with Dupuytren disease 17 patients (with no comorbid painful condition) reported pain in their digit.[16] Pain is also an important factor when comparing early postoperative outcomes between different treatment modalities.

The possibility for patients to add specific activity limitations other than those listed in the questionnaire as well as patient input during the development process are further study strengths. Finally, as previous Dupuytren-specific PROMs were developed using CTT, the use of IRT methodology is another strength in the design of our study. Benefits of IRT methodology include independency between item statistics and the patients, possibility to target the scale to the intended population and the possibility to extract more specific instrument reliability data.[49] In general, use of IRT provides the unique ability to assess reliability at different levels of ability, the development of computerized adaptive tests to best balance between accuracy (reliability) and patient burden, and the development of efficient fixed short forms that generate scores on the same metric as the full item bank, with the ability to tailor the selected items to achieve high reliability at targeted ability levels.

Our aim is to develop a unidimensional measure primarily of activity limitations. Although the conventional belief has been that Dupuytren disease primarily affects hand function and rarely causes symptoms or impacts other health aspects, this has not been clearly established. Therefore, we chose to investigate this question as part of this research by including items about symptoms and other aspects. With the relatively large sample we expect to be able to investigate which and to what degree factors other than activity limitations are relevant to patients with Dupuytren disease. Similar to the development of other PROMs, when this new PROM has been developed it will undergo assessments of validity and responsiveness including determining the minimal clinically important difference using the appropriate methodology.

A possible limitation of our study is that involvement of only professionals working within the Swedish healthcare system during initial expert group meetings and initial item reduction may have influenced the choice of activities included in the questionnaire. Another limitation is that only two

patients recruited from one center provided early feedback on the questionnaire. The items were selected from a large and broad item pool to cover a wide range of activities including those that may be more specific to certain age ranges or sex group. Some of the original PROMs involved varying degrees of patient involvement in their development, including patients with Dupuytren disease. In our previous research on the COPM, 30 patients with Dupuytren Disease scheduled for treatment were interviewed by an occupational hand therapist (a member of the expert group in this study) about the type of activities they had difficulties performing and felt they were most important to them.[17] The patients interviewed mentioned 42 different unique activities with varying frequencies,[17] which was taken into consideration in the first phase of the current study. In the trialed questionnaire the participants are asked to list up to 3 activities they experience to be most problematic and important to them. Thus, we believe the large sample size, the free-response item, and the international multi-center setting lessen the impact of these limitations. Patients responding to the questionnaire prior to their doctor visit may be another potential risk of bias, as patients may exaggerate their activity limitations.

Conclusion

In this study protocol we describe a 3-phase process of developing a novel Dupuytren-specific patient-reported outcome measure using methodology based on item response theory. The recent advances in the treatment of Dupuytren disease and the ongoing research about various treatment methods require evaluation tools with higher accuracy than those currently available. A responsive, disease-specific patient-reported outcome measure can be used in clinical studies to better evaluate current and future treatments. When the final measure has been constructed and validated, this new Dupuytren-specific patient-reported outcome measure will advance clinical research on Dupuytren disease, improve our ability to assess the burden of this disease, and help patients to make well informed decisions about choice of treatment.

Author contribution

IA conceived the study. IA, DE and PEL contributed to study design and data analysis plan. IA, DE, JM, AL and JN contributed to data collection. DE drafted the manuscript with support from IA and PEL. All authors reviewed and approved the final manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Ethical statement

Ethical approval

The study has been approved by the Swedish Ethical Review Authority (Ref: 2023-01,725-01 and 2023-05,885-02). Participants provide written informed consent before enrolment in the study.

ORCID iD

David Eckerdal https://orcid.org/0000-0003-1564-6974

Supplemental Material

Supplemental material for this article is available online.

References

1. Nordenskjold J, Englund M, Zhou C, et al. Prevalence and incidence of doctor-diagnosed Dupuytren's disease: a population-based study. *J Hand Surg Eur* 2017; 42: 673–677.
2. Denkler K. Surgical complications associated with fasciectomy for Dupuytren's disease: a 20-year review of the English literature. *Eplasty* 2010; 10: e15.
3. Nanchahal J, Ball C, Rombach I, et al. Anti-tumour necrosis factor therapy for early-stage Dupuytren's disease (RIDD): a phase 2b, randomised, double-blind, placebo-controlled trial. *Lancet Rheumatol* 2022; 4: E407–e416.
4. Zhao JZ, Hadley S, Floyd E, et al. The Impact of collagenase clostridium histolyticum introduction on Dupuytren treatment patterns in the United States. *J Hand Surg Am* 2016; 41: 963–968.
5. Socialstyrelsen. Statistik databas för operationer. https://sdb.socialstyrelsen.se/if_ope/resultat.aspx (2024), (accessed 19 March 2024).
6. Ball C, Pratt AL and Nanchahal J. Optimal functional outcome measures for assessing treatment for Dupuytren's disease: a systematic review and recommendations for future practice. *BMC Muscoskel Disord* 2013; 14: 131.
7. Nordenskjöld J, Brodén S and Atroshi I. Examiners' influence on the measured active and passive extension deficit in finger joints affected by Dupuytren disease. *BMC Med Res Methodol* 2018; 18: 120.
8. Jerosch-Herold C, Shepstone L, Chojnowski A, et al. Severity of contracture and self-reported disability in patients with Dupuytren's contracture referred for surgery. *J Hand Ther* 2011; 24: 6–10.
9. Gummesson C, Ward MM and Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *BMC Muscoskel Disord* 2006; 7: 44.
10. Budd HR, Larson D, Chojnowski A, et al. The QuickDASH score: a patient-reported outcome measure for Dupuytren's surgery. *J Hand Ther* 2011; 24: 15–20.
11. Beaudreuil J, Allard A, Zerkak D, et al. Unité Rhumatologique des Affections de la Main (URAM) scale: development and validation of a tool to assess Dupuytren's disease-specific disability. *Arthritis Care Res* 2011; 63: 1448–1455.
12. Mohan A, Vadher J, Ismail H, et al. The Southampton Dupuytren's scoring scheme. *J Plast Surg Hand Surg* 2014; 48: 28–33.

13. Bradet-Levesque I, Audet J, Roy JS, et al. Measuring functional outcome in Dupuytren's disease: a systematic review of patient-reported outcome measures. *J Hand Ther* 2022; 35: 613–627.

14. Broekstra DC, Van den Heuvel ER, Lanting R, et al. Measurement properties of the Dutch Unité Rhumatologique des Affections de la Main and its ability to measure change due to Dupuytren's disease progression compared with the Michigan Hand Outcomes Questionnaire. *J Hand Surg Eur* 2018; 43: 855–863.

15. Hensler S, Wehrli M, Herren DB, et al. Measurement properties of the German Unité Rhumatologique des Affections de la Main (URAM) scale in patients treated for Dupuytren's disease. *Hand Surg Rehabil* 2020; 39: 568–574.

16. Rodrigues JN, Zhang W, Scammell BE, et al. What patients want from the treatment of Dupuytren's disease--is the Unité Rhumatologique des Affections de la Main (URAM) scale relevant? *J Hand Surg Eur* 2015; 40: 150–154.

17. Lauritzson A, Eckerdal D and Atroshi I. Responsiveness of the patient-specific Canadian occupational performance measure and a fixed-items activity limitations measure in patients with dupuytren disease. *J Patient Rep Outcomes* 2023; 7: 38.

18. Law M, Baptiste S, McColl M, et al. The Canadian occupational performance measure: an outcome measure for occupational therapy. *Can J Occup Ther* 1990; 57: 82–87.

19. Brodke DJ, Hung M and Bozic KJ. Item response theory and computerized adaptive testing for orthopaedic outcomes measures. *J Am Acad Orthop Surg* 2016; 24: 750–754.

20. Hays RD, Morales LS and Reise SP. Item response theory and health outcomes measurement in the 21st century. *Med Care* 2000; 38: 28–42.

21. Lyrén PE and Atroshi I. Using item response theory improved responsiveness of patient-reported outcomes measures in carpal tunnel syndrome. *J Clin Epidemiol* 2012; 65: 325–334.

22. Forget NJ, Higgins J, Rivard M, et al. Going beyond activity and participation: development of the DIF-CHUM-A patient-reported outcome measure for individuals with Dupuytren's contracture. *J Hand Ther* 2020; 33: 305–313.

23. Boateng GO, Neilands TB, Frongillo EA, et al. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Public Health* 2018; 6: 149.

24. Atroshi I, Lyrén PE and Gummesson C. The 6-item CTS symptoms scale: a brief outcomes measure for carpal tunnel syndrome. *Qual Life Res* 2009; 18: 347–358.

25. Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002; 3: 85–106.

26. Muthén LK and Muthén BO. *Mplus user's guide*. 8th ed. Los Angeles, CA: Muthén and Muthén, 2017.

27. Masters GN. A rasch model for partial credit scoring. *Psychometrika* 1982; 47: 149–174.

28. Andrich D. A rating formulation for ordered response categories. *Psychometrika* 1978; 43: 561–573.

29. Muraki E. A generalized partial credit model: application of an EM algorithm. *Appl Psychol Meas* 1992; 16(2): 159–176.

30. Wu M, Adams R and Wilson M, ConQuest [computer program]. ConQuest: Multi-Aspect Test Software, [computer program] Camberwell: Australian Council for Educational Research. 1997.

31. Wilson M. *Constructing measures: an item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates, 2005.

32. Goretzko D, Siemund K and Sterner P. Evaluating model fit of measurement models in confirmatory factor analysis. *Educ Psychol Meas* 2023; 84(1): 123–144.

33. Hu LT and Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 1999; 6: 1–55.

34. McDonald RP. *Test theory: a unified approach*. Mahwah, NJ: Erlbaum, 1999.

35. Putnick DL and Bornstein MH. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev Rev* 2016; 41: 71–90.

36. Rosales RS and Atroshi I. The methodological requirements for clinical examination and patient-reported outcomes, and how to test them. *J Hand Surg Eur* 2020; 45: 12–18.
37. Karpinski M, Moltaji S, Baxter C, et al. A systematic review identifying outcomes and outcome measures in Dupuytren's disease research. *J Hand Surg Eur* 2020; 45: 513–520.
38. Zyluk A and Jagielski W. The effect of the severity of the Dupuytren's contracture on the function of the hand before and after surgery. *J Hand Surg Eur* 2007; 32: 326–329.
39. Stirling PHC, McEachan JE, Rodrigues JN, et al. QuickDASH questionnaire items behave as 2 distinct subscales rather than one scale in Dupuytren's disease. *J Hand Ther* 2023; 36: 228–233.
40. Degreef I, Vererfve PB and De Smet L. Effect of severity of Dupuytren contracture on disability. *Scand J Plast ReConstr Surg Hand Surg* 2009; 43: 41–42.
41. Räisänen MP, Leppänen OV, Soikkeli J, et al. Surgery, needle fasciotomy, or collagenase injection for Dupuytren contracture: a randomized controlled trial. *Ann Intern Med* 2024; 177(3): 280–290.
42. Dias J, Arundel C, Tharmanathan P, et al. Dupuytren's interventions surgery versus collagenase (DISC) trial: study protocol for a pragmatic, two-arm parallel-group, non-inferiority randomised controlled trial. *Trials* 2021; 22(1): 671.
43. Harrison S, Ogollah R, Hollingworth W, et al. Needle fasciotomy versus limited fasciectomy for the treatment of Dupuytren's contractures of the fingers (Hand-2): study protocol for a randomised controlled trial. *Trials* 2024; 25(1): 398.
44. Dias JJ and Braybrooke J. Dupuytren's contracture: an audit of the outcomes of surgery. *J Hand Surg Br* 2006; 31(5): 514–521.
45. Strömberg J, Ibsen-Sörensen A and Fridén J. Comparison of treatment outcome after collagenase and needle fasciotomy for Dupuytren contracture: a randomized, single-blinded, clinical trial with a 1-year follow-up. *J Hand surg Am* 2016; 41: 873–880.
46. Scherman P, Jenmalm P and Dahlin LB. One-year results of needle fasciotomy and collagenase injection in treatment of Dupuytren's contracture: a two-centre prospective randomized clinical trial. *J Hand Surg Eur* 2016; 41: 577–582.
47. Rodrigues JN, Zhang W, Scammell BE, et al. Recovery, responsiveness and interpretability of patient-reported outcome measures after surgery for Dupuytren's disease. *J Hand Surg Eur* 2017; 42: 301–309.
48. Engstrand C, Borén L and Liedberg GM. Evaluation of activity limitation and digital extension in Dupuytren's contracture three months after fasciectomy and hand therapy interventions. *J Hand Ther* 2009; 22: 21–26.
49. Åström E, Sundström AE and Lyrén PE. Examining the psychometric properties of the Swedish version of clinical outcomes in routine evaluation-outcome measure (CORE-OM) in a clinical sample using classical test theory and item response theory. *Clin Psychol Psychother* 2023; 30: 398–409.