

64708_Auto_Edited.docx

Name of Journal: *World Journal of Hepatology*

Manuscript NO: 64708

Manuscript Type: MINIREVIEWS

Saving time and effort: Best practice for adapting existing patient-reported outcome measures in hepatology

PROMs in hepatology

Abstract

It is increasingly recognised that collecting Patient Reported Outcome Measures (PROMs) data is an important part of healthcare and should be considered alongside traditional clinical assessments. As part of a more holistic view of healthcare provision, there has been an increased drive to implement PROM collection as part of routine clinical care in Hepatology. This drive has resulted in an increase in the number of PROMs currently developed to be used in various liver conditions. However, the development and validation of a new PROM is time-consuming and costly. Therefore, before deciding to develop a new PROM, researchers should consider identifying existing PROMs to assess their appropriateness and, if necessary, make adaptations to existing PROMs to ensure their rigour when used with the target population. Little is written in the literature on how to identify and adapt the existing PROMs in Hepatology. This article aims to provide a summary of the current literature and guidance regarding identifying and adapting existing PROMs in clinical practice.

Key Words: Patient-Reported Outcome Measures, PROMs, adaptation, content validation

Alrubaiy L, Hutchings HA, Hughes S, Dobbs T. Saving time and effort: Best practice for adapting existing patient-reported outcome measures in hepatology. *World J Hepatol* 2022; In press

Core Tip: This paper aims to illustrate ways of identifying existing Patient-Reported Outcome Measures (PROMs) and outlines key considerations and good practice with respect to their adaptation in Hepatology.

INTRODUCTION

What are Patient Reported Outcome Measures (PROMs)?

Patients are treated by healthcare providers with the primary goal of improving their health and well-being. Historically this improvement in health has been judged by improvement in biochemical, histological, radiological or clinical assessments. This approach does not always correlate with improvement from the patient perspective. From a patient perspective improving health is reflected in the documentation of their symptoms and experience of healthcare provision, which are more appropriately collected directly from the patient¹. With a move towards shared-decision making and patient-centred care, there is growing recognition within the healthcare community of the importance of the patient perspective and the need to consider patient reported outcomes (PRO) as a key component of a holistic approach to patient care.

¹ The U.S. Food and Drug Administration (FDA) defines a PRO as “any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else”². The European Medicines Agency state that “Any outcome evaluated directly by the patient himself and based on the patient’s perception of a disease and its treatment(s) is called a patient-reported outcome (PRO)”³. PROMs can be broadly classified as generic or disease-specific instruments. Generic PROMs assess general aspects of health and can be applied across multiple conditions. Disease-specific PROMs, on the other hand,

assess specific aspects that are related to a particular condition. PROMs are designed to measure aspects of health that can neither be directly observed or are not feasible to observe⁴. Broadly speaking, collection of PROMs from the patient can be classified into three main categories based on the outcomes measured:

Health status and quality of life: patients' health and well-being as indicated by patient report

Patient satisfaction: patient-reported satisfaction with their medical treatment or care

Resource use: patients' reported use of health services and resources

Patient knowledge questionnaires: patients' understanding of medical conditions and the treatment

What is currently driving the use of PROMs?

PROMs were initially developed for research use and many regulatory authorities such as the European Medicines Agency (EMA) and the United States (US) Food and Drug Administration (FDA) advocate their use^{2,5,6}. Recent consensus guidance also recommends the inclusion of PROMs in clinical trial designs⁷. The collection of PROMs aligns well with the increased drive within healthcare organisations for value based healthcare, whereby organisations aim to achieve the best possible outcomes for patients with the available resources⁸. As more clinicians recognise the benefit of collecting PROMs in addition to measuring clinical outcomes, PROMs have seen an increased use in routine clinical practice⁹.

As a consequence of the drive to collect PROMs, there has also been an increase in the number of PROMs developed, validated, and used. The King's Fund report reflects on this as "a growing recognition throughout the world that the patient's perspective is highly relevant to efforts to improve the quality and effectiveness of health care" and that PROMs are likely to become "a key part of how all health care is funded, provided, and managed"¹⁰. This has been illustrated in the United Kingdom (UK) when, in 2009, the UK Government implemented the routine collection of PROMs in England for four routine elective surgical procedures—hip and knee replacement, groin hernia repair,

and varicose vein surgery (<https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms>), in order to compare performance between providers. It is likely that routine PROMs collection will be extended to more conditions in the future.

Methods- Literature search to identify best practice for the adaptation of existing PROMs

In order to identify relevant literature regarding best practice and guidance for the adaptation of existing PROMs we undertook a scoping review of the literature. This scoping review aimed to explore the extent of the literature within the PROMs field regarding best practice/guidance for PROM adaptation without describing findings in detail ¹².

We undertook a review of the literature to identify key papers/guidelines for the adaptation of existing PROMs. We searched PubMed and the Cochrane database (<https://www.cochrane.org/>). In order to limit the search, we searched for literature in the English language published within the last 10 years. Reference lists of relevant identified publications were also hand searched to identify further relevant literature. We also undertook a Google™ search to identify relevant publications.

Details of the search strategy are presented in Table 1. The searches were conducted on 17 February 2020. The inclusion and exclusion criteria are listed in Table 2.

We carried out an initial title screening, then abstract screening to identify relevant papers that fitted the inclusion criteria, which we then reviewed fully. We identified specific themes related to the adaptation of existing PROMs which we regarded as recommendations/good practice and have structured the paper according to these identified themes.

Findings

Appendix 1 illustrates the publications identified as part of the scoping review of the literature. The guidance identified within these publications is organised under the specific headings of: defining the requirements of a PROM, identifying and appraising existing tools, adapting existing PROMs, issues of content validity and getting the right people involved.

Defining the requirements of a PROM

1 In order to provide meaningful information, PROMs need to be appropriately developed and validated according to robust criteria. The psychometric validation of PROMs can be complex and time-consuming and requires evidence of numerous facets including validity, reliability and responsiveness¹³⁻¹⁴. Given the growth in the number of available PROMs, even within the same condition, the old adage “don’t reinvent the wheel” should be the first principle applied before taking the decision to embark on the development of a new PROM. Consequently, to enable researchers to appraise the quality of existing measures with the aim of ascertaining whether a new measure is needed, researchers must first establish a clear definition of what is required of the PROM.

The requirements of the PROM need to be identified at the outset ¹⁵. Consideration should include what the PROM aims to measure, whether the PROM should be generic or disease specific, the clinical condition of interest, the specific population for which the PROM will be applied and whether it will be used as part of routine clinical care or research. These factors will help to determine whether existing PROMs are suitable or can be adapted. A useful overview and starting point for deliberations is provided by Wees¹⁶.

A generic PROM may allow comparison of patient outcomes across different conditions, however it will have less focus on specific symptoms relating to a

condition. A disease-specific PROM will have a more defined focus on the condition itself and will be more sensitive to changes in the condition over time and its associated symptoms but may be longer and therefore the burden to the patient may be greater. If a disease-specific PROM is required, one needs to define the specific population. For example, a PROM developed to measure pruritus in Primary Biliary Cholangitis (PBC) may not be suitable for measuring pruritus in Intrahepatic Cholestasis of Pregnancy (ICP) as the patient experience of dyspnea may differ across these two clinical conditions^{2,16}. It is also important to consider whether the PROM will be used within a routine clinical setting or a research setting¹⁷. In a clinic setting where time may be limited, the burden to the patient and the feasibility of completing the PROM need consideration. Within a research setting, time may not be as limited and longer, more detailed PROMs can be considered¹⁷. The proposed method of administration of the PROM is also important and authors planning on using a PROM should ensure that it has been appropriately validated for their proposed administration method². Issues such as respondent and administrator burden- length, formatting, font size, instructions, privacy, literacy levels *etc.* also need to be considered². Figure 1 provides an overview of the first steps required before choosing to develop a new PROM.

Luckett *et al*, have provided useful principles to consider when selecting a PROM¹⁶:

Selection of PROMs should be considered early during study design- selection should be driven by the research objectives, samples, treatment and available resources

3 For the primary outcome, choose as 'proximal' a PROM as will add to knowledge and inform practice- 'proximal' (symptoms) vs 'distal' (overall quality of life)

Identify candidate PROMs primarily on the basis of scaling and content- which items/scales offer best coverage of the impacts of interest and which aspects of score distribution will be most meaningful to consider?

Appraise the reliability, validity and 'track record' of candidate PROMs- look beyond articles that focus on evidence of validity and reliability

3 Look ahead to practical considerations- patient and staff burden, methods of administration, cost, availability of translated versions, guidelines for scoring and interpretation

3 Take a minimalist approach to ad hoc items- where content is similar, PROMs with proven psychometric properties are preferable to ad hoc measures developed by the researcher²¹

The EAPC ¹⁷ have provided guidance regarding the selection and measurement of PROMs in palliative care. This guidance reflects on the content of the PROM and the feasibility of its collection within the clinical environment. Some of the points made are particularly useful when considering adapting existing measures and can be usefully applied to other conditions:

16 Use PROMs that have been validated with relevant populations and make sure these are sufficiently brief and straightforward

Use multi-dimensional measures

6 Use measures that have sound psychometric properties

Use measures that are suited to the clinical task being delivered and are also suited to the aims of your clinical work and the population you work with

Use valid and reliable measures in research that are relevant to the research question and consider patient burden when using measures^{17,18}.

Identifying and appraising existing tools

Once one has defined the scope of the PROM, possible candidate measures can be identified. This process will determine whether there is a need for a new PROM. It will also allow for the identification of PROMs that could, be adapted, shortened, translated or expanded.

Given the large number of available PROMs, there are several ways in which possible candidate PROMs may be identified. Identifying systematic reviews of PROMs for a particular clinical area may prove to be particularly fruitful as good reviews will assess

the methodological quality ¹³⁻¹⁶ of the PROMs identified and provide a summary of the PROMs that offer the most promise. In addition to undertaking literature reviews, there are also databases and online resources that can be searched to identify existing PROMs. Some of these resources are generic and cover many conditions, whilst others provide a resource for disease-specific PROMs. Table 3 provides some examples of resources that can be used to identify candidate PROMs for adaptation.

If these strategies do not identify any PROMs, conducting a new systematic review may uncover PROMs for consideration or items/questions in existing PROMs that could be included in the development of a new PROM¹⁴. Prinsen *et al*, have formulated a useful ten step process for conducting such systematic reviews of PROMs ^{13,14}. Such a systematic review should be conducted in accordance with the guidelines outlined by internationally recognised ²⁵ COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN). COSMIN provide detailed information and ²⁴ tools to aid this process on their website (<https://www.cosmin.nl/tools/guideline-conducting-systematic-review-outcome-measures/>). This will ensure that the methodological rigor of the PROMs identified is appropriately appraised.

Once PROMs have been identified, the tools should be reviewed for their content and appropriateness for the desired application ^{13,14-22}. This process will also help to identify relevant questions/items that could be used to develop a new PROM or adapt an existing one.

Adapting existing PROMs

Researchers need to consider the existing PROM literature to ⁴ determine whether an adequate instrument exists to assess and measure the concepts of interest. If no PROM exists, a new PROM can be developed or in some situations, a PROM can be adapted by modifying an existing instrument.².

Examples of instrument modifications include: 1) making minor cultural/Language adaptations within the same source language; 2) undertaking a cross-cultural adaptation that includes translation into a different language; 3) including additional items/questions; and 4) shortening the original instrument. Such PROM modification may be necessary to enable the PROM to be used with a different population or with a different population age group (for example, modification of an adult PROM for a paediatric population), to facilitate its use in a different language, for use in a different disease stage or treatment (for example cancer stage, or for a newly diagnosed condition rather than a pre-existing condition), or to reduce patient burden.

The FDA state that when a PROM is modified, evidence of adequacy for its new intended use should be provided and that “additional qualitative work may be adequate” to test such modifications². Such changes include:

4 Changing an instrument from paper to electronic format

Changing the application to a different setting, population or condition

Changing the order of items, item wording, response options, or recall period or deleting portions of a questionnaire

Changing the instructions or the placement of instructions within the PROM.

Snyder *et al*, outline some requirements to revalidate a PROM when changes such as these are made to an existing PROM ²¹.

The search for PROMs may identify existing instruments that have proven validity for the population being studied and can be applied without requiring any adaptation. Alternatively, a PROM may be identified that appears appropriate but requires modification. Before engaging in any adaptation, it is important first to contact the PROM developer/copyright holder to ask for permission to make changes to the original PROM. Anfray *et al*, have recently published guidance from the International Society for Quality of Life Research (ISOQOL) Translation and Cultural Adaptation

Special Interest Group (TCA-SIG) regarding copyright of PROMs ²². Failure to gain appropriate permissions for use and adaptation may result in legal challenges due to breaches in copyright. The authors present recommendations to prevent future conflict that includes:

- protecting the copyright of the original PROM
- writing a contract
- taking care when publishing
- establishing rules
- making the copyright notice visible
- maintaining copyright of the PROM and any derivatives with the original author
- centralising distribution
- getting legal counsel
- clarifying the copyright situation with respect to legacy PROMs.

It is therefore prudent for researchers considering the adaptation (including the translation of existing PROMs) to identify and obtain agreement from the copyright holder prior to any adaptation²².

Cross cultural adaptation

If a suitable PROM is identified and has appropriate content validity for the population of interest, but was developed and validated in a different language, cross cultural adaptation represents an efficient way of adapting an existing PROM. Cross-cultural adaptation manages language translation and cultural adaptation issues with the aim of ensuring a PROM is sensitive to the linguistic and cultural needs of the target population ²². A PROM that has undergone rigorous cross-cultural adaptation is suitable for use in multinational and multicultural studies.

It is important that any cross-cultural adaptations of PROMs are undertaken rigorously. Guidance regarding the process of cross-cultural adaptation has been

described in a wide range of publications ²²⁻²⁵. The lack of 'gold standard' guidance for cross-cultural validation prompted the Patient Reported Outcome (PRO) Consortium, to update and develop further guidelines for best practice in the translation process ³⁶. These guidelines are based on the ISPOR Task Force guidelines, updated with greater detail through a further consensus process ²².

¹⁸ The aim of cross-cultural adaptation is to provide equivalence between the source PROM and the adapted version. Equivalence has many definitions; however, most current guidelines follow the universalist approach proposed by Herdman and colleagues (1997) ²⁷, which gives consideration to the influence of culture on how people respond to any given item on a questionnaire. Questions therefore not only require linguistic translation, but they must also be adapted to fit ⁹ culturally to the target country ²⁶. For example, a question about difficulty using a fork in eating may not be applicable in a country where a fork is not used in eating ²⁸. Equivalence can be divided into five categories plus a summary category ^{22,28} (see Table 4), and this has formed the basis of many guidelines for the cross-cultural adaptation of outcome measures.

Ultimately, all of the available guidelines are broadly based on a core set of principles that need to be considered when cross-culturally adapting an existing PROM:

Preparation. The initial stage of the process is to identify the team that will be responsible for the work, identify suitable translators and gain permission from the original instrument design team to carry out a translation process

Forward Translation. Translation of the original language version into the new, target language. It is considered best practice for this to be performed at least twice by different translators from the target country

Reconciliation. Comparison of multiple forward translations and merging them into one translated version

Back Translation. The newly translated version is translated back into the original source language

Back Translation review. The back-translated version is compared to the original version to assess for equivalence in text and meaning

Harmonisation. All translated versions are reviewed for consistency in language and conceptual meaning

Proofreading. All copies of the questionnaire are proofread to remove mistakes

Cognitive interviewing. The newly translated questionnaire is piloted on a minimum of five people in the target population. Cognitive interviews are performed to identify problems with the questionnaire, difficulties in understanding and meaning of items and any other concerns

Cognitive interview review. Results of the cognitive interviews are reviewed and changes made to the questionnaire if required

Final review and publication. The final version of the translated questionnaire is agreed upon and published for use in the target population

Cross-cultural validation. Following the production of a culturally and linguistically valid and similar version of the original questionnaire, the adapted questionnaire must then be psychometrically validated against internationally recognised criteria ^{28,29}.

18

The degree of cross-cultural adaptation required varies depending on the proposed use of the adapted PROM. The intended use of the PROM may influence the number of steps of the above that require completion ²⁹. Table 5 illustrates five different scenarios where differing adaptation needs are required ^{25,29}. These range from a situation in which no adaptation is required (i.e., the questionnaire used in the same population, in the same culture and language as originally designed), to full translation and cross-cultural adaptation (i.e., where the questionnaire is to be used in a different country and language).

Adding to existing PROMs

If an existing PROM is identified as largely meeting the requirement for the population of interest but following patient and expert consultation and/or exploration of the

literature it is perceived to be missing in one or more key areas, there is the potential to adapt the PROM by adding new questions/items. There are various ways in which items can be sourced ^{17,28}:

By asking patients. Patients can be asked to identify additional items and domains that do not exist in the current version of the PROM. Patients are essential to item generation, ensuring item content is both relevant and provides full coverage of the target construct. Qualitative methods such as patient focus groups, interviews and surveys are useful for generating potential new items ³⁰⁻³².

By evaluating the PROMs identified as a result of reviewing the literature or online resources. This can be an efficient way to generate new items. There are benefits to sourcing items in this way, most notably that there are likely to be a limited number of ways to ask questions about a specific problem such as abdominal pain, vomiting, etc. Moreover, items in existing PROMs have been repeatedly used and validated in many studies and trials.

By identifying possible items from clinical observations. These items can be derived by clinicians based on their experience.

By asking experts. This is a commonly used approach to generate new items. Similar methods (for example interviews, focus groups and surveys) to those used with patients can be used for gathering information about possible items for inclusion. Although useful for generating items, expert involvement should be used in tandem with other methods and should not be used in place of patient input.

By utilising item banks. Item banks are a source of validated items that can be added to existing PROMs. One such item bank, the Patient-Reported Outcomes Measurement Information System (PROMISTM) initiative was established in 2004, with the main goal of developing and evaluating, for the clinical research community, a set of publicly available, efficient and flexible measurements of PROs ²³. PROMISTM, (<http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis/List-of-adult-measures>) ² provides item banks that offer the potential for PRO measurement that is efficient (minimizes item number

without compromising reliability), flexible (enables optional use of interchangeable items), and precise (has minimal error in estimate) measurement of commonly-studied PROs³³. The PROMIS group has developed and tested several hundred items measuring 11 health domains³³. These core PROMIS domains reflect common, generic symptoms and experiences that are likely to apply to people in a variety of contexts or with a variety of diseases³³. With additional validation, these banks may provide a common metric of represented constructs across a range of patient groups, thereby reducing the large number of different measures currently used in research and allowing researchers to compare these constructs across patient groups in different studies³³.

Shortening of existing PROMs

Although many single-item and short-form symptom measures exist, one reason for adapting an existing PROM is to shorten it and reduce the number of items included in it. This can result in reduced patient burden and facilitate the use of a PROM as part of routine clinical care. As with other aspects of adaptation, it is essential to ensure that a shortened PROM is comprehensible to patients, includes all the relevant items and is fit for purpose. Like cross-cultural adaptation and adding existing items to a PROM, shortening will require further psychometric testing according to recognised criteria³⁰.

Issues of content validity

Where any adaptation is planned, the PROM will still need to show evidence that it is 'fit for purpose' with the intended population³³. In 2007, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Health Science Policy Council recommended that an ISPOR Patient-Reported Outcomes taskforce on the use of existing instruments and their modifications be established. This resulted in the publication in 2009 of their report detailing good research practices¹⁶. A major aspect of this report related to the content validity of PROMs, and stated "evidence of content validity should be obtained from an analysis of the relationship between the

instrument's content and the construct it intends to measure"¹⁶. This report highlights the key issues relating to content validity issues that should be considered when selecting and modifying existing PROMs ¹⁶:

Name and define the concept

Target population and end point

Identify candidate PROMs

7 Identify or formulate a conceptual framework for the PROM

Assemble and evaluate information on development methods- elicitation focus groups and interviews; cognitive interviews; transcripts

Conduct any needed qualitative work

Assess adequacy of content validity for purpose

Determine the need for modifications or new PROM development.

This guidance has since been updated to include further recommendations from the ISPOR good practice task force ³⁴. Additional guidance regarding content validity and its consideration with respect to PROM development and adaptation have also been published and includes best practices for undertaking qualitative research to explore content validity, including differences between establishing content validity for new measures compared with existing measures ^{14, 35}.

Assessment of PROM content is an important process when adapting an existing PROM and this should involve engagement with, most importantly, patients and also clinicians.

Getting the right people involved

Having identified a candidate PROM for adaptation it is important to ensure that it is appropriate for the patient population being studied. This is particularly important to undertake if the PROM is being adapted for use with a new clinical population. Pre-

testing the PROM with patients, clinicians, and subject-matter experts will provide evidence of the PROM's content validity and help to ensure that any problems are rectified prior to applying the PROM in a large-scale study or implementing the instrument in routine clinical practice.

Getting patients involved

¹⁵ In 2009 FDA guidance suggested that an important first step in establishing that a measure is fit for purpose is to develop a conceptual framework for the PROM and generate relevant items on the basis of direct input from patients with the clinical disease^{2,34, 35}.

Recent guidance^{36,37} highlights the various roles that patients and patient advocates can play in PROM studies. These include:

¹³ PROM design and selection- bringing knowledge of the disease, symptoms and attributes of care with the greatest impact on patients' lives

²² PROM implementation and administration- the patient can bring insights based on their experience to guide practical decisions around PROM administration and implementation

Linguistic and cultural input- patients can contribute to the language used in the PROM to ensure it is straightforward and understandable to patients.

Guidance regarding how the patients can be recruited to PROM studies, how to engage with patients, defining the role, provision of training and remuneration³⁸ has also been provided.

In addition, a framework for fully incorporating public involvement (PI) into PROMs has recently been published³⁸ which illustrates the extent to which patients can be involved in the adaptation process (see Table 6).

Existing measures can be reviewed to ensure they match the domains of interest and if further modification may be required¹⁶. Recent research that explored the level of involvement of patients in the development of PROMs has concluded that what patients consider important can differ from what health-professionals regard as important^{30,31}. Content validity is often cited as a PROM's most important measurement property as unless the PROM can be shown to be measuring the construct of interest from the patient perspective, all other measurement properties may be considered inconsequential¹³. This highlights the importance of engaging with patients as part of the PROM adaptation process.

A variety of qualitative methods can be used by researchers to engage with patients with the aim of maximising a candidate PROM's content validity (relevance and comprehensiveness) and to pre-test an adapted PROM for comprehensibility and acceptability of instructions to respondents, its items and response format(s). In a recent study examining the developers' perspective of including patients, the methods used were interviews and/or focus groups, cognitive interviews and feedback questionnaires^{30,31}. Maitland and Presser advocate a diverse range of methods, both qualitative and quantitative, for appraising the quality of PROM items and the ability of the items to generate reliable and valid responses³⁹.

Interview and focus groups are often used to gain insight into the experiences of the target population in relation to the construct of interest and, therefore can be used to generate content for new or additional questions. Cognitive interviews, on the other hand, are normally used to refine item candidates and their response scales^{55,56,57}. Cognitive interviews capture problems with the cognitive processes associated with item response⁴⁰, thereby enabling the developers to evaluate the relevancy, comprehensiveness, comprehensibility and acceptability of the instrument's items and response scales.

Feedback questionnaires can also provide patients insight regarding their experience of using a health status questionnaire. The QQ10 is one such validated, self-completed questionnaire. It is made up of 10-items scored using a 5-point Likert scale (0 = strongly disagree to 4 = strongly agree) covering two factors, “value” and “burden”. It contains specific items developed to assess a PROM’s content validity (i.e., relevance, comprehensiveness) from the patient perspective^{41,42}.

Getting experts involved

The assessment of a candidate PROM from the expert clinical, researcher and academic perspective is also important. This can be achieved *via* focus groups and interviews, by questionnaire survey methods or by employing expert review panels³⁴. Ideally, these panels should include clinicians with experience of treating the defined population, PROMs methodologists and researchers. The COSMIN standards recommend a minimum sample size of seven professionals for studies evaluating a PROM’s content validity¹⁴

Experts can also be utilised to calculate Content Validity Indices (CVI) based on ratings of item relevance. A minimum of three experts is recommended for the purposes of calculating a CVI⁴³. A CVI is a consensus indicator of the content validity of an item or scale⁴⁴. It represents the proportion of reviewers who agree that an item is content valid, adjusted for chance agreement. If all reviewers are in agreement, the CVI value for an item (I-CVI) will be 1.00.

Employing different psychometric methods to PROMs development and adaptation

Traditionally the development and psychometric evaluation of PROMs has been based on classical test theory (CTT). CTT is probably still the most commonly applied method in validation studies^{20,45}. CTT assumes that the expected value of all the random error will equal zero⁴⁶. There are, however, some disadvantages with CTT, such as sample dependency. This is where the item and scale statistics can only in theory, apply to the

specific group of patients who took the test and as such further validation is required for a different population²⁹. There is also the assumption of item equivalence, where it is assumed that all items contribute equally to the final test score and no item weightings are applied^{46,47}.

As a result of the disadvantages of CTT, ²⁷ modern psychometric methods of item response theory (IRT) and Rasch measurement theory (RMT) have been developed⁴⁸. Rather than considering the questionnaire as a whole, as in the case of CTT, these methods allow analysis at the individual item level⁴⁸. They also provide sample-free measurements (i.e. the results are applicable to all similar groups once the validation process has occurred). In IRT, additional model parameters are used to model the relationship between the individual's trait, the item property and the probability of endorsing an item. The assumption in IRT is that the "probability of answering any item in the positive direction is unrelated to the probability of answering any other item positively for people with the same amount of the trait"^{28,29}. RMT differs somewhat in that the data are assessed to see if they fit the Rasch model. RMT allows for the creation of linear, interval-level measurement from categorical data. In the case of non-fitting data (items or persons), data can be further examined to understand why they do not fit or removed from the data set. Rasch analysis can be used to examine the properties of previously constructed scales as well as in the construction of new scales, and is important in making interval scales⁵⁰.

Although there has been a general shift towards using IRT in more recent years ⁸ for developing and validating a PROM, there are some drawbacks to its use over CTT. One issue relates to the sample size required. It is recommended that ¹⁴ sample sizes based on CTT ¹⁴ should be large enough for the descriptive and exploratory pursuit of meaningful estimates from the data, starting with a sample of 30 to 50 subjects may be reasonable in some circumstances⁵¹. At later stages of psychometric testing, various recommendations have been given for exploratory factor analyses with

recommendations of at least five cases per item and a minimum of 300 cases or to enlist a sample size of at least 10 times the number of items being analysed, so a 20-item questionnaire would require at least 200 subjects⁵¹. For IRT, sample sizes of a minimum of 150-250 patients has been proposed, with around 500 patients recommended for the latter stages of validation^{51,52}.

In addition to the inflated sample size recommendations for IRT, additional expertise in the study team is often required, and this may consequently result in greater development costs. Furthermore, strict assumptions in the model can mean that items may be rejected even when they have good content validity if they do not fit the IRT model. CTT should therefore not be disregarded and indeed, most authorities will agree that aspects of both CTT and IRT have a role to play in the validation process of a modern PROM.

Clinical use of PROMs in Hepatology

Most health providers treat patients with the aim of improving radiological, biochemical, histological and clinical assessments¹. Historically, health outcomes have focused on death or clinical indicators such as infection rates, readmissions, re-operations and adverse events¹⁰. Many of the symptoms that are experienced by patients in Hepatology may be undetected by clinical tests or underestimated by clinicians⁵³. The need to assess the impact of health treatment on patients and to demonstrate the value of the care provided to the patient by the provider is now recognised⁹. There is constant pressure on healthcare providers to improve the quality of healthcare provided and make it more patient-centered⁵⁴. Given how much money is spend on treatment, it is important to assess if the treatment given offers value for the money. Clinical applications of PROMs can be divided into:

Clinical research and trials: Health regulatory bodies such as the FDA and the National Institute for Health and Care Excellence (NICE) require PROMs to be incorporated into the assessment of new treatments, health technologies or medical devices

Quality improvement projects: PROMs can be very helpful in assessing the impact of a new service or project from the patient perspective. However, PROMs must be integrated into clinical practice with strong incentives to encourage their routine use in such quality improvement projects

Clinical practice. Measuring PROMs in clinical practice contributes to patient-centeredness and measures clinical effectiveness from a patient perspective.

There has been widespread adoption of PROMs use within the research field, especially since FDA and EMA recommended that PROMs should form part of the outcome assessment for new drug trials^{2,5}. Reporting guidance for PROMs has also now been incorporated as an extension to CONSORT reporting for trials⁷. The value of collecting PROMs data routinely is now recognised as an important part of driving the delivery and organisation of healthcare and can thereby help to improve healthcare quality⁹.

Although individual hospitals and clinicians have started to implement routine PROM collection, widespread adoption is largely restricted to England, Sweden and parts of the United States^{9,55}. PROMs have now been implemented in England for the routine collection following some elective surgery (<https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms>). Their potential to be used in other clinical areas, such as oncology^{56,57}, multiple sclerosis⁷⁴ is now also recognised. The routine collection of PROMs is not without its challenges however^{1,9,55,59,60}.

Some of the practical challenges to routine integration include: the selection of the most appropriate tool; difficulties with patient completion (for example, lack of comprehension, elderly and frail or sick populations); clinical reluctance; achieving high rates of patient participation; operational difficulties; lack of clarity about the PROM; times pressure for patients and clinicians; lack of human resources; recognition of the

three dimensions of quality (safety, effectiveness and experience); attributing outcomes to the quality of care; providing meaningful outputs from PROMs data for differing audiences; and avoiding misuse of PROMs^{1,9,55,59,60-65}.

Calvert *et al*, have developed a series of key strategic priorities that should be considered when implementing PROMs in real-world situations⁶⁶:

Ensure international collaboration across multiple stakeholders to agree on a standardised approach to PROM assessment

Develop a comprehensive standard set of recommendations, methods and tools that are applicable to the generation of real-world evidence

Formulate a clear governance process including an ethical framework for how patients should be consented, who selects patients, who has access to data and how data will be used

Establish standard sets of PROMs, electronic tools and administration schedules

Develop and use electronic PROMs where possible

Minimise workload and technical complexity for patients and clinicians

Consider the objectives of the PROM assessment, timings, length of follow-up, strategies for managing missing data and inclusion of diverse patient populations

Ensure data collection adheres to the FAIR (findable, accessible, interoperable and reusable) guidance

Provide guidance on interpretation and use of the data

Ensure both patients and clinicians gain value from PROM collection to tailor their needs.

Future of PROMs

With new treatment and technologies, mortality is reducing and more patients are living with their illness for longer. As such, there is a growing need to develop and implement PROMs to facilitate the translation of clinical research into practice and, in keeping with the principles of shared clinical decision-making as part of routine clinical practice.

The increased use of digital media presents an exciting opportunity for PROM capture and adaptation. By utilising new technologies to aid PROMs capture and support interpretation, more clinicians may be encouraged to use PROMs as part of their routine clinical care. For example, innovative delivery methods using app or web-based based technology (for example, through data platforms such as REDCap- Research Electronic Data Capture (<https://www.project-redcap.org/>)) are helping to streamline data capture from patients by facilitating PROM completion on tablets, mobile phones and the internet.

Employing digital media also allows novel methods such as Ecological Momentary Assessment (EMA)⁶⁴⁻⁶⁸ to be used. EMA refers to a collection of methods often used in behavioural medicine research where a patient repeatedly reports on, for example, their symptoms or quality of life close in time to when they experience them and in their own environment⁶⁴. EMA data can be collected in various ways, including written diaries, electronic diaries and telephone. EMA using mobile phones, for example, could facilitate the collection of PROM data in real-time and overcomes some of the inherent problems of PROMs, such as patient recall accuracy.

The burden of data collection associated with the routine collection of PROMs data in practice can be reduced by simplifying data collection using techniques such as computerised adaptive testing (CAT). CAT involves using a computer to administer a PROM, one question at a time. The CAT then uses an algorithm to choose the subsequent question based on the previous answer given. For example if a PROM is assessing hand function and in response to the first question a patient answers that their hand function is 'normal', then there is little to be gained from asking increasingly granular questions about hand problems, which may be more appropriate to someone who has 'non-normal' hand function. By pre-selecting questions, the PROM score can be determined without having to ask all of the questions^{61,62,69-70}.

Another future technology is that CAT and electronic PROMs could be administered by virtual assistants (such as Siri or Alexa, or similar) using voice recognition software to avoid the need for manual form filling to further reduce the manpower required for data collection⁵⁹⁻⁶².

CONCLUSION

The process of developing a new PROM is often a complex and resource-intensive process. If possible, researchers should first consider whether any existing PROMs could be suitable candidates for use, or if they could be adapted. This review provides a general introduction to PROMs and some background regarding the recent drive to collect PROM data. It then reports on findings from a scoping review that identified good practice and issues that should be considered prior to adapting existing PROMs. These issues are organised under the specific headings of: defining the requirements of a PROM, identifying and appraising existing tools, adapting existing PROMs, issues of content validity and getting the right people involved. The review ends with some insights into different psychometric methods, clinical use of PROMs and future PROMs developments.

16%

SIMILARITY INDEX

PRIMARY SOURCES

1	link.springer.com Internet	136 words — 2%
2	cdr.lib.unc.edu Internet	94 words — 1%
3	T. Luckett, M.T. King. "Choosing patient-reported outcome measures for cancer clinical research – Practical principles and an algorithm to assist non-specialist researchers", European Journal of Cancer, 2010 Crossref	74 words — 1%
4	www.scribd.com Internet	70 words — 1%
5	pure-oai.bham.ac.uk Internet	57 words — 1%
6	Walshe, Catherine, Preston, Nancy, Johnston, Bridget. "EBOOK: Palliative Care Nursing: Principles and Evidence for Practice", EBOOK: Palliative Care Nursing: Principles and Evidence for Practice, 2018 Publications	55 words — 1%
7	onlinelibrary.wiley.com Internet	52 words — 1%
8	cronfa.swan.ac.uk Internet	

50 words — 1%

9 Laith Alrubaiy, Hayley A Hutchings, John G Williams. "Assessing patient reported outcome measures: A practical guide for gastroenterologists", United European Gastroenterology Journal, 2014 47 words — 1%
Crossref

10 www.nature.com 47 words — 1%
Internet

11 Luke Geoghegan, Thomas D Dobbs, Jeremy N Rodrigues. "PART 2: Can improvement in outcome measurement in plastic surgery be achieved in today's health systems?", Journal of Plastic, Reconstructive & Aesthetic Surgery, 2019 37 words — 1%
Crossref

12 Thomas D Dobbs, Jeremy Rodrigues, Andrew M Hart, Iain S Whitaker. "Improving measurement 1: Harnessing the PROMise of outcome measures", Journal of Plastic, Reconstructive & Aesthetic Surgery, 2019 34 words — < 1%
Crossref

13 www.ncbi.nlm.nih.gov 34 words — < 1%
Internet

14 "Overview of Classical Test Theory and Item Response Theory for the Quantitative Assessment of Items in Developing Patient-Reported Outcomes Measures", 'Elsevier BV' 33 words — < 1%
Internet

15 Loretta A. Williams, Meagan S. Whisenant, Tito R. Mendoza, Shireen Haq, Karen N. Keating, Brian Cuffel, Charles S. Cleeland. "Modification of existing patient-

reported outcome measures: qualitative development of the MD Anderson Symptom Inventory for malignant pleural mesothelioma (MDASI-MPM)", Quality of Life Research, 2018

Crossref

16 journals.sagepub.com 31 words — < 1%
Internet

17 David Cella, William Riley, Arthur Stone, Nan Rothrock et al. "The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008", Journal of Clinical Epidemiology, 2010 26 words — < 1%
Crossref

18 Thomas D. Dobbs, Maarten Ottenhof, John A.G. Gibson, Iain S. Whitaker, Hayley A. Hutchings. "The Patient-Reported Outcome Measures In Skin Cancer Reconstruction (PROMISCR) study: Anglicisation and initial validation of the FACE-Q Skin Cancer Module in a UK cohort", Journal of Plastic, Reconstructive & Aesthetic Surgery, 2021 26 words — < 1%
Crossref

19 Cappelleri, Joseph C., J. Jason Lundy, and Ron D. Hays. "Overview of Classical Test Theory and Item Response Theory for the Quantitative Assessment of Items in Developing Patient-Reported Outcomes Measures", Clinical Therapeutics, 2014. 25 words — < 1%
Crossref

20 Thomas D. Dobbs, Sarah Hughes, Nicholas Mowbray, Hayley A. Hutchings, Iain S. Whitaker. "How to decide which patient-reported outcome measure to use? A practical guide for plastic surgeons", Journal of Plastic, Reconstructive & Aesthetic Surgery, 2018 18 words — < 1%
Crossref

21	jpro.springeropen.com Internet	18 words — < 1%
22	Bonnie Addario, Jan Geissler, Marcia K. Horn, Linda U. Krebs et al. "Including the patient voice in the development and implementation of patient - reported outcomes in cancer clinical trials", Health Expectations, 2019 Crossref	17 words — < 1%
23	www.tandfonline.com Internet	15 words — < 1%
24	Clarissa W. Ong, Jennifer Krafft, Michael E. Levin, Michael P. Twohig. "A systematic review and psychometric evaluation of self-report measures for hoarding disorder", Journal of Affective Disorders, 2021 Crossref	14 words — < 1%
25	researchoutreach.org Internet	14 words — < 1%
26	pada.psycharchives.org Internet	13 words — < 1%
27	www.frontiersin.org Internet	12 words — < 1%

EXCLUDE QUOTES ON
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES < 12 WORDS