Patient-reported outcome measure translation: An overview

Cate Talley[1] and **Shawn McKown**[2]
1 RWS Life Sciences, Chicago, Illinois, USA
2 RWS Life Sciences, East Hartford, Connecticut, USA

Correspondence to:
Shawn McKown
RWS Life Sciences
101 E River Dr.
East Hartford, CT 06108
USA
+1 860–503–1586
shawn.mckown@rws.com

Abstract
The unique nature of patient-reported outcome (PRO) measures presents unique challenges for translation. Regulators emphasise the importance of maintaining conceptual equivalence across all languages in multilingual and multinational trials, while making necessary cultural adaptations. This article will provide an overview of the central issues affecting PRO measure translation, best practices for PRO measure translation, and ways to improve the translatability of PRO measures at the development stage.

Patient-reported outcome (PRO) measures make unique contributions to clinical research. By asking patients to report their own experiences directly, PRO measures allow researchers to evaluate effects on patients' quality of life in ways that reporting by healthcare professionals simply cannot. But this unique role entails unique challenges. PRO measures must be written in non-technical language to ensure they are understood by diverse populations of laypeople (including those with low levels of education and literacy) and must be interpreted consistently enough to provide meaningful data. The need for translation in multilingual and multinational trials only compounds these challenges, as the concepts being measured (often subjective) must be rendered in ways that are both understandable to each local population (itself diverse) and consistent across all languages in the trial.

The translation of PRO measures must achieve a balance of cultural adaptation and conceptual equivalence.

The translation of PRO measures thus proceeds differently from that of other clinical outcomes assessments in order to achieve this delicate balance of cultural adaptation and conceptual equivalence. In the following article, we will provide an overview of the translation process for PRO measures and its regulatory basis, as well as reflections on the difficulties most commonly encountered while translating PRO measures and how those difficulties can be mitigated by considering translatability during PRO measure development.

Regulatory guidance and industry standards

EMA and FDA guidance

Translation and cultural adaptation of PRO measures are addressed by both the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA), as they affect data collection for primary and secondary endpoints.[1,2] Both bodies indicate the importance of evaluating the process used in translating PRO measures in order to assess whether content validity has been maintained across all languages in a trial. However, neither body provides detailed requirements or guidelines for the process to be used.

ISPOR and ISOQOL guidelines

In lieu of detailed regulatory guidance, industry

standards for translating PRO measures have been shaped by two non-regulatory organisations. In 2005, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Taskforce for Translation and Cultural Adaptation released a report outlining best practices for PRO measure translation.[3] This was echoed in 2013 when the International Society for Quality of Life Research (ISOQOL) published recommended minimum standards for PRO measures.[4] The full process outlined in the following section, commonly referred to as linguistic validation (LV), reflects the best practices detailed by ISPOR.

Both the ISPOR and ISOQOL recommendations emphasise the need to verify quality through qualitative assessment, most likely through cognitive interviewing or debriefing. In cognitive debriefing, each translation is tested with patients from the target population (or alternately with laypeople) who are interviewed to confirm that the translation is clearly understood, accurately interpreted, and perceived to be culturally relevant. Best practices in the industry thus dictate that feedback from the target audience of translated PRO measures be employed to validate the quality of the translations.

Overview of the full linguistic validation process

The full LV process through which translations of PRO measures are developed includes a number of safeguards to ensure conceptual accuracy and equivalence, as well as cultural relevance. Different questionnaire developers and translation companies employ slightly different processes, which can be considered variations on the model below:

1. Explanation of concepts: A document is developed by the questionnaire's developer or the translation company, during or after the development of the PRO measure, to clarify the intended meaning of the key concepts in the source text of the questionnaire. The explanation of concepts serves as a reference for translators throughout the LV process to guide their interpretation of the source text when rendering it in the target language.
2. Forward translation and reconciliation: Two translators independently translate the source text, then collaborate to reconcile their translations and create a single forward translation. This reconciliation process serves as a check on each translator's interpretation of the source text, selecting for a more accurate and appropriate translation.

3. Back translation and review: A third, independent translator then back-translates the reconciled forward translation to English (or other source language). The back translation is compared to the source text to identify discrepancies in the rendering of key concepts. Any discrepancies are resolved through discussions with the team of translators, revision of the forward translation, and back translation of the revised wording.
4. International harmonisation: Back translations from all languages in the scope of the trial or translation project are reviewed together to identify conceptual discrepancies across languages and verify a coherent approach to rendering concepts. Where problematic concepts are identified, further guidance may be sought from the questionnaire's developer.
5. Expert reviews: Draft translations may be reviewed by experts in the target countries (e.g. clinicians, client subsidiaries, developer subsidiaries) to ensure that the translation aligns with local usage.
6. Cognitive debriefing: Interviewers in the target countries test the draft translations with diverse subject populations of laypeople or patients, identifying concepts that are not accurately understood or that are perceived not to be culturally relevant. Through discussion with the team of translators, and in view of problems identified in other languages, highly problematic concepts may be revised through a process similar to the one described in #3.
7. Proofreading: The final product is proofread by a translator to ensure it is free of errors.

Key concerns in translating PRO measures

As reflected in the process outlined above, PRO measure translation differs from other forms of medical translation in fundamental ways. PRO measures must achieve conceptual validity using nontechnical language, and must do so consistently across languages. Translators of PRO measures must carefully assess the best ways to reflect the content and register of an everyday expression using what may be a very different set of everyday expressions in another language. For this reason, translation of PRO measures reflects a distinct area of expertise within the field of medical translation. Both the LV process described above and the specific competencies of the translators involved contribute to achieving quality PRO measure translations.

Allowing cultural adaptation

The task of cultural adaptation during the PRO measure translation process falls first to the forward translators. While creating their initial translations and a single reconciled translation, the forward translators must consider which concepts in the source text of the questionnaire can be translated literally without loss of meaning, and which must be adapted to suit the cultural context of the target language. Such adaptations can range from using different linguistic structures or idiomatic expressions with similar meaning to providing entirely different examples (e.g. culturally adapted examples of sports or food). Whatever the scale of the adaptation, translators must consider how to increase cultural relevance in a way that maintains the integrity of the concept being measured.[5]

The effectiveness of the forward translators' cultural adaptation is assessed during cognitive debriefing, wherein subjects in the target country may provide feedback about the translation's cultural relevance and offer suggestions for better adapting the translation.

Ensuring conceptual equivalence

Of course, cultural adaptation of PRO measures cannot be pursued at the expense of maintaining

conceptual validity across all languages in the trial. The LV process thus includes a number of safeguards to ensure that translations are not inaccurate (either due to translators' misinterpretation or by over-adaptation).

The first of these safeguards is the explanation of concepts, which prevents inaccuracies by providing the translators with a clear definition of each item at every stage of the process.

Back translation review allows for conceptual equivalence to be directly assessed. Conceptual discrepancies that are identified during back translation review may be easily resolved if they reflect a misinterpretation of the concepts in the source text, or they may require further discussion with the team of translators to determine what translation best reflects the source concepts without unduly sacrificing cultural relevance.

Here again, cognitive debriefing can evaluate the success of the negotiations between adaptation and equivalence, and help make revisions to the text where needed.

Developing PRO measures with translatability in mind

Though the LV process and the expertise of PRO measure translators ensure a quality translation, the translatability of the source text of the questionnaire fundamentally affects how well it can be rendered in other languages. The clarity and discreteness of the concepts being measured in the source text directly impact the degree to which strict conceptual equivalence is possible across languages.

When developing PRO measures, avoiding two common pitfalls can greatly increase the translatability of the questionnaire.

Pitfall #1: Semantically rich concepts

Many concepts of interest to quality of life research, and so commonly assessed by PRO measures, are compound, describing many symptoms and experiences. Such concepts may already be ambiguous in the source text, and subjects may place emphasis on different aspects of the concept and therefore interpret the item differently. The ambiguity is only amplified in translation, since the symptoms and experiences that make up the compound concept may be grouped differently in other languages and it may therefore be difficult to articulate the same complexity without introducing concepts,

eliminating elements, or shifting emphasis. For these reasons, compound concepts such as *fatigue*,[6] *bother*,[7] *frustration*,[8] and *distress*[9] should be avoided where possible.

Pitfall #2: Overlapping concepts

Response sets are integral to the data-collection function of PRO measures but can present a particular challenge for translation.[10] Gradations of amount or degree are distinguished differently from language to language, making it difficult to maintain the differences between response options without departing from the concepts in the source text. It is therefore preferable to use distinct concepts where possible, rather than gradations of amount or degree. For example, the response set "None of the time / A little of the time / Some of the time / A lot of the time / Most of the time / All of the time" should be replaced by "Not at all / Rarely / Sometimes / Often / Always". Where using distinct concepts is not possible, the shift to a numeric rating scale can allow for replacing potentially overlapping concepts of amount or degree with clearly differentiated numerical responses.

Tools for improving translatability

To assess the translatability of PRO measures at the development stage and to identify items that will benefit from revision, it is recommended that PRO measures undergo the processes of face validation and/or translatability assessment.[11–13] In face validation, an expert reviews the questionnaire to ensure that all concepts are clear, discrete, and unambiguous, while in translatability assessment, translators review the questionnaire to identify areas of potential difficulty for translation. Items identified as problematic by either process can then be referred back to the questionnaire's developer for revision.

Attending to PRO measure translatability at the development stage creates source texts that contain less ambiguity and fewer culturally-specific concepts, making it easier to maintain consistency and make appropriate cultural adaptations during the translation process. This up-front investment leads to fewer delays during the translation process and higher quality data in the trial.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Committee for Medicinal Products for Human Use. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products. London, European Medicines Agency. 2005 Jul 27.
2. US Department of Health and Human Services FDA Center for Drug Evaluation and Research, US Department of Health and Human Services FDA Center for Biologics Evaluation and Research, US Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labelling claims: draft guidance. Health Qual Life Outcomes. 2006 Dec;4:1–20.
3. Wild D, Grove A, Martin M, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health. 2005 Mar;8(2):94–104.
4. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Qual Life Res. 2013 Oct 1;22(8):1889–905.
5. Gawlicki MC, Angun Ç, Brandt BA, McKown S, Schulz C, Talbert M. Localization of activities and equivalence of movements in clinical outcomes

assessments. Value Health. 2013 Nov 1;16(7):A597.
6. Gawlicki MC, Brandt BA, McKown S, Talbert M. The translatability of fatigue across languages. Presented at the International Society for Quality of Life Research, 21st Annual Conference, Berlin, October 15–18, 2014 (poster).
7. Gawlicki MC, McKown SM, Talbert MJ, Brandt BA. Application of Bother in patient reported outcomes instruments across cultures. Health Qual Life Outcomes. 2014 Dec;12(1):18.
8. McKown S, Angun C, Talbert M, Brandt BA, Gawlicki MC. Assessing the translatability of the term "frustrated". Value Health. 2014 Nov 1;17(7):A516–7.
9. Talbert M, Brandt B, McKown S, Gawlicki M. Assessing the translatability of the term "distressed". Value Health. 2015 Nov 1;18(7):A710.
10. Gawlicki MC, Brandt B, McKown S, Talbert M. Translatability of response sets used in patient reported outcomes and best practices for development. Value Health. 2012;7(15):A484.
11. Gawlicki MC, Brandt BA, McKown S, Talbert M. Face Validation and Translatability Assessment of Clinical Outcomes Assessments. Presented at the 20th Annual Conference of the International Society for Quality of Life Research, Miami, October 9 - 12, 2013 (poster).
12. Gawlicki MC, Handa M, McKown S. Preempting difficulties in linguistic validation: the use of face validation to create more sound translations. Value Health. 2010;13(7):A336.
13. Acquadro C, Patrick DL, Eremenco S, et al. Emerging good practices for translatability assessment (TA) of patient-reported outcome (PRO) measures. J Patient Rep Outcomes. 2017 Dec 1;2(1):8.

Author information

Cate Talley, PhD, is a linguistic validation consultant with RWS Life Sciences.

Shawn McKown, MA, is the Senior Director of Linguistic Validation at RWS Life Sciences.

My experience attending EMWA conferences

It all started few years ago in a "career day event" organised by my previous postgraduate programme. I knew nothing about medical writing, let alone EMWA. Among several speakers, who were describing different job opportunities with various responsibilities and roles, it was the lecture about medical writing that made me the most curious. I knew I liked science and I liked writing, specifically scientific writing, so "medical writing" stuck in my head. About a year later, I applied for a medical writing job, although the agency apparently needed an experienced writer. I did not get the job, but still, the idea stuck in my head.

After that, while I had started a new postdoctoral position, I decided that I still would like to know more about medical writing. Although I was a bit sceptical and, to be honest, scared, I registered for EMWA's conference and signed up for a couple of workshops. To my surprise, the people at the EMWA conference were really friendly and welcoming. They were coming to me after seeing my green badge, introducing themselves, and insisting that I not be shy. They explained what they did and how they ended up in medical writing. This happened throughout the conference – from the first networking event to the coffee breaks between workshops, at the breakfast tables, and of course during the social events. All of this made me feel comfortable during the conference. Moreover, the workshops with their precise pre-workshop assignments and well-organised lectures convinced me that it was the right decision to register for this EMWA conference and that I should register for forthcoming conferences. The various workshop topics, from regulatory writing to proofreading techniques, and even writing for the internet, make it possible for almost everyone to have choices and benefit even without any previous knowledge.

To my surprise, the people at the EMWA conference were really friendly and welcoming. They were coming to me after seeing my green badge, introducing themselves, and insisting that I not be shy.

Though I still do not know where my career will take me, for someone who works most of the time in the laboratory, attending EMWA conferences, symposia, and workshops is a valuable investment and experience. Needless to say, meeting old friends and finding new ones is also a pleasant part of it.

Mona Saffarzadeh, PhD
Center for Thrombosis and Hemostasis,
Johannes Gutenberg University Medical Center,
Mainz, Germany
monasaffarzadeh81@gmail.com