VALUE IN HEALTH ▮ (2017) ▮▮▮–▮▮▮

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Standards for Instrument Migration When Implementing Paper Patient-Reported Outcome Instruments Electronically: Recommendations from a Qualitative Synthesis of Cognitive Interview and Usability Studies

Willie Muehlhausen, DVM[1], *Bill Byrom, PhD*[2,*], *Barbara Skerritt*[1], *Marie McCarthy, MBA*[1], *Bryan McDowell, MBA*[3], *Jeremy Sohn, BA*[3]

[1]*ICON Clinical Research, Dublin, Ireland;* [2]*ICON Clinical Research, Marlow, Buckinghamshire, UK;* [3]*Novartis Pharmaceuticals, Basel, Switzerland*

ABSTRACT

Objectives: To synthesize the findings of cognitive interview and usability studies performed to assess the measurement equivalence of patient-reported outcome (PRO) instruments migrated from paper to electronic formats (ePRO), and make recommendations regarding future migration validation requirements and ePRO design best practice. **Methods:** We synthesized findings from all cognitive interview and usability studies performed by a contract research organization between 2012 and 2015: 53 studies comprising 68 unique instruments and 101 instrument evaluations. We summarized study findings to make recommendations for best practice and future validation requirements. **Results:** Five studies (9%) identified minor findings during cognitive interview that may possibly affect instrument measurement properties. All findings could be addressed by application of ePRO best practice, such as eliminating scrolling, ensuring appropriate font size, ensuring suitable thickness of visual analogue scale lines, and providing suitable instructions. Similarly, regarding solution usability, 49 of the 53 studies (92%) recommended no changes in display clarity, navigation, operation, and completion without help. Reported usability findings could be eliminated by following good product design such as the size, location, and responsiveness of navigation buttons. **Conclusions:** With the benefit of accumulating evidence, it is possible to relax the need to routinely conduct cognitive interview and usability studies when implementing minor changes during instrument migration. Application of design best practice and selecting vendor solutions with good user interface and user experience properties that have been assessed in a representative group may enable many instrument migrations to be accepted without formal validation studies by instead conducting a structured expert screen review.

Keywords: cognitive interview, electronic patient-reported outcomes, ePRO, measurement equivalence, patient-reported outcomes, PRO.

Introduction

Because of significant improvements in the integrity, quality, and timeliness of data collected and increased awareness of the potential benefits of electronic collection, a growing number of clinical trials are using electronic media (smartphones and tablets) to collect patient-reported outcomes (PROs). Because many PRO instruments were developed and validated on paper, care is needed when migrating them to electronic formats (ePRO) to ensure that the measurement properties of the instrument are unchanged and that the electronic version is easy to use in the target group of patients. In 2009, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) ePRO Good Research Practices Task Force published recommendations on the evidence needed to support measurement equivalence when migrating from paper to electronic formats [1]. This task force recommended that minor changes to an instrument because of migration should require a cognitive interview and usability study in the target patient population to demonstrate measurement equivalence. Such minor changes include, for example, minor formatting changes such as presenting only a single question per screen or wording changes such as changing question response instructions from “tick or circle” on pen and paper to “select” on an electronic implementation. These recommendations have been largely adopted by the industry and regulators.

In this context, cognitive interviews typically involve developing a semistructured interview that is conducted by a trained qualitative interviewer to collect information about patient experience after they completed the instrument both on paper (or its original form) and in the electronic format. Structured, probing

*Address correspondence to: Bill Byrom, ICON Clinical Research, Unit 2 Globeside Business Park, Marlow, Buckinghamshire SL7 1HZ, UK. E-mail: bill.byrom@iconplc.com.

1098-3015$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2017.07.002

questions help to identify whether changes in format and presentation might affect the way patients respond to the questions and, thus, whether the modality provides equivalent patient responses. These studies are typically carried out in a small sample (n = 5–10) of the target patient population and interviews are transcribed and summarized qualitatively [2].

The purpose of our synthesis was to explore whether routine performance of cognitive interview and usability studies should remain a recommendation for all migrations requiring minor modifications or whether the benefit of growing evidence obtained from conducting these evaluations is supportive of other less arduous approaches. We also used our synthesis to confirm ePRO design best practice recommendations.

This is not the first review exploring learnings from previous migration studies. Two meta-analyses of equivalence studies performed on instruments migrated from original to electronic formats have been reported [3,4]. Both analyses concluded that there is no meaningful evidence that migration to alternative formats affects instrument measurement properties (the analyses considered 46 and 72 equivalence studies, respectively [3,4]).

One of the fundamental aspects of our analysis has been to consider instruments as a collection of response scale types as opposed to a combination of items. Common response scale types include the following [5]:

1. *Verbal response scales (VRSs):* These comprise a question prompt and an associated list of response options ordered in a logical scale order, for example, mild, moderate, and severe.
2. *Numeric response scales (NRSs):* These scales combine question text with a horizontal list of ordered numbers reflecting the degree of association with the construct measured, such as severity or agreement. The scale interpretation is typically anchored using a text description to describe the first and the last number of the scale. An NRS to measure pain severity might, for example, ask the subjects to rate their pain on a scale from 0 to 10, where 0 represents "no pain" and 10 represents "worst possible pain."
3. *Likert scales:* These scales measure a concept ranging from a positive to a negative rating, with the center option being neutral, for example, measuring satisfaction from very satisfied to very unsatisfied. These can be presented using a VRS or an NRS.
4. *Visual analogue scales (VAS):* These scales use a straight horizontal line on which the respondents mark their assessment of a specific construct. The scale interpretation is typically anchored using a text description to describe each end of the horizontal line, for example, "no pain" to "worst possible pain."

Additional response options sometimes included in electronic clinical outcome assessment instruments include yes/no fields, number entry fields, free-text fields, multiple choice fields, and time and date fields. Because these response types are common in everyday usage of a mobile device and personal computer (PC) applications, we did not consider it necessary to evaluate them specifically in this work.

The rationale for considering migration assessment by response scale types is founded in the hypothesis that potential changes in an instrument's measurement properties, after minor formatting and layout changes due to migration, are primarily concerned with understanding whether subjects can interact with each response scale type appropriately and in the same way on both modalities, independent of the specific question item or construct that each item evaluates. Ensuring each item is an appropriate measure of the required construct has already been assessed thoroughly in the development and psychometric validation activity performed by instrument authors, and so when changes are minor there is no requirement to re-assess this in cognitive interviews associated with migration assessment. This may mean that previous migration studies on instruments using the same response scale types can provide evidence of migration acceptability for new instruments, so long as ePRO design best practice standards are followed.

Methods

We synthesized findings from all cognitive interview and usability studies performed between 2012 and 2015 by a contract research organization (CRO) to which a number of the authors belong: 53 studies comprising 68 unique instruments and 101 instrument evaluations. These studies are rarely published in the scientific literature, but are routinely included in drug approvals by sponsor organizations to regulatory authorities to support the appropriate use of ePRO for clinical trials within the submission [1].

In all studies, cognitive interview and usability assessment was performed using a standardized semistructured interview conducted by an experienced qualitative interviewer. Patients were asked to read and complete both modes of instrument administration. Interviewers probed whether any perceived differences in the self-report task or aspects of the changes between formats—such as overall appearance, text size, instructional information, moving from question to question, and how responses were selected—may, in the patients' perception, have caused them to potentially answer differently between formats. Usability questions explored the clarity of text and images, ease of navigation, use of touch screen, and whether participants felt they would be able to use the electronic solution without help. All interviews were recorded and transcribed, and findings were summarized.

For each study, we identified the instruments studied and the response scale types they contained, the patient population and sample size, and the electronic modality compared with the original paper instrument. Reported findings of the cognitive interviews were summarized, specifically identifying whether changes in the way patients responded to instrument items because of migration differences were reported, and any additional recommendations. For each usability testing report, we summarized findings relating to display clarity, navigation, use of touch screen/stylus (where applicable), and the ability of patients to use the electronic solution without help. We synthesized the findings across all studies included.

Results

Description of Studies and Instruments

The 53 studies were conducted in samples ranging from 5 to 30 patients (median sample size: 10 patients) and included patients from a broad range of therapeutic areas. Out of these studies, 6 studies (11%) were conducted in healthy volunteers; 9 studies (17%) included patients with respiratory conditions including asthma and chronic obstructive pulmonary disease; 7 studies (13%) included patients with gastrointestinal conditions such as ulcerative colitis, Crohn disease, and constipation; and 7 studies (13%) included oncology patients including those with breast cancer, melanoma, and gastric/bladder cancer. A further 6 studies (11%) included rheumatology patients all involving patients with osteoarthritis of the knee; 4 studies (8%) included central nervous disease indications including migraine; and 14 studies (26%) involved patients with other conditions (Fig. 1). Patients were aged between 5 and 84 years (Table 1). Four studies included