Principles for the Development of New Clinical Outcome Assessments

Scope

This document is intended to provide considerations that would allow for any newly developed clinical outcome assessment (COA) tools to be implemented on any available data collection mode. This document is intended to apply to the various types of COAs, namely observer-reported outcomes (ObsRO), clinician-reported outcomes (ClinRO), and patient-reported outcomes (PRO).

The assumptions and nomenclature related to the application of this document to a new instrument development project are listed below.

Assumptions

1. This document is intended for use by those embarking on de novo instrument development
2. The considerations below are independent of the mode used for data-collection
3. This document does not replace or supersede guidance for the development of COAs for use as endpoints in clinical trials or the translation and linguistic validation of COAs
    a. References for Content Validity:
        - FDA PRO Guidance
        - ISPOR Task Force Report - Content Validity Part I
        - ISPOR Task Force Report - Content Validity Part II
    b. References for Cultural and Linguistic Translation:
        - ISPOR Task Force Report – Translation and Cultural Adaptation
        - Brislin RW. The wording and translation of research instruments. In: Lonner WJ, Berry JW, eds. Field methods in cross-cultural research. Beverly Hills: Sage, 1986:137-164.

Nomenclature

Throughout this document certain terms will be used synonymously. For example, when referring to a COA, the terms *instrument, tool, questionnaire, and scale* may be used interchangeably to refer to the COA.

The use of the words 'mode' and 'method' in this document departs from that used in the FDA's PRO Guidance. The PRO Guidance defines *modes of administration* as self-administration, interview, or a combination of both. Further, *data collection methods* are defined to include paper-based, computer-assisted, and telephone-based assessments. However, an informal review of the articles included in a meta-analysis conducted by Gwaltney et al (2008) suggests a different use of this terminology in the ePRO literature. For the purposes of this document, the term *data collection mode* refers to various platforms available for instrument administration (i.e., paper and electronic platforms).

General Considerations

- Consider characteristics of the target population and therapeutic area for which the instrument is intended for use.
    - Consider functional abilities associated with the target population (e.g., diabetes-related vision problems, Ménière's disease-related hearing loss, Parkinson's disease-related tremors, stroke-related physical or cognitive impairment)
    - Consider how frequently the concepts to be measured occur or change. The recall period and administration frequency should reflect this consideration.
- Anticipate that the instrument will be translated into other languages.
    - The cultural appropriateness of the instrument's items and responses should be considered (e.g., avoid idiomatic expressions)
    - Translated text is likely to be longer than English text
- Consider patient burden with respect to the length of the instrument and the time needed for completion.
    - Consider whether other COAs are likely to be administered alongside the new COA.
- In addition to the characteristics of the instrument, it is also important to make additional considerations to the appropriateness of migration to each respective platform.
    - Consider the patient burden of completing the questionnaires (e.g., the amount of time to complete the questionnaire, cognitive burden)
    - Instrument developers should provide recommendations about the appropriate modes of administration.
- Consider regions of the world where the instrument will be utilized.
    - Trial planning should consider infrastructure for collection of data electronically (e.g., internet connectivity variation).
    - Assume that translated text will take more space than US English so keep the content compact/ brief where possible.
        - There are direct implications of wording/phrasing changes for how the instrument is formatted on the mode of administration
- Consider whether branching or skip logic, real-time edit checks, calculations, and reminders will be incorporated into the measure.
    - Make explicit recommendations for administration window

Instructions

- Use platform neutral language in instructions where possible. There are often key words and phrases within instructions that pertain to specific platforms (including paper). Platform-neutral language is words and phrases that can be used and understood on various administration platforms.

    - Examples of platform-specific language: "Using the pen / pencil, circle the response …", "Using the numbers on your phone, press the number of the response…", "Using the mouse, click on the response…"
    - Examples of platform-neutral language: "Select the response…", "Choose the response…", "Enter the response…"
- Instructions should be clear and succinct.
- If there is a time frame for recall, be sure that it appears, or is heard with every item for which it is relevant, not just once at the beginning of a series of items.

Item Stems

- Item stems are generally incomplete statements or direct questions. If the same incomplete statement is used for multiple items, each item should be self-contained (i.e., include the full stem and response options, avoid split-stems). Instructional language should be left out of item stems where possible and appropriate.
- Example of split stem (to be avoided)

  During the past 4 weeks, how much has your pain interfered with:
  1. Vigorous activities such as running or heavy lifting?
  2. Moderate activities such as climbing a flight of stairs?
- Example of complete items
  1. During the past 4 weeks, how much has your pain interfered with vigorous activities such as running or heavy lifting?
  2. During the past 4 weeks, how much has your pain interfered with moderate activities such as climbing a flight of stairs?

Response Sets

- The use of discrete response sets (e.g., yes/no, true/false) should be limited to constructs possessing nominal or ordinal level data
- Continuous response sets (e.g., numeric rating scale, adjectival scale) are appropriate for use with ordinal, interval, and ratio level data
- Consider the feasibility of implementing the response set across data collection platforms
    - Different technologies may offer the use of response aides (e.g., spinner/counter) that cannot be operationalized on all platforms

Considerations for Usability Testing

- Test instruments with the intended target population with the intended level of supervision (e.g. test diaries in an unsupervised setting with patients, site instruments at the site and ClinROs with clinicians and patients if appropriate)
- Use the actual platforms the instruments will be deployed on for usability testing, not screenshots (e.g., If the instrument will be deployed on a tablet, use a tablet and not a handheld device for testing).