Exploration of the utility of linguistic validation cognitive debriefing interviews in generating evidence for adding or modifying source COA content

Tim Poepsel, Chryso Hadjidemetriou, Rebecca Israel, Kwanele Masuku, Nkateko Ngalele, Rachael Browning

ISPOR Europe 2025 #ISPOREurope

Introduction

An often overlooked but vital role of linguistic validation (LV) is continuous evaluation of clinical outcome assessments (COAs) for completeness and relevance to the patient experience.

FDA notes that sponsors may consider modifications to improve a COA's ability to reflect the concept of interest.

Importantly, FDA notes that patient input can:

(1) identify which aspects of a concept are most impactful for patients

(2) be used to develop or modify COAs accordingly

The cognitive debriefing (CD) component of the LV process is well-positioned to evaluate and update COAs in this way.

Namely, LV can occur frequently following instrument development and ahead of clinical trials

- engaging participants from a range of cultural, socioeconomic, and health literacy backgrounds
- who can routinely provide feedback on and relevance checks for COA content as understandings of conditions, language use, culture, technology and treatments change across time

Methods

This pilot tested a novel probe assessing completeness / appropriateness of COA content, in a study involving cognitive debriefing of one 200-word chronic lung condition patient reported outcome (PRO) measure in 12 languages for 15 countries, for a total of 21 language-country pairs (see Table 1). There were 5 patients tested per language-country pair, for a total of 105 (61 females; 44 males; age range: 18-95 years; average educational attainment: 12.2 years, range 7-20 years – see Table 1 for fine-grained demographic data for each sample).

The instrument consisted of 6 items covering different symptom areas, such as cough, fatigue, and breathlessness. Notably, the instrument lacked instructions or introductory text, and items themselves consisted of a symptom area (i.e., cough) followed directly by a set of response options unique to each item.

Upon CD interview completion, a probe question was administered to patients asking them to identify any significant missing instrument elements (e.g., questions, topics, response options).

Probe: "Are any important questions or topics related to your disease missing from this questionnaire? If 'yes', what would you add?"

Results

Positive feedback to the probe about missing content was received from 11/21 (52%) locales in the CD sample (see Table 2), identifying 20 potentially missing symptom / QoL categories or other COA content (most frequent: sleep quality: wheezing; chest pain; missing response options – see Table 3). Overall, 24/105 (23%) patients offered feedback to the novel probe, for a total of 34 unique pieces of feedback on COA completeness. See Table 4 for selected examples of patient feedback to the probe about instrument completeness. A majority of feedback was about missing items or symptom areas, but one of the most frequent mentions was about a perceived missing response option for an item about sputum color.

Missing Content by Locale

	Locale	Amount of Feedback
1	Arabic-Israel	1
2	Chinese-China	1
3	Dutch-Belgium	2
4	English-Australia	7
5	English-Canada	2
6	French-Canada	2
7	German-Germany	1
8	Hebrew-Israel	2
9	Italian-Italy	1
10	Spanish-Spain	13
11	Spanish-US	2

Table 2: The amount of missing content feedback received from each locale

Discussion

Rigorous instrument validation methods and pre-testing produce robust and clinically relevant COAs, but COA completeness and patient-acceptability may degrade across time with evolving language use and clinical practice. Pilot testing of a novel probe assessing instrument completeness was well-received by patients, demonstrating effectiveness in capturing novel CD data useful for improving or updating COAs. Use of such probes provides a simple way to continuously monitor for COA relevance and completeness post-development and across time. Linguistic validation studies provide an ideal environment in which to perform this continuous content validation, as they occur relatively frequently in comparison to other types of psychometric assessment or validation (which typically occur once during instrument development). Due to their global nature, LV studies can provide access to samples with variation in cultural, geographic, linguistic, and educational factors that may be potentially even more useful for holistic assessments of the continuing relevance and completeness of COAs.

Conclusion

We propose the standard addition of the completeness probe discussed above for LV CD interviews, to strengthen their ability to contribute to the development and maintenance of maximally patient-centered and relevant COAs in a usage-context marked by constant change in technology, language and culture, and health literacy.

Missing Content Categories by Frequency and Locale

	Missing Content Category	Mentions	Locales
1	ADL impact	2	Italian-Italy Spanish-Spain
2	Bladder control	1	English-Australia
3	Chest pain	3	English-Australia French-Canada Spanish-US
4	Clubbed fingers	1	Spanish-US
5	Cough	2	Chinese-China English-Australia
6	Hospitalization	1	Spanish-Spain
7	Information about disease progression	1	English-Australia
8	Item clarification	2	Spanish-Spain
9	Missing response option	3	Spanish-Spain
10	Mood and depression	1	English-Australia
11	Muscle pain	1	Spanish-Spain
12	Physical function	1	Spanish-Spain
13	Phlegm consistency	1	Spanish-Spain
14	Reduced appetite	2	Dutch-Belgium Hebrew-Israel
15	Self-consciousness or embarrassment	2	English-Australia
16	Sleep quality	3	Arabic-Israel German-Germany Spanish-Spain
17	Smoking	1	Hebrew-Israel
18	Taste and Smell	1	Dutch-Belgium
19	Weight loss	2	English-Canada Spanish-Spain
20	Wheezing	3	French-Canada Spanish-Spain English-Canada

Table 3: All missing content categories, with number of mentions and locales the feedback came from

CD Sample Characteristics

	Language-Country Pair	Age Range (years)	Male:Female	Average Education (years)	Education Range (years)
1	Arabic-Israel	36-75	3:2	12.8	10-16
2	Chinese-China	36-75	3:2	11.4	9-15
3	Chinese-Taiwan	26-75	2:3	13	9-16
4	Danish-Denmark	18-85	2:3	14	10-18
5	Dutch-Belgium	18-85	2:3	12.2	10-16
6	English-Australia	56-85	1:4	14.4	8-19
7	English-Canada	56-85	2:3	13	11-16
8	English-US	26-75	1:4	11.6	10-14
9	French-Belgium	46-85	2:3	11.8	10-15
10	French-Canada	56-95	2:3	12.6	10-16
11	French-France	46-75	2:3	9.4	8-12
12	German-Germany	26-65	2:3	10.8	9-13
13	German-Austria	36-85	1:4	9.6	9-12
14	German-Belgium	26-65	2:3	10.2	9-12
15	Hebrew-Israel	36-75	3:2	12.6	10-16
16	Italian-Italy	46-85	2:3	12.2	10-16
17	Russian-Israel	36-75	3:2	12.6	10-15
18	Spanish-US	36-75	2:3	12	10-14
19	Spanish-Argentina	36-75	2:3	14	7-20
20	Spanish-Spain	36-75	2:3	15	12-18
21	Vietnamese-Vietnam	36-65	3:2	10.4	7-14

Table 1: Language-country pairs in the sample

Examples of Patient Feedback to Completement Probe

	Locale	Representative Feedback on Missing Content
1	Italian-Italy	The impact on daily life/daily activities is not fully covered; yes, you have a question about fatigue. But the impact on daily life activities is not just caused by fatigue, but by the frailty/vulnerability that come with this condition. Patients with [condition] can't go out when the weather is cold and rainy; we run the risk of becoming sick… So the impact is huge and shall be better analyzed.
2	English-Australia	S2 considered the topic of MOOD AND DEPRESSION should have been included in this questionnaire because "[condition] is a difficult illness socially". Fits of uncontrolled coughing in a movie theater, restaurant, shop or plane provoke negative reactions from other people, which make social activities very stressful and difficult; S2 felt depressed by the complications of her illness, but she said, "I have never once had a doctor whom I was seeing about my [condition], ask me about how the illness was effecting my mood. Or did I feel depressed."
3	French-Canada	Yes, I think there should be questions about chest pain, because I think a lot of people with my disease experience it.
4	Spanish-Spain	In the section on SPUTUM COLOUR, I was expecting mention of the "brown" colour.
5	Spanish-Spain	In the section "SPUTUM VOLUME", she says it should be made clear whether the volume is daily or refers to each expectoration
6	Dutch-Belgium	I would add a question about lack of appetite because your taste and smell are affected.

Table 4: Examples of feedback on missing COA content from patients in the sample

rws