



Published in final edited form as:

Qual Life Res. 2014 August ; 23(6): 1767–1775. doi:10.1007/s11136-013-0618-z.

USE OF COGNITIVE INTERVIEWS IN THE DEVELOPMENT OF THE PLUS-M ITEM BANK

Sara J. Morgan, CPO, Dagmar Amtmann, PhD, Daniel C. Abrahamson, CPO, Andre J. Kajlich, and Brian J. Hafner, PhD

University of Washington Department of Rehabilitation Medicine 1959 NE Pacific Street, Box 356490 Seattle, WA 98195

Abstract

Purpose—Measuring constructs such as mobility with patient-reported outcomes (PROs) can enhance clinical and scientific understanding of how health conditions, like lower limb amputation, impact patients' lives. When developing PRO questionnaires, cognitive interviews (CIs) are used to examine if survey items are understandable, clear, and meaningful. The aim of this study was to use CIs to inform item development for the Prosthetic Limb Users Survey of Mobility (PLUS-M), a PRO that measures mobility in prosthetic limb users.

Methods—Thirty-six CIs were conducted with 30 prosthetic limb users. Each participant responded to up to 30 items from the PLUS-M candidate item set. Each item was reviewed by a minimum of five participants who differed in self-reported mobility, literacy, level of amputation, and time since amputation. Items were revised based on participant feedback and substantially revised items were re-evaluated through additional CIs.

Results—Feedback from CIs identified substantial issues in 76 of the total 156 items. These items were subsequently modified or eliminated.

Conclusion—CIs were an essential qualitative step in the development of the PLUS-M item bank and resulted in better functioning items.

Keywords

cognitive interviews; artificial limb; qualitative research; mobility; patient-reported outcome measure

INTRODUCTION

Clinicians and researchers traditionally measure rehabilitation outcomes, such as mobility following lower limb amputation, using standardized performance measures (e.g., timed tests or clinical assessments of function [1-6]). However, performance on these measures may not adequately represent all aspects of an outcome of interest (e.g., mobility) [7]. Patient-reported outcomes (PROs) are health surveys that are designed to be answered directly by the respondent and are intended to help clinicians and researchers understand

how health conditions impact fundamental aspects of the respondents' lives. PROs are well suited to assess complex health outcomes because they solicit information about the respondents' perspectives, opinions, and/or experiences without interpretation by another individual [8]. Well-developed PROs complement performance measures when assessing important health outcomes like mobility.

A variety of PROs are available to rehabilitation clinicians and researchers to assess mobility and mobility-related constructs (e.g. balance confidence and physical function) in persons with lower limb loss. However, many of these instruments possess inherent psychometric limitations (e.g., floor and ceiling effects, poor validity or reliability, and inability to detect clinically meaningful changes) that make them less practical for use in routine clinical care or research [9]. The Prosthetic Limb Users Survey of Mobility (PLUS-M) is being developed with modern psychometric methods to mitigate these common limitations. The resulting instrument is expected to be a valid, reliable, sensitive, and brief instrument for measuring self-reported mobility in prosthetic limb users.

Cognitive interviews (CIs) are used to solicit feedback from respondents to examine if PRO items (i.e. survey questions) are understood as intended by the PRO developers [10]. Information obtained in CIs can be used to infer the cognitive thought processes participants use when they respond to items as well as to assess if questions are interpreted similarly across participants [10-15]. Interviews are purposefully designed to explore the cognitive processes used by respondents as they comprehend the question, recall information needed to answer the question, make decisions about how best to answer the question, and select an appropriate and meaningful response [10].

Evaluation of PRO items through CIs is considered standard practice in the development of psychometrically sound PRO instruments [15, 16]. CIs allow for assessment of individual items' function early in the development process. The information obtained during CIs may then be used to modify or remove items that perform poorly prior to administration of the instrument to a large group of respondents in order to develop scoring. Benefits of cognitive interviewing include (1) identification of problems with comprehension, recall, decision, or response processes required for participants to meaningfully answer a survey question, (2) in-depth exploration of the construct/topic of interest through a semi-structured discussion of survey items with respondents, (3) detection of structural defects that may affect performance of the questionnaire (e.g. unclear layout, illogical questions), and (4) investigation of problems in the survey instructions and response options [10]. The goal of cognitive interviewing in PLUS-M development was to enhance the overall quality of the instrument.

The primary aim of this study was to inform the development of PLUS-M survey items through use of CIs with a diverse group of prosthetic limb users. Data gathered in CIs were used to assess the content validity of candidate items, identify and modify problematic items, and eliminate those items that could not be modified to function as intended. Prior to CIs, the investigators formed an advisory panel of methodological and content experts and conducted focus groups to assist in conceptualization of a PLUS-M measurement model. Following CIs, remaining and revised items were tested in over 1000 prosthetic limb users

to assess item functioning. Details regarding these efforts will be provided in future publications.

METHOD

Data collection and analysis protocols used in this study were based on recent outcome measure development efforts (e.g., PROMIS and Neuro-QoL) [8, 15, 17, 18] and contemporary standards for developing PROs [19].

PARTICIPANTS

Participants were recruited from sites in Seattle, WA; Miami, FL; and Chicago, IL, through print, email, and internet postings. To be eligible for inclusion, interested persons were required to (1) be at least 18 years of age, (2) have a lower limb amputation at or below the hip and at or above the ankle, (3) regularly use a lower limb prosthesis, (4) report the ability to speak and read English, and (5) allow the interview to be audio recorded.

In addition to the criteria listed above, purposive sampling techniques were used to ensure that each question was evaluated by at least five participants who collectively possessed the following characteristics: (1) self-report of high mobility, (2) self-report of fair-to-low mobility, (3) lower than grade 12 education or self-report of cognitive impairment, (4) bilateral lower limb amputation, (5) amputation above the knee, and (6) less than one year since their amputation. Single participants could represent more than one of these traits (e.g. self-report of high mobility and bilateral lower limb amputation) as the intent of the described sampling procedure was to include a wide range of perspectives and experiences. Additional CIs were conducted for items that were modified as well as those that targeted specific populations (e.g. prosthetic limb users who participate in sport activities) to ensure that items are reviewed by people who have experienced the activity described in the item.

The sample size was deemed sufficient as CIs in this study were used to qualitatively assess individual items prior to their inclusion in the PLUS-M item bank, not to derive statistical estimations [14]. In qualitative research (e.g. CIs), the sample must be sufficient to elicit a variety of the perspectives that are relevant to the construct of interest (i.e. mobility with a prosthetic limb). That is, the objective of qualitative research designs is to reduce the chances of discovery failure rather than recruit a sample that is representative of the population [16]. In addition, many of the items tested were modifications of existing items that had presumably been tested by respondents during development of their parent instruments. Similar sampling protocols have been deemed acceptable in previous standardized instrument development efforts [8, 15, 17] and small sample sizes are recognized as an inherent attribute of cognitive interviewing [14].

The study was approved by the University of Washington Institutional Review Board. All participants gave informed consent prior to the interview.

INITIAL PLUS-M SURVEY

The investigators conducted a thorough literature review to identify self-report instruments of mobility or mobility-related constructs (e.g. physical function). More than 1000 candidate

items from 41 unique instruments were identified. Many items addressed similar concepts (e.g., walking up stairs) and were therefore subjected to a process of binning and winnowing [15]. New items were then created by investigators based on professional experience and feedback from focus groups held with prosthetic limb users. Ultimately, 130 items were considered for inclusion in the initial PLUS-M survey. Following initial rounds of CIs, twenty-six additional items were added based on participants' suggested additions and/or revisions. Thus, a total of 156 items were tested with CIs. Most candidate items were modified or created to share the following general structure: they began with the context "Are you able to" and the available response options were "without any difficulty", "with a little difficulty", "with some difficulty", "with much difficulty", and "unable to do". The exceptions included one item that assessed confidence ("How confident are you") and another item that assessed frequency ("How often do you"). The initial survey instructions and examples of candidate items can be found in Tables 1 and 2.

PROCEDURE

CIs were conducted in person (n=8) or via telephone (n=28) by study investigators to accommodate involvement by both local and national participants. The protocol was the same for both types of interviews. Participants were first asked to complete a hard-copy of the initial survey. The survey was either mailed to participants with instructions to complete the survey at the time of interview (for telephone interviews) or provided to them at the interview (for in-person interviews). Immediately following completion of the survey, the investigators asked open-ended follow-up questions about each survey item. Thus, the time between completing the written survey and beginning the interview process was generally less than 5 minutes. The interview process required, on average, 56 minutes (standard deviation = 23 minutes) to complete. Due to the large number of candidate items (n=156), they were divided into seven sets, each with between 25-30 items. Each set included items that spanned a range of mobility activities (i.e. not all items were related to a single topic, such as stair climbing). In the process of conducting interviews, the investigators noted that a number of questions (e.g. "Are you able to run 10 miles?") measured athletic activities that most respondents had never experienced. Therefore, two additional prosthetic users who regularly participated in athletic activities were recruited to review these questions. Participants also responded to questions related to the respondent's general health, activity level, prosthesis, and cognitive function.

Three investigators were trained by an experienced qualitative researcher to conduct CIs using retrospective verbal probing. In this approach, the participant is asked to fill out the survey without any guidance from the interviewer just prior to an item-by-item oral interview. Verbal probing is then used to solicit information about the thought processes used by the participant to answer each question. Willis recommends retrospective verbal probing when evaluating written surveys because it allows the respondent to complete the measure in the intended manner (i.e., unaided, as if in a medical office), thus minimizing the effect of the interview on the respondent's choices [14]. In this study, the number of items administered at each interview was restricted (i.e., each respondent was given 30 or fewer survey questions) to allow for more accurate recall by participants and to limit the total time required for the interview. Retrospective verbal probing has also been adopted by large

national initiatives for the development of recent self-report measures (e.g., PROMIS and Neuro-QoL) [8, 15, 17].

Interview guides were compiled for each interviewer in order to standardize the CI process across investigators. Each guide contained the subset of the items that were to be administered to the participant and an associated list of scripted questions (i.e. probes) that were to be used for the oral interview. These probes were designed to assess comprehension/interpretation (e.g. “what does the term ‘transfer’ mean to you?”) and recall (e.g. “when was the last time you walked continuously for 15 minutes?”) as well as to solicit general information about the participant's cognitive process when answering the item (e.g. “tell me what you were thinking when you answered that question”) [10]. Interviewers also asked unscripted questions (i.e. spontaneous probes) when appropriate to better understand the participant's response. The interviewing team included two researchers who are certified and licensed prosthetists and one researcher who is a prosthetic limb user.

Interviewers audio-recorded all interviews and took field notes to supplement the recordings. Interviewers compiled a detailed summary of each interview within one day of the interview, referencing the audio recording and/or field notes as needed. Interviewers marked each item as “no change” if the item appeared to function as intended or “change and/or learn more” if the item did not function as intended. Investigators discussed items marked as “change and/or learn more” during item review.

ITEM REVIEW

Following the first round of CIs (n = 5 participants per item), interviewer summaries were organized into a single document that detailed collective participants’ feedback for each item. Investigators, as a group, reviewed feedback across respondents and discussed potential concerns with each item. Group consensus determined how items were managed. If the item did not function as intended, the item was revised to address concerns identified in the CIs.

Items that were substantially revised were re-reviewed in a second round of CIs. ‘Substantial revision’ was defined as any change to the item that involved more than (1) addition or removal a supportive word, (2) addition or removal of a word that did not change the meaning of the phrase, (3) substitution of a word to simplify the item semantics, or (4) a change to the order of the words in the item. Revised items were administered to participants who provided useful feedback in the first round of CIs (n=3 participants per item). In addition, a third round of CIs was performed with two athletic participants to further assess well functioning items from the first round or revised items from the second round that involved running or participation in sport activities. Finally, a fourth round of CIs (n=5 participants per item) was performed to assess new items that were constructed by the investigators following feedback given in the first two rounds of CIs. Second-, third-, and fourth-round CIs were conducted in the same manner as the initial CIs. Group consensus by the study investigators was again used to decide if the item was to be revised or if it should be removed from inclusion in the candidate item bank. Improvement in revised items was gauged by participants’ ability to easily understand the intent of each item and provide a meaningful response.

Items that did not have identified concerns and items that were successfully revised were compiled into a candidate item bank to be tested with a large sample of prosthetic limb users.

RESULTS

Thirty participants were involved in a total of 36 CIs. Participants were diverse with respect to current age, age at amputation, years since limb loss, and etiology of amputation. About two-thirds of the participants were male. This is consistent with the gender ratio of persons with lower limb loss reported in the literature [21, 22]. Approximately half of the participants were from Washington State and half from other locations in the United States. Demographic characteristics of the study sample are provided in Table 3.

The quality of PLUS-M items, instructions, and response options were examined and improved based on participant feedback. A total of 205 items were examined across four rounds of CIs (Figure 1). This included both the original 156 items that were tested in the first round of CIs and those items (e.g. alternate versions of original items) that underwent additional rounds of testing. Of the 156 original items, 80 (51%) were not substantially revised, 22 (14%) were substantially revised, and 54 (35%) were deleted based on CI feedback. It should be noted that although 80 items were not substantially revised, minor edits were made when necessary to simplify the language of the items. The resulting 105 items were included in the candidate PLUS-M item bank. The item revision process is described below and examples are included in Table 2.

INSTRUCTIONS

Participants were asked about the usefulness of the instructions. Many participants reported that they did not carefully read the instructions prior to answering the PLUS-M questionnaire. Those that reviewed them often stated that the instructions were “good”, “helpful”, and “made sense.” Suggestions for improvement included shortening the instructions and addressing concerns of participants who may view the described activity (e.g., walking while carrying a small child) as unsafe. In response to this feedback, the instructions were simplified and formatted with bullet points to reduce respondent burden. In addition, language was added instructing respondents to choose “unable to do” for situations where safety concerns would restrict engaging in that activity.

RESPONSE OPTIONS

Participants were also asked to comment on the appropriateness of the response options. Most participants made comments about the response options, such as “they were reasonable” or “I didn’t notice any problems.” However, one participant suggested that respondents be asked if they had performed the activity and another stated that the addition of “not applicable” response option would be helpful to accommodate for activities that the respondent had not tried. It is understandable why respondents wished to have a “not applicable” or “do not know” response available. However, including such options in a survey instrument introduce issues related to scoring and do not enhance the quality of data

[23]. For these reasons, “not applicable” and “do not know” response options were not added.

CANDIDATE ITEMS

Participants provided detailed feedback on problematic candidate items. Common problems seen in items included issues with clarity, inappropriate assumptions, and double-barreled items. These problems are discussed in further detail below.

Item intent is unclear—Problems with item clarity were most commonly identified in PLUS-M candidate items. In such situations, respondents perceived items differently than what was originally intended by the investigators. For example, “Are you able to walk on a sideways incline (e.g. a sidewalk that slopes toward the street)?” was understood by respondents as a question about walking on a ramp or a curb cutout. However, the question was intended to assess walking on a surface when the feet are at unequal heights. Because the original item was so challenging to understand, it was rewritten as, “Are you able to walk on a surface that slants sideways where one side is higher than the other?” Subsequent rounds of CIs showed the intent of the revised item was more clearly understood by survey respondents.

Item assumptions are inappropriate.—Another problem identified in candidate items was an implied assumption that experiences related to the item content were similar across respondents. In reality, experiences differed in ways that markedly affected participants’ responses to some questions. For example, participants’ responses to the item, “Are you able to open and close windows in your home?” varied based on the type and condition of windows in their home. Some respondents replied that this task was quite difficult because their windows stuck, while others found it easier because they had newer windows. As such, responses were strongly influenced by the respondents’ experiences, rather than their mobility. This question was therefore deleted.

Items are double-barreled—Some items were identified as problematic because they were double-barreled, (i.e. they combined two different questions). Double-barreled items were challenging for participants to answer because they wished to respond differently to different parts of the item. For example, respondents reported challenges responding to the item “Are you able to walk when you are unable to see your feet or the ground (e.g. deep leaves or tall grass)?” They indicated that it was more difficult to walk in tall grass than it was to walk in deep leaves. Although both tall grass and deep leaves were reported to restrict view of the respondents’ feet, tall grass was noted to resist swing phase of the prosthetic limb far more than did deep leaves. Investigators therefore chose to split this question into two items, one that specifically addressed respondents’ ability to walk without looking at their feet and another that addressed their ability to walk when motion of the prosthesis was inhibited.

APPROPRIATENESS OF ITEMS FOR SUBPOPULATIONS OF PERSONS WITH LOWER LIMB LOSS

Persons with less than one year of experience—New prosthetic limb users varied in their responses to candidate PLUS-M items based on the amount of experience they had using a prosthesis. Two participants with between 0.7 years and 1.0 years experience easily related to the scenarios asked of them and thus had no difficulty responding to survey items. Two other participants with limited prosthetic experience (i.e. each with 0.3 years) reported that they had not experienced many of the activities and could not provide meaningful answers. One of these individuals stated multiple times that she would prefer to answer “not applicable” because she had never tried the activities. These two participants often envisioned performing the activities while sitting in a wheelchair, rather than while walking with their prosthetic limb. This suggested that activities performed recently after amputation required the use of wheeled assistive devices and that responses to items did not reflect the developers’ intent of measuring prosthetic mobility.

Persons with upper extremity involvement—Two individuals in this sample noted that their upper extremity involvement (i.e. injury or amputation of the upper limbs) affected the way that they answered PLUS-M items. One person with shoulder dysfunction stated that his shoulder injury, rather than his amputation, made “playing catch with a ball” difficult. Another person with upper limb loss noted that the difficulty that he reported “weeding a garden” was due to his upper limb loss and not mobility problems related to use of his lower limb prostheses. This feedback suggested that some items functioned differently in persons with lower limb loss *and* upper extremity involvement than they did in those with lower limb loss alone.

Persons with bilateral amputation—Prosthetic limb users with bilateral amputation also provided feedback about all candidate items during CIs. Interestingly, PLUS-M items were understood similarly both by participants with bilateral amputation and participants with unilateral amputation.

Persons with low literacy—The candidate items were tested in persons with either self-reported cognitive dysfunction and/or by persons who did not graduate from high school or pass the general education development (GED) exam. In general, items were read and understood by those in this group. One individual with self-reported cognitive dysfunction took longer to read and process items, which was reflected in a longer CI time (110 minutes) than other participants (16-94 minutes). However, once he finished reading the questions, he correctly understood their intent and discussed his thought process for deriving an answer with the interviewer. Although CI results suggested that the PLUS-M items were comprehended by persons with low literacy, the investigators also modified items whose Lexile® score was higher than 1100L to ensure that they would be understood by those at an eighth grade reading level [24].

DISCUSSION

CI resulted in content modifications that substantially improved the clarity, comprehensibility, and quality of PLUS-M items. Problems identified in CIs were addressed through revision and/or deletion of items that could not be modified to function as intended by the PLUS-M investigators. Results of this study suggest that minor changes in the wording and format of the items, instructions, and response options can substantially change respondents' interpretation of item content. Addressing problems identified by respondents (e.g. items that are unclear, make inappropriate assumptions, or are double-barreled) is an essential and useful step in instrument development. The importance of CIs is indicated by Federal Drug Administration (FDA) patient reported outcomes (PRO) guidelines that require documentation of content validity through qualitative studies with members of target respondents [25]. Data that show instrument items are understandable and appropriate for the intended population and use are considered evidence of content validity [26]. The FDA guidance therefore recommends use of CIs to evaluate respondents' understanding of the questions, instructions, and response options [25].

The CIs administered in this study also provided information about the usefulness of items for subpopulations within lower prosthetic limb users. Some items appeared to function differently for persons with limited prosthetic experience and for persons with upper extremity involvement. Persons with limited prosthetic experience reported that they have not been exposed to many of the activities described in the candidate items and therefore could not meaningfully respond. Persons with upper extremity involvement noted that answers to items that inquired about mobility tasks that included upper limb use were mainly driven by their upper limb concerns and not their lower limb mobility. Future research is therefore warranted to better understand the function of PLUS-M items in persons with limited prosthetic experience or upper extremity involvement. Until more evidence is available, use of this instrument in these populations should be approached with caution. The candidate items functioned appropriately for persons with bilateral lower limb loss, though additional research is needed to ensure the appropriateness of the PLUS-M for those in this population. Lastly, candidate PLUS-M items are appropriate for respondents with low literacy. However, the administration of the survey to persons with lower literacy may take longer than is typically expected and Lexile® readers below 1100L may experience challenges in comprehension with some items.

This summary of use of CIs to inform development of the PLUS-M item bank represents part of a larger project. This project to-date has included development of a conceptual framework, review of existing measures of mobility and physical function, and conduct of focus groups of prosthetic limb users (details regarding those methods will be published elsewhere). In addition, following CIs, the modified items were administered to a group of over 1000 prosthetic limb users to further evaluate the function of the items and to develop scoring for the PLUS-M.

LIMITATIONS

Thirty individuals with limb loss participated in the cognitive interviewing process, but only about five participants reviewed each individual item. Five participants cannot provide all

perspectives and interpretations. However, the goal of this qualitative step was to better understand how the items function when reviewed by individuals with a variety of backgrounds and experiences [8]. It is more important at this step in item development to have a range of experiences represented through a diverse sample than a large number of people [14-16]. Nevertheless, it is possible that CIs with additional respondents would have identified problems not identified by the participants in this study.

Study participants included individuals the investigators considered representative of persons with low literacy. Low literacy in this study was established via participants' self-report of educational experience and/or difficulties with thinking, memory or concentration. The investigators acknowledge that direct testing of literacy through established measures such as the Wide Range Achievement Test (WRAT) [27] or the Rapid Estimate of Adult Literacy in Medicine (REALM) [28] may provide more accurate assessments of literacy. Cognitive testing of PLUS-M items with persons deemed to be of low literacy based on scores from these or similar instruments should be considered in future investigations to verify that the survey is suitable for persons with low literacy.

Another potential limitation common to all CIs that use retrospective probing is participant recall [10]. A participant may not remember how they arrived at their response if follow-up questions are presented long after they respond to the tested item. To minimize this limitation, we chose to administer few items (i.e., 25-30) and held the interviews immediately after the participant filled out the questionnaire. These steps reduced the time between completing the questionnaire and answering the interview questions, thus increasing the likelihood that participants were able to remember and comment on their experience when responding to the item.

The study investigators recognize that CIs are only one step towards establishing evidence of a measure's content validity. Thus, additional work has been undertaken (e.g., advisory panel reviews, conceptual framework development, construct definition, focus groups, and literature review) but is not detailed in this manuscript. This work will be described in future publications regarding development of the PLUS-M instrument.

CONCLUSION

CIs were effective in identifying and resolving concerns with candidate PLUS-M items. Participant feedback resulted in considerable improvement of the measure and resulted in 105 well-functioning items for measuring self-reported mobility of lower limb prosthetic users. Although engagement in item development processes like CIs was time- and resource-consuming, taking these steps ensure that PLUS-M will be a valid and useful tool to measure mobility in prosthetic limb users. CIs substantially improved the relevance of the instrument to intended users, and ultimately make the measure more clinically meaningful.

Acknowledgments

This research is supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NIH grant number HD-065340). The authors wish to thank Rana Salem, MS, for performing the descriptive data analysis, Silvia Christian, BA, for performing the Lexile® analysis, and Meighan Rasley, BA, for scheduling participant interviews.

Funding/Support: This research is supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NIH grant number HD-065340).

REFERENCES

1. Sansam K, O'Connor R, Neumann V, Bhakta B. Can simple clinical tests predict walking ability after prosthetic rehabilitation? *Journal of Rehabilitation Medicine*. 2012; 44(11):968–974. doi: 10.2340/16501977-1046. [PubMed: 23037865]
2. Lin SJ, Bose NH. Six-minute walk test in persons with transtibial amputation. *Archives of Physical Medicine and Rehabilitation*. 2008; 89(12):2354–2359. doi:10.1016/j.apmr.2008.05.021. [PubMed: 18976979]
3. Deathe AB, Miller WC. The L Test of Functional Mobility: Measurement Properties of a Modified Version of the Timed “Up & Go” Test Designed for People With Lower-Limb Amputations. *Physical Therapy*. 2005; 85(7):626–635. [PubMed: 15982169]
4. Gailey RS, Roach KE, Applegate EB, Cho B, Cunniffe B, Licht S, Maguire M, et al. The amputee mobility predictor: an instrument to assess determinants of the lower-limb amputee's ability to ambulate. *Archives of Physical Medicine and Rehabilitation*. 2002; 83(5):613–627. [PubMed: 11994800]
5. Brooks D, Parsons J, Hunter JP, Devlin M, Walker J. The 2-minute walk test as a measure of functional improvement in persons with lower limb amputation. *Archives of Physical Medicine and Rehabilitation*. 2001; 82(10):1478–1483. doi:10.1053/apmr.2001.25153. [PubMed: 11588757]
6. Datta D, Ariyaratnam R, Hilton S. Timed walking test: an all-embracing outcome measure for lower-limb amputees? *Clinical Rehabilitation*. 1996; 10(3):227–232. doi: 10.1177/026921559601000307.
7. Stratford PW, Kennedy D, Pagura SMC, Gollish JD. The relationship between self- report and performance-related measures: questioning the content validity of timed tests. *Arthritis and Rheumatism*. 2003; 49(4):535–540. doi:10.1002/art.11196. [PubMed: 12910560]
8. Amtmann D, Cook KF, Johnson KL, Cella D. The PROMIS initiative: involvement of rehabilitation stakeholders in development and examples of applications in rehabilitation research. *Archives of Physical Medicine and Rehabilitation*. 2011; 92(10 Suppl):S12–19. doi:10.1016/j.apmr.2011.04.025. [PubMed: 21958918]
9. Condie E, Scott H, Treweek S. Lower limb prosthetic outcome measures: a review of the literature 1995 to 2005. *JPO: Journal of Prosthetics and Orthotics*. 2006; 18(6):P13.
10. Willis, GB. Course manual prepared for a short course at the 1999 Meeting of the American Statistical Association. Research Triangle Institute; 1999. *Cognitive Interviewing: A “How To” Guide*.
11. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*. 2003; 12(3):229–238.
12. Jobe JB. Cognitive psychology and self-reports: models and methods. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*. 2003; 12(3):219–227.
13. Presser S, Couper MP, Lessler JT, Martin E, Martin J, Rothgeb JM, Singer E. Methods for testing and evaluating survey questions. *Public opinion quarterly*. 2004; 68(1):109–130.
14. Willis, GB. *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications; 2005. Incorporated
15. DeWalt DA, Rothrock N, Yount S, Stone AA. Evaluation of item candidates: the PROMIS qualitative item review. *Medical Care*. 2007; 45(5 Suppl 1):S12–21. doi:10.1097/01.mlr.0000254567.79743.e2. [PubMed: 17443114]
16. Beatty PC, Willis GB. Research Synthesis: The Practice of Cognitive Interviewing. *Public Opinion Quarterly*. 2007; 71(2):287–311. doi:10.1093/poq/nfm006.
17. Gershon RC, Lai JS, Bode R, Choi S, Moy C, Bleck T, Miller D, et al. Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon

- clinical and general population testing. *Quality of Life Research*. 2012; 21(3):475–486. doi: 10.1007/s11136-011-9958-8. [PubMed: 21874314]
18. Cella D, Nowinski C, Peterman A, Victorson D, Miller D, Lai J-S, Moy C. The neurology quality-of-life measurement initiative. *Archives of Physical Medicine and Rehabilitation*. 2011; 92(10 Suppl):S28–36. doi:10.1016/j.apmr.2011.01.025. [PubMed: 21958920]
 19. Patient Reported Outcomes Measurement Information System (PROMIS®). [19 Mar 2013] PROMIS® instrument development and psychometric evaluation scientific standards. 2012. http://www.nihpromis.org/Documents/PROMIS_Standards_050212.pdf.
 20. Franchignoni F, Giordano A, Ferriero G, Orlandini D, Amoresano A, Perucca L. Measuring mobility in people with lower limb amputation: Rasch analysis of the mobility section of the prosthesis evaluation questionnaire. *Journal of Rehabilitation Medicine*. 2007; 39(2):138–144. doi: 10.2340/16501977-0033. [PubMed: 17351696]
 21. Ziegler-Graham K, MacKenzie EJ, Ephraim PL, Travison TG, Brookmeyer R. Estimating the prevalence of limb loss in the United States: 2005 to 2050. *Archives of Physical Medicine and Rehabilitation*. 2008; 89(3):422–429. doi:10.1016/j.apmr.2007.11.005. [PubMed: 18295618]
 22. Ephraim PL, Dillingham TR, Sector M, Pezzin LE, Mackenzie EJ. Epidemiology of limb loss and congenital limb deficiency: A review of the literature. *Archives of Physical Medicine and Rehabilitation*. 2003; 84(5):747–761. [PubMed: 12736892]
 23. Krosnick JA, Holbrook AL, Berent MK, Carson RT, Hanemann WM, Kopp RJ, et al. The impact of “no opinion” response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quarterly*. 2002; 66(3):371–403. doi:10.1086/341394.
 24. Stenner, AJ.; Horablin, I.; Smith, DR.; Smith, M. *The Lexile framework*. MetaMetrics. Inc.; Durham, NC: 1988.
 25. US Department of Health and Human Services (USDHHS). [1 Dec 2010] Guidance for industry. Patient-reported outcome measures: Use in medical product development to support labeling claims (US FDA guidance for industry related to the development and review of PRO measures). 2009. Retrieved from www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf.
 26. Patrick DL, Burke LB, Gwaltney CJ, Leidy NK, Martin ML, Molsen E, et al. Content validity—Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: Part 2—Assessing respondent understanding. *Value in Health*. 2011; 14(8):978–988. doi:10.1016/j.jval.2011.06.013. [PubMed: 22152166]
 27. Wilkerson, GS.; Robertson, GJ. *WRAT4 Wide Range Achievement Test Professional Manual*. Psychological Assessment Resources, Inc.; Lutz, FL: 2006.
 28. Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, Crouch MA. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Family Medicine*. 1993; 25(6):391–5. [PubMed: 8349060]

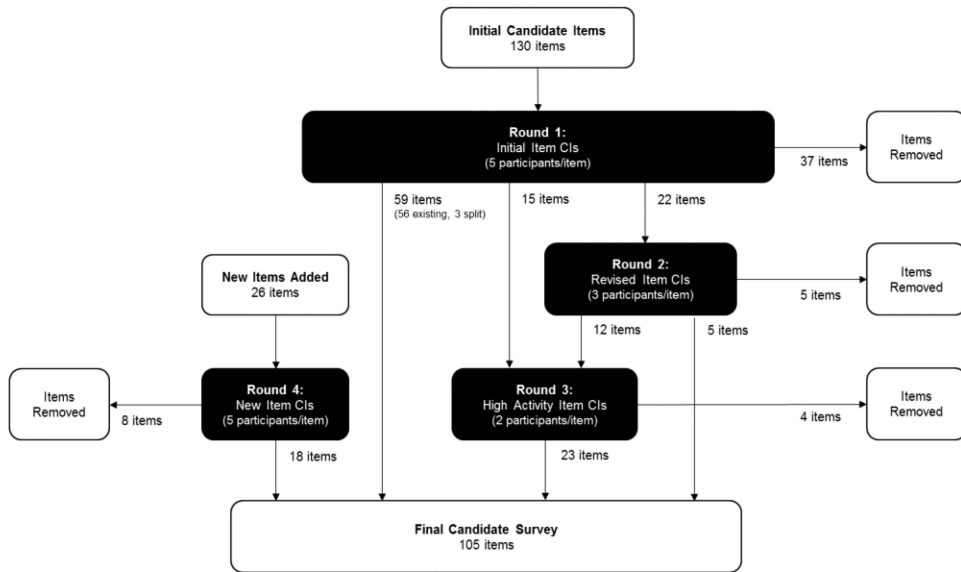


Fig. 1. Schematic depicting the CI process used in the development of the PLUS-M candidate item bank

Table 1

PLUS-M Instructions

Original instructions:	We want to know how well you can move around using your prosthetic leg. Please respond to all questions as if you were wearing the prosthesis you would normally use to perform the task. If you would normally use a device that helps you walk or balance (e.g. a cane, crutch, or walker) while performing the task, please answer the questions as if you were using that device. Please do not answer questions as if you are sitting in a wheelchair or receiving support from another person.
Final instructions:	Please respond to all questions as if you were wearing the prosthetic leg you use most days. If you would normally use a cane, crutch, or walker to perform the task, please answer the questions as if you were using that device. Please choose "unable to do" if you: 1)Would need help from another person to complete the task 2)Would need a wheelchair or scooter to complete the task 3)Feel the task may be unsafe for you

Table 2

Sample Questions and Examples of the Revision Process

Type of Problem	Original Question “Are you able to...”	Decision Made	Revised Item “Are you able to...”
Unclear intent	... walk on a sideways incline (e.g. a sidewalk that slopes toward the street)?	Item revised to clarify the surface characteristics intended by the developers.	... walk on a surface that slants sideways where one side is higher than the other?
	... walk in a narrow hallway?	Item was revised. Developers intended the question to reflect a situation where the respondent would need to narrow their base of support, which is not how respondents interpreted the question.	... walk with your feet close together.
Inappropriate assumptions	... open and close windows in your home?	Item was deleted. Respondents reported that the item reflected window function rather than mobility.	N/A
	... stand up on tiptoes?	Item was deleted. Respondents reported that standing on tiptoes is not appropriate for persons who have typical prosthetic feet.	N/A
Double-barreled	... walk when you are unable to see your feet or the ground (e.g. deep leaves or tall grass)?	Item was split into two independent items. Respondents reported that deep leaves were easier to walk in than tall grass.	... walk without looking down at your feet? ... walk in tall grass (e.g. a field)?
	... walk on slippery surfaces (e.g. wet tile, ice, a rainy street, or a wet deck)?	Item was revised. Respondents felt that the examples given differed in their difficulty level.	... walk on slippery surfaces (e.g. wet tile or ice)?
No problems identified	... walk on hardwood floors?	No changes made.	N/A

N/A = not applicable

Table 3

Participant Characteristics

Participant Characteristics (n=30)		Mean (SD)	Range
Age (yrs)		50.2 (13.9)	(26-77)
Years since first lower limb loss (yrs)		14.9 (13.9)	(0.4-52.3)
Prosthetics Evaluation Questionnaire – Mobility Subscale (PEQ-MS) *		2.9 (0.9)	(0.8-4.0)
		N	Percent
Gender	Female	10	33.3%
	Male	20	66.7%
Sample characteristics	Bilateral	10	33.3%
	Unilateral	20	66.7%
	Transfemoral	14	46.7%
	Transtibial	17	56.7%
Race/Ethnicity	Non-Hispanic White	18	60.0%
	Non-Hispanic Black	8	26.7%
	Hispanic	3	10.0%
Education	Some High School	6	20.0%
	High School Grad/GED	5	16.7%
	Some College	7	23.3%
	College Degree	6	20.0%
	Advanced Degree	4	13.3%
Employment Status	Employed	9	30.0%
	Retired	6	20.0%
	On disability	6	20.0%
	Unemployed	4	13.3%
	Homemaker	2	6.7%
	Other	2	6.7%
Cause of Limb Loss (May select multiple)	Trauma	15	50.0%
	Infection	10	33.3%
	Diabetes	6	20.0%
	Other	7	23.3%
Participant Location	Washington	16	53.3%
	Illinois	8	26.7%
	Alaska	1	3.3%
	Arizona	1	3.3%
	Florida	1	3.3%
	Ohio	1	3.3%
	Utah	1	3.3%

* Prosthetics Evaluation Questionnaire score (PEQ-MS) has a possible range of 0-4; higher scores indicate better functioning [20].