

Original Article
Medicine General & Policy



Patient-Reported Outcomes Measurement Information System: Translation and Linguistic Validation of Six Profile Domains for Korean Adults

Youngha Kim ,^{1*} Junghee Yoon ,^{2*} Nayeon Kim ,^{2,3} Mangyeong Lee ,^{1,4}
Danbee Kang ,^{1,2} Hye Yun Park ,^{2,5} Dongryul Oh ,⁶ Ki-Sun Sung ,⁷
Gee Young Suh ,⁸ Jin Seok Ahn ,⁹ and Juhee Cho ^{1,2,3,4}

OPEN ACCESS

Received: Apr 28, 2021

Accepted: Jul 18, 2021

Address for Correspondence:

Juhee Cho, PhD

Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, 81 Irwon-ro, Gangnam-gu, Seoul 06315, Korea.
E-mail: jcho@skku.edu

*Youngha Kim and Junghee Yoon equally contributed equally to this work.

© 2021 The Korean Academy of Medical Sciences.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Youngha Kim
<https://orcid.org/0000-0002-6002-9433>
Junghee Yoon
<https://orcid.org/0000-0003-1228-7435>
Nayeon Kim
<https://orcid.org/0000-0002-7515-347X>
Mangyeong Lee
<https://orcid.org/0000-0002-6266-5978>
Danbee Kang
<https://orcid.org/0000-0003-0244-7714>
Hye Yun Park
<https://orcid.org/0000-0002-5937-9671>
Dongryul Oh
<https://orcid.org/0000-0002-5643-5519>
Ki-Sun Sung
<https://orcid.org/0000-0003-1314-1171>

¹Center for Clinical Epidemiology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

²Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, Seoul, Korea

³Cancer Education Center, Samsung Medical Center, Seoul, Korea

⁴Department of Digital Health, SAIHST, Sungkyunkwan University, Seoul, Korea

⁵Division of Pulmonary and Critical Care Medicine, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

⁶Department of Radiation Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

⁷Department of Orthopedic Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

⁸Department of Critical Care Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

⁹Division of Hematology/Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

ABSTRACT

Background: The purpose of the study was to translate and linguistically validate a Korean language version of the PROMIS (K-PROMIS) for the six profile adult domains: Fatigue, Pain Intensity, Pain Interference, Physical Function, Sleep Disturbance, and Ability to Participate in Social Roles and Activities.

Methods: A total of 268 items were translated into Korean according to the Functional Assessment of Chronic Illness Therapy multilingual translation methodology. Participants first completed approximately 27 to 35 items and were then interviewed to evaluate the conceptual equivalence of the translation to the original English language source. The K-PROMIS items that met the a priori threshold of $\geq 20\%$ of respondents with comprehension difficulties in the cognitive interview.

Results: 54 of the 268 items were identified as difficult items to comprehend for at least 20% of respondents in Round 1. The most frequently identified K-PROMIS domain on difficult items to comprehend was the Physical function (24.5%). Most items with linguistic difficulties were Fatigue and Physical function. Cultural difficulties were only included the Physical function and Ability to Participate in Social Roles and Activities domains. 25 of 54 items were slightly revised, and then these revised items were tested with additional six participants in Round 2, and most participants had no problems to understand modified items.

Conclusion: The six profile adult domains of K-PROMIS have been linguistically validated. Further psychometric validation of the K-PROMIS items will provide additional information of meaningful outcomes for chronic disease and clinical setting.

Gee Young Suh <https://orcid.org/0000-0001-5473-1712>Jin Seok Ahn <https://orcid.org/0000-0002-2880-6730>Juhee Cho <https://orcid.org/0000-0001-9081-0266>**Funding**

This research was supported by a grant (18182MFDS407) from the Ministry of Food and Drug Safety in 2018, and Samsung Medical Center Research and Development Grant (SMX1210381).

Disclosure

The authors have no potential conflicts of interest to disclose.

Authors Contributions

Conceptualization: Kim Y, Yoon J, Kim N, Lee M, Kang D, Park HY, Oh D, Sung KS, Suh GY, Ahn JS, Cho J. Data curation: Kim Y, Yoon J. Formal analysis: Kim Y, Yoon J. Funding acquisition: Cho J. Investigation: Kim Y, Cho J. Methodology: Kim Y, Yoon J, Cho J. Project administration: Kim Y, Yoon J, Cho J. Resources: Cho J. Supervision: Cho J. Validation: Kim Y, Yoon J. Writing - original draft: Kim Y, Yoon J. Writing - review & editing: Kim Y, Yoon J, Kim N, Lee M, Kang D, Park HY, Oh D, Sung KS, Suh GY, Ahn JS, Cho J.

Keywords: Patient-Reported Outcomes; Patient-Reported Outcomes Measurement Information System; PROMIS; Validation; Korean Translation

INTRODUCTION

Over the years, patient-reported outcomes (PROs) have become important endpoints in clinical trials, health care, and the social sciences.^{1,2} PROs are health outcomes that are directly reported by the patient without interpretation by a clinician or anyone else. Many studies have confirmed that PROs should be considered as the standard for measuring perceived health condition, symptoms, functional status, and health-related quality of life to adequately guide treatments and improve health outcomes.³⁻⁶ In addition, PROs may provide information relevant to treatment choices, disease management, reimbursement decisions, and health policies in various clinical settings.^{7,8}

However, most PRO measurement instruments have encompassed problems such as a narrow scope of health status and difficulties in standardized interpretation owing to irrelevant, incomprehensible, or poorly formulated questions.⁹ Thus, in 2004, the Patient-Reported Outcomes Measurement Information System (PROMIS[®]) was established as a multi-center, collaborative project funded under the US National Institutes of Health Initiative to improve the measurement of clinically important symptoms and outcomes.¹⁰ The PROMIS is based on the conceptual framework advocated by the World Health Organization, and has three key domains: physical, mental, and social health.¹¹ These domains have a set of item banks to facilitate standardized assessment of PROs in adults and children. Initial development of the constructs and domains of PROMIS is detailed elsewhere.¹² The item banks for the measurement of a large number of clinically important outcomes can be administered either as short forms (fixed subsets of the best items from an item bank) or through computerized adaptive testing (where a computer algorithm selects items from an item bank one by one based on answers to previous items, leading to a variable subset of items per person). The item banks were developed using item response theory (IRT) methods, which model the relationship between the level of the domain and the probability of giving a certain response to an item. Based on the IRT model, items are ordered in an item bank based on the item difficulty and this ordering allows for the selection of subsets of items for short forms or computer adaptive tests (CATs) and ensures that all short forms and CATs are measured on the same scale and that scores can be compared directly.

The PROMIS has been translated and validated in various languages, and the domains of the adult item banks have been internationally adopted in clinical research and practice.¹³⁻²³ The PROMIS has the potential to be global health scale considering as different item parameters could be used across countries and because its psychometric properties enable efficient assessments, which is valuable in clinical research and the evaluation of medical care.²⁴⁻²⁶ Thus, we aim to translate and linguistically validate a Korean version of the PROMIS (K-PROMIS) for six profile adult domains.

METHODS**Translation procedure**

We translated and validated six of the eight profile domains of the PROMIS (“fatigue,” “pain intensity,” “pain interference,” “physical function,” “sleep disturbance,” and “ability to

participate in social roles and activities”; **Supplementary Table 1**) because the items in the depression and anxiety domains have already been linguistically validated. In addition, items available in Korean from Short Forms before this study (e.g., Short Form 4a items) were also excluded from the study.

Fig. 1 provides an overview of the translation and linguistic validation procedure. A total of 268 items from six profile domains were translated into Korean according to the Functional Assessment of Chronic Illness Therapy (FACIT) multilingual translation methodology including forward and back translations, independent reviews of translation quality, and pilot testing including cognitive interviews.^{27,28} This methodology is consistent with the guidelines for translation of PRO instruments and previously published translation and validation procedures of the PROMIS for non-English speaking populations. During the translation, we focused on semantic and conceptual equivalence of meaning with the original English PROMIS.

First, each English source item was forward translated into Korean by three independent bilingual professional translators who have licenses for Korean-English interpretation and translation. This step focused on capturing the essence of items rather than performing literal translations. Then, a fourth independent bilingual translator who was not involved with the forward translation process selected one of the three translations or created a hybrid version of the translated items. The reconciled version was then back translated by another set of three bilingual professional translators fluent in both English and Korean. Translators did not have medical background. While they were provided brief background information about PROMIS measurement system before the translations, no other information was

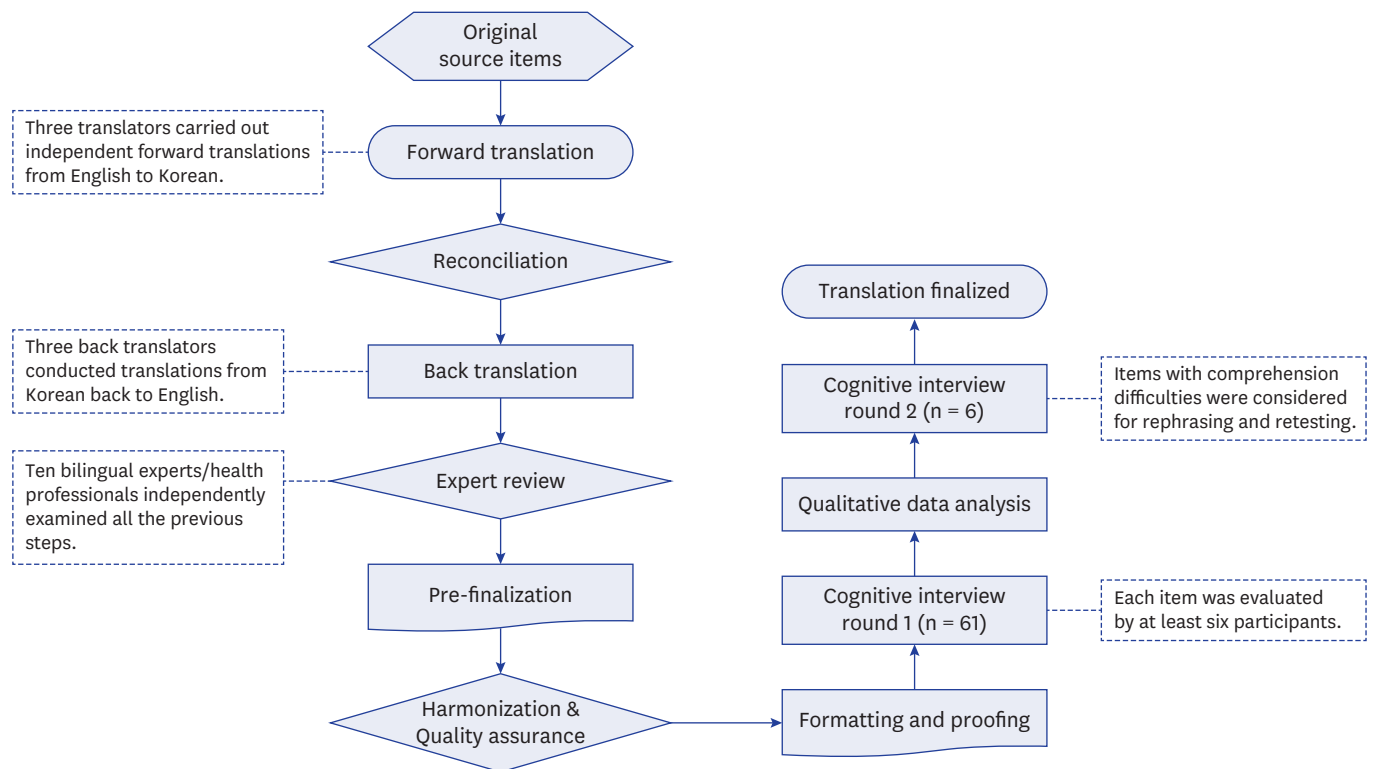


Fig. 1. Translation and linguistic validation process for the Korean version of the PROMIS profile domains.

provided to the translators. For quality assurance of the translation procedure, we organized an expert group consisting of 10 health professionals (six clinicians, two nurses, and two behavioral scientists) who were native speakers of Korean and fluent in English. The expert group reviewed the forward translations, reconciliations, and back translations based on the concept and definition provided by the PROMIS Statistical Center and identified inadequate expressions and concepts in the translation as well as any discrepancies between the original source items in English and the translated Korean items. If the previous translation steps were unacceptable, the experts selected the most appropriate translation for each item or provided alternate translations. In addition, the Korean Language and Culture Institute reviewed the translation for spelling errors and conceptual equivalence. The harmonization and quality assurance of the pre-final translation were performed by the PROMIS Statistical Center to address consistency and conceptual equivalence with existing translations in other languages if applicable, as well as between items.

Cognitive interviewing procedure

We conducted cognitive debriefing interviews to determine whether participants found any of the K-PROMIS items ambiguous or difficult to understand. This is an interview method designed to evaluate the comprehensibility, ease of response, and acceptability of the terminology, phrasing, response options, and format of a newly developed PRO measure.²⁹

Questionnaire scripts

Debriefing scripts were created to guide the cognitive interviews. Each interview was conducted using a standardized script consisting of open-ended questions developed by the researchers. Each script contained 27 to 35 items to avoid burdening the participants in a single interview. The interviewer asked a series of open-ended questions, following a script, seeking comments with regard to the item stem, the response options, and the time frame (e.g., in the past seven days, now, etc.). Prior PROMIS cognitive debriefing studies have successfully used a similar approach.^{18,23,30-32} Data concerning participants' sociodemographic characteristics, including marital status, educational level, and monthly household income, were also gathered by self-report. Clinical information was obtained from electronic medical records. We calculated the total time (in minutes) taken to complete the K-PROMIS survey, excluding any breaks taken during the interview.

Participants

Sixty-seven adults aged 18 years or older who had seen a physician for chronic diseases or pain within the past five years, who had no concurrent psychiatric condition or cognitive or other impairments (e.g., visual) that would interfere with completing an interview, and who could speak, read, and write Korean as their primary language were recruited. Prespecified enrollment goals included the following: at least 30% of participants with an educational level of high school or older than 65 years, approximately equal gender representation, and at least five participants per item.²⁸ Accrual of the sample was monitored prospectively to achieve these enrollment goals.

Interviews

Two rounds of cognitive interviews were conducted. Sixty-one participants were included in Round 1. After making the necessary revisions based on feedback from Round 1, six participants were included in Round 2. The cognitive interviews were performed in a private area of the outpatient clinic and consisted of the following two parts: administration of the K-PROMIS survey composed of a subset of K-PROMIS items, and a semi-scripted debriefing

interview with cognitive probing for comprehension, clarity, and ease of judgment. Interviews were conducted by two experienced interviewers in conducting cognitive interviews.

After completing the paper-and-pencil K-PROMIS survey, participants were asked to indicate items they found difficult to comprehend and those for which they had difficulty selecting a response option. The interviewer did not provide any assistance or advice and encouraged patients to complete the survey to the best of their ability on the basis of the instructions provided. After completing the survey, the cognitive debriefing interview was conducted. The script was similar to those used in previous PROMIS studies.^{30,31,33} Participants identified confusing text from the K-PROMIS items and rephrased them in their own words. The interviewer asked the participants to suggest alternative wording or phrasing to improve the clarity of items that were difficult to comprehend. The interviewer took detailed notes, and the interviews were recorded and transcribed for analysis.

Iterative cycles of analysis and retesting

For analysis of the individual K-PROMIS items, interview field notes and transcripts were compiled, abstracted, and summarized item by item. The interview data were examined for semantic and conceptual equivalence to the original English items. The expert group reviewed the results of each round of interview data analysis. The proportion of respondents exhibiting any level of difficulty or hesitation for an item or response option was calculated. K-PROMIS items that met the a priori threshold of $\geq 20\%$ of respondents with comprehension difficulties in Round 1 were considered for rephrasing and retesting in Round 2.

Items that at least 20% of the respondents found difficult to comprehend were classified under either linguistic or cultural difficulties. Linguistic difficulties were categorized into four issues: 1) “Do not know the meaning of the word” referred to a situation where the participants did not know or understand the meaning of a few words or sentences; 2) “Ambiguous meaning” referred to a situation where the participants knew the words but had a different interpretation than the intended meaning; 3) “Confused context” referred to a situation where the participants had an idea of the words but were confused about the exact meanings or were uncertain about the meanings in the correct context; and 4) “Others” referred to difficulties not included in the above three categories. Cultural difficulties were categorized into three issues: 1) “Differences in concepts or unfamiliar activities” referred to a situation where the participants did not understand the phrasing owing to conceptual differences or unfamiliar experiences; 2) “Differences in lifestyle and living environment” referred to a situation where the participants had difficulties responding because of differences in lifestyles and living conditions; and 3) “Different metric units” referred to a situation where the participants had difficulties understanding metric units.

Item revision was considered on the basis of a detailed review of participants' responses and in the context of an effort to produce a final version that would be well comprehended by diverse respondents, including those who were older and had lower educational levels. Following this, the PROMIS Statistical Center performed the final quality review.

Ethics statement

This study was approved by the Institutional Review Board (IRB) of the Samsung Medical Center (IRB approval No. SMC-2017-03-103-012), and written informed consent was obtained from all participants.

Table 1. Characteristics of cognitive interview participants (n = 67)

Characteristics	Overall (n = 67)	Round 1 (n = 61)	Round 2 (n = 6)
Age	57.1 (15.7)	57.1 (15.4)	56.7 (19.7)
Men	37 (55.2)	34 (55.7)	3 (50.0)
Educational level			
Below high school	34 (50.7)	28 (45.9)	6 (100.0)
Above college	33 (49.3)	33 (54.1)	-
Monthly household income (US dollars)			
< \$3,000	29 (43.3)	25 (41.0)	4 (66.7)
≥ \$3,000	38 (56.7)	36 (59.0)	2 (33.3)
Marital status			
Married	48 (71.6)	44 (72.1)	4 (66.7)
Single, divorced, separated, or widowed	19 (28.4)	17 (27.9)	2 (33.3)
Area of residence			
Seoul metropolitan area	39 (58.2)	33 (54.1)	6 (100.0)
Medical department			
Orthopedics	32 (47.8)	28 (45.9)	4 (66.6)
Pulmonology	14 (20.9)	13 (21.3)	1 (16.7)
Oncology	21 (31.3)	20 (32.8)	1 (16.7)

Values are presented as mean ± standard deviation or number (%).

RESULTS

A total of 67 outpatients participated in the cognitive interviews (61 participants in Round 1 and six participants in Round 2). The mean age was 57.1 years (age range: 18 to 86 years, standard deviation: 15.7), 55.2% were men, 71.6% were married, and 50.7% had an educational level of below high school (Table 1). Of the study participants, 38.8% either had an educational level of below middle school or were aged above 65 years.

Although participants understood most of the K-PROMIS items during the cognitive interview in Round 1 (n = 61), there were several items that were difficult to comprehend. Fifty-four (20.2%) of the 268 K-PROMIS items were identified as difficult to comprehend for at least 20% of the respondents (Table 2). Among these 54 items, four items had both linguistic and cultural difficulties. Most K-PROMIS items involved linguistic difficulties, but items under “physical function” and “ability to participate in social roles and activities” involved difficulties with regard to both. Among the items that were difficult to comprehend, the most frequently identified K-PROMIS domain was “physical function” (24.5%). In particular, items which more than 60% of participants had difficulty in comprehending fell under “physical function” (10 items), “fatigue” (five items), and “ability to participate in social roles and activities” (one item).

Table 2. Number of translated items presenting comprehension difficulties during the cognitive interview in Round 1

Domain	Number of translated items in bank	Overall > 20%	Item with difficulties ^a			
			Linguistic difficulties		Cultural difficulties	
			20–60%	> 60%	20–60%	> 60%
Fatigue	69	16 (23.2)	11 (15.9)	5 (7.2)	-	-
Pain intensity	3	-	-	-	-	-
Pain interference	36	4 (11.1)	4 (11.1)	-	-	-
Physical function	106	26 (24.5)	8 (7.5)	3 (2.8)	8 (7.5)	7 (6.6)
Sleep disturbance	23	3 (13.0)	3 (13.0)	-	-	-
Ability to participate in social roles and activities	31	5 (16.1)	4 (12.9)	-	-	1 (3.2)

Values are presented as number (%).

^aItems either noted by respondents as difficult to comprehend or identified through cognitive interview responses as posing problems with comprehension, clarity, or ease of response.

Linguistic and cultural difficulties were specifically categorized into four (Table 3) and three (Table 4) issues, respectively. Most items with linguistic difficulties were related to “fatigue” (11 items) and “physical function” (11 items). In terms of the issue “Do not know the meaning of the word,” participants were unfamiliar with the foreign words or Chinese characters, such as “mental energy,” “restrict,” or “feel limited.” Regarding “Ambiguous meaning,” four items (“household chores,” “feel less alert,” “flights of stairs,” and “labor”) in the “fatigue” and “physical function” domains presented difficulties for 60% or more of the participants. For instance, for the item “How often were you too tired to do your household chores?” five out of six participants (83.3%) said that they understood the meaning of the term but had no experience with household chores. Most participants (85.7%) understood “flights of stairs”

Table 3. Key cognitive interview findings regarding items with linguistic difficulties in Round 1^a

Source of difficulty	Domain	Source item (English)	% of participants with difficulties in Round 1 ^b (n = 61)	Examples of difficulties experienced by participants in Round 1	Resolution	Decision after Round 2
Do not know the meaning of the word	Fatigue	How much mental energy did you have on average?	33.3	Participants did not know what was meant by “mental energy.”	Phrasing retained as no suitable alternative was found.	
		How often did you feel your fatigue was beyond your control?	28.6	Participants did not understand what was meant by this sentence.	Phrasing retained as no suitable alternative was found.	
	Pain interference	How often did pain restrict your social life to your home?	28.6	Participants had difficulty understanding “restrict.”	Phrasing retained as no suitable alternative was found.	
		Sleep disturbance	I had trouble staying asleep.	28.6	Participants had difficulty understanding “staying asleep.”	Phrasing retained.
Ambiguous meaning	Fatigue	I feel limited in the amount of time I have for my family.	28.6	Participants had difficulty understanding “feel limited.”	Rephrased and retested.	Phrasing retained.
		How often were you too tired to do your household chores?	83.3	Participants had no experience in performing household chores.	Phrasing retained as no suitable alternative was found.	
		How often did your fatigue make you feel less alert?	66.7	Participants understood “feel less alert” to mean “distractibility.”	Rephrased and retested.	Phrasing retained.
		How often did your fatigue make you more forgetful?	50.0	Participants associated “more forgetful” with “amnesia.”	Phrasing retained as no suitable alternative was found.	
	Physical function	To what degree did your fatigue make you feel slowed down in your thinking?	33.3	Participants associated “feel slowed down in your thinking” with “amnesia.”	Rephrased and retested.	Phrasing retained as no suitable alternative was found.
		Are you able to climb up 5 flights of stairs?	85.7	Participants were confused by “five flights of stairs.” They thought of it as “five steps.”	Rephrased and retested.	Rephrased in a simpler manner.
		Does your health now limit you in doing eight hours of physical labor?	71.4	Participants understood “labor” in a broad sense.	Phrasing retained.	
		Does your health now limit you in bending, kneeling, or stooping?	57.1	Participants considered “bending” and “stooping” to mean the same thing and wanted to exclude duplicate expressions.	Rephrased and retested.	Phrasing retained as no suitable alternative was found.
		Are you able to wipe yourself after using the toilet?	42.9	Participants understood “using the toilet” in a broad sense.	Rephrased and retested.	Phrasing retained.
		Sleep disturbance	I had trouble sleeping.	28.6	Participants considered this the same as “I had trouble staying asleep.”	Rephrased and retested.

(continued to the next page)

Table 3. (Continued) Key cognitive interview findings regarding items with linguistic difficulties in Round 1^a

Source of difficulty	Domain	Source item (English)	% of participants with difficulties in Round 1 ^b (n = 61)	Examples of difficulties experienced by participants in Round 1	Resolution	Decision after Round 2
Confused context	Fatigue	How often did your fatigue make you feel slowed down in your thinking?	50.0	Participants understood “slowed down in your thinking” to mean “I couldn’t think.”	Rephrased and retested.	Phrasing retained.
		To what degree did you feel tired even when you hadn’t done anything?	42.9	Participants understood “hadn’t done anything” to mean “don’t have a job.”	Rephrased with a detailed explanation and did not retest.	
		How often did you have trouble starting things because of your fatigue?	33.3	Participants said that the answer could differ depending on the “things” in question.	Phrasing retained.	
		How often were you too tired to take a short walk?	33.3	Participants were confused about whether “short” was in reference to distance or time.	Phrasing retained as no suitable alternative was found.	
	Pain interference	How difficult was it for you to take in new information because of pain?	57.1	Participants were confused about whether “new information” was specific to pain or referred to general information.	Rephrased and retested.	Reverted to phrasing tested in Round 1.
		How much did pain interfere with your enjoyment of life?	28.6	Participants had difficulty conceptualizing “enjoyment of life.”	Phrasing retained.	
		How much did pain interfere with doing your tasks away from home (e.g., getting groceries, running errands)?	28.6	Participants understood “your task away from home” as “social activities.”	Phrasing retained.	
	Physical function	Are you able to walk up and down two steps?	85.7	Participants understood this as climbing two steps at a time.	Rephrased and retested.	Phrasing retained as the problem was with subjective interpretation.
		Are you able to wash your back?	57.1	Participants wanted a clearer meaning by using “by [your] self.”	Phrasing retained as it was identical to the English source item.	
		Are you able to exercise for an hour?	42.9	Participants said that the answer could differ depending on the type of exercise.	Phrasing retained as the question did not pose any limitations with regard to types of exercise.	
		Are you able to transfer from a bed to a chair and back?	42.9	Participants asked about different transfer methods, such as moving without standing up (on the floor).	Phrasing retained.	
		Are you able to lift one pound (0.5 kg) to shoulder level without bending your elbow?	28.6	Participants said that the answer could differ depending on the location of the object (i.e., on the table or on the floor).	Phrasing retained.	
		Are you able to carry a shopping bag or briefcase?	28.6	Participants understood “carry” to mean “carry a bag on one’s shoulders.”	Phrasing retained as this was a special case.	
		Are you able to use your hands, such as for turning faucets, using kitchen gadgets, or sewing?	28.6	Participants were confused by the word “sewing.” They said they could not sew because of poor vision.	Rephrased and retested.	
Sleep disturbance	I was afraid I would not get back to sleep after waking up.	28.6	Participants had difficulty understanding “get back to sleep.”	Phrasing retained as the problem was with subjective interpretation.		
Ability to participate in social roles and activities	I have to do my work for shorter periods of time than usual (include work at home).	28.6	Participants understood this as having to work quickly.	Phrasing retained.		
Other	Fatigue	How fatigued were you when your fatigue was at its worst?	66.7	Participants felt there was no connection between “at its worst” and “how fatigued were you?.”	Rephrased and retested.	Phrasing retained.

^aWe excluded items with less than 20% of respondents by source of difficulties among items with 20% or more of respondents with comprehension difficulties;

^bNumbers represent the proportion of participants who had difficulties with the item out of the total number of participants who were debriefed in cognitive interviewing.

Table 4. Key cognitive interview findings regarding items with cultural difficulties in Round 1^a

Source of difficulty	Domain	Source item (English)	% of participants with difficulties in Round 1 ^b (n = 61)	Examples of difficulties experienced by participants in Round 1	Resolution	Decision after Round 2
Differences in concepts or unfamiliar activities	Ability to participate in social roles and activities	I have to limit social activities at home.	71.4	Participants did not understand what was meant by “social activities at home.”	Phrasing retained and retested.	Phrasing retained as no suitable alternative was found.
Differences in lifestyle and living environment	Physical function	Are you able to cut your food using eating utensils?	85.7	Participants did not understand what was meant by “cut using eating utensils.” Koreans are familiar with spoons and chopsticks.	Rephrased and retested.	Rephrased using simpler words.
		Are you able to open a can with a hand can opener?	85.7	Participants thought of a tuna can opened using a pull tab.	Rephrased and retested.	Phrasing retained as no suitable alternative was found.
		Are you able to open previously opened jars?	57.1	Participants understood “jar” to mean “bottle.”	Phrasing retained.	
		Are you able to do yard work like raking leaves, weeding, or pushing a lawn mower?	42.9	Participants did not do “yard work.”	Phrasing retained as no suitable alternative was found.	
		Are you able to carry a laundry basket up a flight of stairs?	28.6	Most Korean houses are single storied, so there are no stairs to carry a laundry basket.	Rephrased and retested.	Phrasing retained.
Different metric units	Physical function ^c	Does your health now limit you in hiking a couple of miles (3 km) on uneven surfaces, including hills?	85.7	Participants had difficulty envisioning a distance that constituted “3 km.” They wanted suitable examples.	Phrasing retained.	
		Are you able to carry a heavy object (over 10 pounds/5 kg)?	71.4	Participants had difficulty visualizing a “heavy object (over 5 kg).” They wanted suitable examples.	Phrasing retained.	

^aWe excluded items with less than 20% of respondents by source of difficulties among items with 20% or more of respondents with comprehension difficulties;

^bNumbers represent the proportion of participants who had difficulties with the item out of the total number of participants who were debriefed in cognitive interviewing; ^cWe showed only high-proportion items with comprehension difficulties.

to mean “steps.” We conducted a retest by modifying the phrase to “climb five floors with several stairs,” but most participants still did not understand. Through further review, we modified the phrase “climb five floors with several stairs” to “climb up five floors.”

For the issues grouped under “Confused context,” the K-PROMIS domain that involved the most difficulties in comprehension was “physical function” (seven items). Most participants (85.7%) understood the items “walk up and down two steps” and “climb up five steps” to mean going up and down “two steps” and “five steps” at a time. We added the word “one step at a time” to make the intended meaning clearer, and then conducted the cognitive interview in Round 2. At this point, more than 80% of the participants understood the intended meaning.

Issues classified as “Others” were identified only under “fatigue.” We tried to use the same phrases or words that were previously translated in other domains in order to consistently match the meaning in whole domain in Korean. For instance, “How fatigued were you on average?” and “I feel fatigued” were translated into the Korean “How exhausted were you on average?” and “I was exhausted” in the FACIT-Fatigue scale. Of the participants, 66.7% said that there was no connection between “How fatigued were you” and “at its worst.”

Cultural difficulties were only identified in the “physical function” (seven items) and “ability to participate in social roles and activities” (one item) domains. In terms of the issue of “Differences in concept or unfamiliar activities,” regarding the item “I have to limit social

activities at home,” five out of seven participants (71.4%) said that they do not socialize at home. This was because in Korea, “social activities” refer to activities involving interaction with people outside the home.

For the issue of “Differences in lifestyle and living environment,” all problematic items fell under “physical function.” As an item relevant to a different lifestyle, 85.7% of the participants did not understand what “cut food using eating utensils” meant. Americans cut food using eating utensils such as knives and forks, whereas Koreans use spoons and chopsticks. After discussing this issue with the PROMIS Statistical Center, we rephrased this item to “cut food using spoons or chopsticks.”

Regarding the issue of “Different metric units,” many participants (85.7%) indicated that they found it difficult to judge exactly what distance “two miles (3 km)” or “10 miles (16 km)” represented and that they needed suitable examples. Koreans are familiar with the concept of judging distances in terms of the number of bus stops or blocks on the way to one’s destination, but the distances between bus stops or blocks can vary widely depending on the area of residence. For a similar reason, about 71.4% of the participants needed practical examples for measurement units like a “2 kg object” or “heavy object (5 kg).”

Of the 54 items, 25 were slightly revised after consultation with the PROMIS Statistical Center while 29 remained unchanged. These revised items were tested with an additional six participants in Round 2. During Round 2, most participants had no problems understanding the modified items. Three raters determined the types of issues associated with comprehension difficulties. Differences between raters were resolved by discussion until a consensus was reached.

DISCUSSION

In this study, we translated and linguistically validated the six adult profile domains of the PROMIS into Korean based on the standard FACIT multilingual translation methodology.^{27,28} The majority of items in the K-PROMIS were well comprehended by Korean speakers, including older adults and those with lower educational levels. After two rounds of cognitive debriefing, we concluded that the K-PROMIS possesses satisfactory psychometric properties for use in clinical research and the care of Korean-speaking patients. However, despite adhering to rigorous procedures to maintain semantic and conceptual equivalence with the original PROMIS items, some linguistic and cultural issues persisted.

Overall, the prevalence of linguistic issues in all domains, with the exception of “pain intensity,” ranged from 10% to 20%. However, only 15 items (14.1%) under “physical function” and one item (3.2%) under “ability to participate in social roles and activities” had cultural issues. This might be because, by conferring with the PROMIS Statistical Center, we were able to eliminate common cultural issues identified in PROMIS linguistic validation studies from other countries during the translation process. For example, measurement units (e.g., miles, blocks), definitions of “leisure activities” and “recreational activities,” and extended definitions of “yard work” for people living in apartments were discussed during translation. Yet, cultural adaptations were particularly required for “physical function” items in different living environments (e.g., yard work, carrying a laundry basket up) and tools used in daily life (e.g., utensils, can openers). In terms of metrics, while we employed those used in Korea (e.g., km, kg), patients still experienced difficulties because they do not use metrics in daily life. In Korea,

people are more likely use examples instead of exact metrics or units to express an amount or distance. For example, people prefer to use the number of steps or time taken to get to a destination (e.g., 30 minutes) than exact units (e.g., 1.5 km). This is similar to previous studies. In an Arabic version, “walking more than a mile” was replaced by “walking a distance more than a bus stop” because the mile is not a well-known measure in Arabic culture.³⁴

The PROMIS was developed quantitatively with scientific accuracy for the purpose of making an evidence-based objective outcome measure available in clinical practice.^{35,36} Yet, in this study, we found that individuals' experiences or situations affected their interpretations of the items. For example, regarding one item in the “fatigue” domain, “How often were you too tired to do your household chores?” some patients, especially older males, did not respond because they did not have experience of doing household chores. This is similar to a previous study, where patients could not answer some questions as they could not visualize the tasks/demands proposed.¹⁸ Thus, our study team considered including examples of household chores. However, the cognitive debriefing in the Dutch PROMIS validation study found that giving examples might lead to confusion or misunderstanding of items. For example, in the context of a question about physical activities, “Does your health now limit you in participating in active sports such as swimming, tennis, or basketball?,” patients were not able to think beyond these specific sports.³⁷ Thus, after conferring with the PROMIS Statistical Center, we decided not to include examples or reference points. The issue of how patients' interpretations can vary based on such uncertainties should be further explored.

The PROMIS item bank covers a broad range of levels because the items need to be evaluated based on IRT methods.³⁸ However, participants had difficulties distinguishing between similar idiomatic expressions and sentence constructions in the “fatigue” domain. There were 10 words (tired, fatigued, sluggish, run-down, physically drained, exhausted, bushed, totally drained, wiped out, extremely exhausted) but patients could not differentiate between them. This was even more problematic with people with limited vocabulary.^{23,30} Translation of these 10 different words to describe increasing levels of fatigue may have resulted in a different ordering (item difficulties) of the items in the translated language, potentially introducing differential item functioning. Further language-specific item calibration would be necessary if important differential item functioning is found.

Similarly, there were some words/expressions that, though slightly different, were considered the same by participants. The study team acknowledged this issue during forward translation and tried to choose the translation with the greatest conceptual equivalence. However, participants still had difficulties in distinguishing between words. For example, participants could not differentiate “flight of stairs” from “steps,” “labor” from “work,” and “bending” from “stooping.” While not being able to differentiate between these words would not pose a problem in daily life, this ambiguity is detrimental to the evaluation based on IRT methods. It might be helpful for patients if they are provided with the exact definitions of words, or given pictures as guidance for words with ambiguous meanings.

Furthermore, participants, especially those with low educational levels, had difficulties with loanwords such as “energy” or “recreational activities,” but we had to use them because they had no Korean equivalents. These items could not be substituted because the suggestions offered were mostly colloquial. This issue has been mentioned in previous studies involving testing PRO measures such as the Common Terminology Criteria for Adverse Events that have been translated into Korean.³⁹

This study has some limitations. First, as the setting was a single hospital, the participants were not representative of all Koreans. However, the sample possessed the demographic and clinical diversity to provide rigorous evidence of content validity in clinical settings. Therefore, additional testing to confirm its comprehensibility and cultural acceptability with regard to the general population with no experience of symptoms could be considered. Second, some items could not be modified because of the absence of Korean substitutes. Especially, four items included in the standard short forms were identified as difficult to comprehend for at least 20% of the respondents but items could not be modified as there were no suitable alternatives. Therefore, it is necessary to encourage clinicians and researchers who use the translated item banks in the target countries to share any culturally important content that might be missing. Furthermore, it might be reasonable to select different items in the standard short forms for use in Korea as these items could be problematic for Korean population. It is, therefore, necessary to encourage clinicians and researchers who use the translated item banks in the target countries to share any culturally important content that might be missing. Third, we conducted cognitive interviews with only six to seven participants per item; it might be better to increase this number to assess the level of comprehensibility and cognitive equivalence of the translation. However, this is higher than the number of participants recommended in the FACIT translation methodology for the debriefing of new items in the target country (five participants). In addition, we had an expert group who participated in the entire process, from linguistic validation to quality assurance to validation.

Regardless of these limitations, this study has several strengths. First, we performed purposive sampling to include people with low health literacy and cognitive abilities. Of the sample, 38.8% either had an educational level of high school or below or were older than 65 years, and we sought an approximately equal gender representation. Second, we included patients who were experiencing symptoms and difficulties included in the PROMIS items. This increased the validity of the study as individuals without symptoms might have had more difficulties in comprehending PRO items. For example, orthopedic patients would report problems with physical function more accurately than people who do not experience those limitations or symptoms. Next, we matched the words and phrases to increase the consistency of meaning within and between domains, except for the mental health profile, for which linguistic validation had already been performed by another Korean team.¹⁴ In addition, we designed common questionnaires and principles for the cognitive interviews and coding methods to reduce bias among interviewers.³⁹ Lastly, we followed standard methodology and guidelines, including the item definitions provided by the PROMIS team, and tried to achieve consensus with the PROMIS Statistical Center throughout the process to ensure measurement invariance across languages.

Our study can contribute to the formation of a basis for the international measurement of PROs and formal process of PROMIS adoption across countries. Further psychometric validation of the K-PROMIS items will provide additional information regarding meaningful outcomes for chronic diseases and clinical settings.

ACKNOWLEDGEMENTS

We thank the participants who participated in this study.

SUPPLEMENTARY MATERIAL

Supplementary Table 1

Translated Korean PROMIS item banks for adults

[Click here to view](#)

REFERENCES

1. U.S. Food and Drug Administration. Guidance for industry: patient-reported outcomes measures: use in medical product development to support labeling claims. <https://www.fda.gov/media/77832/download>. Updated 2009. Accessed December 23, 2019.
2. Trotti A, Colevas AD, Setser A, Basch E. Patient-reported outcomes and the evolution of adverse event reporting in oncology. *J Clin Oncol* 2007;25(32):5121-7.
[PUBMED](#) | [CROSSREF](#)
3. Chen J, Ou L, Hollis SJ. A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Serv Res* 2013;13(1):211.
[PUBMED](#) | [CROSSREF](#)
4. Howell D, Molloy S, Wilkinson K, Green E, Orchard K, Wang K, et al. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. *Ann Oncol* 2015;26(9):1846-58.
[PUBMED](#) | [CROSSREF](#)
5. Denis F, Basch E, Septans AL, Bennouna J, Urban T, Dueck AC, et al. Two-year survival comparing web-based symptom monitoring vs routine surveillance following treatment for lung cancer. *JAMA* 2019;321(3):306-7.
[PUBMED](#) | [CROSSREF](#)
6. Ishaque S, Karnon J, Chen G, Nair R, Salter AB. A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). *Qual Life Res* 2019;28(3):567-92.
[PUBMED](#) | [CROSSREF](#)
7. Snyder CF, Jensen RE, Segal JB, Wu AW. Patient-reported outcomes (PROs): putting the patient perspective in patient-centered outcomes research. *Med Care* 2013;51(8 Suppl 3):S73-9.
[PUBMED](#) | [CROSSREF](#)
8. Ahmed S, Berzon RA, Revicki DA, Lenderking WR, Moinpour CM, Basch E, et al. The use of patient-reported outcomes (PRO) within comparative effectiveness research: implications for clinical practice and health care policy. *Med Care* 2012;50(12):1060-70.
[PUBMED](#) | [CROSSREF](#)
9. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012;21(8):1305-14.
[PUBMED](#) | [CROSSREF](#)
10. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The Patient-reported outcomes measurement information system (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007;45(5 Suppl 1):S3-11.
[PUBMED](#) | [CROSSREF](#)
11. Riley WT, Rothrock N, Bruce B, Christodolou C, Cook K, Hahn EA, et al. Patient-reported outcomes measurement information system (PROMIS) domain names and definitions revisions: further evaluation of content validity in IRT-derived item banks. *Qual Life Res* 2010;19(9):1311-21.
[PUBMED](#) | [CROSSREF](#)
12. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol* 2010;63(11):1179-94.
[PUBMED](#) | [CROSSREF](#)
13. Haan EA, Terwee CB, Van Wier MF, Willigenburg NW, Van Deurzen DF, Pisters MF, et al. Translation, cross-cultural and construct validity of the Dutch-Flemish PROMIS® upper extremity item bank v2.0. *Qual Life Res* 2020;29(4):1123-35.
[PUBMED](#) | [CROSSREF](#)

14. Choi H, Kim C, Ko H, Park CG. Translation and validation of the Korean version of PROMIS® pediatric and parent proxy measures for emotional distress. *J Patient Rep Outcomes* 2019;3(1):36.
[PUBMED](#) | [CROSSREF](#)
15. Devine J, Schröder LA, Metzner F, Klasen F, Moon J, Herdman M, et al. Correction to: Translation and cross-cultural adaptation of eight pediatric PROMIS® item banks into Spanish and German. *Qual Life Res* 2018;27(11):3057.
[PUBMED](#) | [CROSSREF](#)
16. Westmoreland K, Reeve BB, Amuquandoh A, van der Gronde T, Manthalu O, Correia H, et al. Translation, psychometric validation, and baseline results of the Patient-Reported Outcomes Measurement Information System (PROMIS) pediatric measures to assess health-related quality of life of patients with pediatric lymphoma in Malawi. *Pediatr Blood Cancer* 2018;65(11):e27353.
[PUBMED](#) | [CROSSREF](#)
17. Devine J, Schröder LA, Metzner F, Klasen F, Moon J, Herdman M, et al. Translation and cross-cultural adaptation of eight pediatric PROMIS® item banks into Spanish and German. *Qual Life Res* 2018;27(9):2415-30.
[PUBMED](#) | [CROSSREF](#)
18. Schnohr CW, Rasmussen CL, Langberg H, Bjørner JB. Danish translation of a physical function item bank from the Patient-Reported Outcome Measurement Information System (PROMIS). *Pilot Feasibility Stud* 2017;3(1):29.
[PUBMED](#) | [CROSSREF](#)
19. Jakob T, Nagl M, Gramm L, Heyduck K, Farin E, Glattacker M. Psychometric properties of a German translation of the PROMIS® depression item bank. *Eval Health Prof* 2017;40(1):106-20.
[PUBMED](#) | [CROSSREF](#)
20. Haverman L, Grootenhuis MA, Raat H, van Rossum MA, van Dulmen-den Broeder E, Hoppenbrouwers K, et al. Dutch-Flemish translation of nine pediatric item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS)®. *Qual Life Res* 2016;25(3):761-5.
[PUBMED](#) | [CROSSREF](#)
21. Silva EC, Pinto Rde M, Mendonca TM, Silva CH. Brazilian-Portuguese translation and cultural adaptation of the sleep and wake disturbances domains of the Patient-reported-outcomes measurement information system (PROMIS). *Cad Saude Publica* 2014;30(7):1391-401.
[PUBMED](#) | [CROSSREF](#)
22. Alves FS, Pinto RM, Mendonça TM, Silva CH. Portuguese-language translation and cross-cultural adaptation of the fatigue domain of patient-reported-outcomes measurement information system (PROMIS). *Cad Saude Publica* 2014;30(5):1103-10.
[PUBMED](#) | [CROSSREF](#)
23. Terwee CB, Roorda LD, de Vet HC, Dekker J, Westhovens R, van Leeuwen J, et al. Dutch-Flemish translation of 17 item banks from the patient-reported outcomes measurement information system (PROMIS). *Qual Life Res* 2014;23(6):1733-41.
[PUBMED](#) | [CROSSREF](#)
24. Bevans M, Ross A, Cella D. Patient-reported outcomes measurement information system (PROMIS): efficient, standardized tools to measure self-reported health and quality of life. *Nurs Outlook* 2014;62(5):339-45.
[PUBMED](#) | [CROSSREF](#)
25. Fischer HF, Wahl I, Nolte S, Liegl G, Brähler E, Löwe B, et al. Language-related differential item functioning between English and German PROMIS depression items is negligible. *Int J Methods Psychiatr Res* 2017;26(4):e1530.
[PUBMED](#) | [CROSSREF](#)
26. Choi JS, Park YS, Kim JA, Park CS. International trends on patient-reported outcome measures for improving care quality and its implication for South Korea: focus on OECD PaRIS. *Qual Improv Health Care* 2019;25(1):11-28.
[CROSSREF](#)
27. Bonomi AE, Cella DF, Hahn EA, Bjordal K, Sperner-Unterweger B, Gangeri L, et al. Multilingual translation of the functional assessment of cancer therapy (FACT) quality of life measurement system. *Qual Life Res* 1996;5(3):309-20.
[PUBMED](#) | [CROSSREF](#)
28. Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Eval Health Prof* 2005;28(2):212-32.
[PUBMED](#) | [CROSSREF](#)
29. Willis GB. *Cognitive Interviewing: a Tool for Improving Questionnaire Design*. New York, NY, USA: SAGE Publications; 2004.

30. Solorio R, Ayala NC, Paez E, Skalicky AM, Morales LS. Use of cognitive interviews to adapt PROMIS measurement items for Spanish speakers living with HIV. *Aids Res Treat* 2016;2016:8340863.
[PUBMED](#) | [CROSSREF](#)
31. Christodoulou C, Junghaenel DU, DeWalt DA, Rothrock N, Stone AA. Cognitive interviewing in the evaluation of fatigue items: results from the patient-reported outcomes measurement information system (PROMIS). *Qual Life Res* 2008;17(10):1239-46.
[PUBMED](#) | [CROSSREF](#)
32. Liegl G, Rose M, Correia H, Fischer HF, Kanlidere S, Mierke A, et al. An initial psychometric evaluation of the German PROMIS v1.2 physical function item bank in patients with a wide range of health conditions. *Clin Rehabil* 2018;32(1):84-93.
[PUBMED](#) | [CROSSREF](#)
33. Liu Y, Hinds PS, Wang J, Correia H, Du S, Ding J, et al. Translation and linguistic validation of the pediatric patient-reported outcomes measurement information system measures into simplified Chinese using cognitive interviewing methodology. *Cancer Nurs* 2013;36(5):368-76.
[PUBMED](#) | [CROSSREF](#)
34. Mahmoud GA, Rady HM, Mostafa AM. Cross cultural adaptation and validation of an Arabic version of selected PROMIS measures for use in rheumatoid arthritis patients. *Egypt Rheumatol* 2019;41(3):177-82.
[CROSSREF](#)
35. Rhee Y, Jun S, Choi SE. Concepts and applications of patient-reported outcomes & quality of life measure: practical recommendations for Korea. *J Health Technol Assess* 2015;3(1):48-58.
36. Cho J, Yoon J, Kim Y, Park S. *Development and Clinical Validation of Patient-Reported Outcome Measurements for Chronic Disease Patients for Precision Medicine*. Cheongju, Korea: Korea Ministry of Food and Drug Safety; 2018.
37. Oude Voshaar MA, Ten Klooster PM, Taal E, Krishnan E, van de Laar MA. Dutch translation and cross-cultural adaptation of the PROMIS® physical function item bank and cognitive pre-test in Dutch arthritis patients. *Arthritis Res Ther* 2012;14(2):R47.
[PUBMED](#) | [CROSSREF](#)
38. Bae JM. Indices for the responsiveness and interpretability in patient-reported outcomes. *Korean J Fam Pract* 2015;5(3):161-6.
39. Cho J, Yoon J, Kim Y, Oh D, Kim SJ, Ahn J, et al. Linguistic validation of the US national cancer institute's patient-reported outcomes version of the common terminology criteria for adverse events in Korean. *J Glob Oncol* 2019;5(5):1-10.
[PUBMED](#) | [CROSSREF](#)