Cross-cultural adaptation and construct validity of the Korean version of a physical activity measure for community-dwelling elderly

BONGSAM CHOI, PT, MPH, PhD

1) Department of Physical Therapy, Woosong University: 17-2 Jayang-dong, Dong-gu, Daejeon 34606, Republic of Korea

Abstract. [Purpose] This study aimed to cross-cultural adapt and validate the Korean version of a physical activity measure (K-PAM) for community-dwelling elderly. [Subjects and Methods] One hundred and thirty-eight community-dwelling elders, 32 males and 106 female, participated in the study. All participants were asked to fill out a fifty-one item questionnaire measuring perceived difficulty in the activities of daily living (ADL) for the elderly. One-parameter model of item response theory (Rasch analysis) was applied to determine the construct validity and to inspect item-level psychometric properties of 51 ADL items of the K-PAM. [Results] Person separation reliability (analogous to Cronbach’s alpha) for internal consistency was ranging 0.93 to 0.94. A total of 16 items was misfit to the Rasch model. After misfit item deletion, 35 ADL items of the K-PAM were placed in an empirically meaningful hierarchy from easy to hard. The item-person map analysis delineated that the item difficulty was well matched for the elderlies with moderate and low ability except for high ceilings. [Conclusion] Cross-cultural adapted K-PAM was shown to be sufficient for establishing construct validity and stable psychometric properties confirmed by person separation reliability and fit statistics.

Key words: Cross cultural adaptation, Item response theory, Validity

INTRODUCTION

Functional status assessment is an essential component of geriatric care. The community-dwelling frail elderly is more likely than healthy elderly to have underlying comorbidities and impairments that may increase the risk of functional decline1,2. Rehabilitation professionals who care for the elderly should be able to detect any changes in functional status by the use of effective assessments. A myriad of assessments measuring activity of daily living (ADL) has been developed and elaborated to assist rehabilitation professionals in detecting any changes. While designed to approximate the capability to accomplish functional tasks in relation to the ADLs, many of those assessments fail to reflect the multitude of functional changes in the elderly population3. In addition, those assessments may or may not be sensitive to a broad range of the elderly’s ability levels although they have adequate psychometric properties. That is, the assessments commonly target the average person. Accordingly, they tend to only be sensitive at the center of a broad spectrum of the ability level4.

In contrast to classical test (CTT) theory-based assessments, item response theory (IRT)-based assessments focus on the psychometric properties of the items of the assessment, rather than the assessment as a whole. This approach would allow one to estimate the probability of selecting a particular response category of assessment items (e.g., no difficulty, some difficulty, or a lot of difficulty). These estimates of person ability are now invariant regardless of the item used within an assessment. In addition, IRT places item difficulty and person ability (i.e., both are presented in terms of a unit of measure called a “logit”) onto the same linear scale and links the difficulty level to person ability5. The item-person map delineates how
the distributions of persons and items relate to one another. By way of explanation, one may anticipate that a person would experience greater difficulty in estimating person ability as well as item locations using the map. Likewise, Rasch analysis (i.e., one-parameter IRT model) permits to examine the validity of an assessment by carefully scrutinizing item ordering (i.e., item difficulty).

ICFmeasure.com activity measure is a prototype web-based computer adaptive testing program that aims for measuring activity limitations of community dwelling elderly. It was created through the procedures of focus group presentation, professional panel verification, cognitive interviewing and paper/pencil field test. The assessment with 255 item pool appears to possess adequate psychometric properties as an activity measure. The item-level psychometric properties, person and item characteristic, are estimated by the Rasch model. The estimated measures are presented as a unit of measurement called a logit (i.e., log-odds unit). The logit scale is invariant. That is, the scale never change to whomever one may apply and with whichever the measures assess. The activity measure is a reliable and clinically useful measure in relation to measurement efficiency. The Korean version of the physical activity measure (K-PAM) is a translated version of the activity measure of which all assessment items from the item pool of icfmeasure.com target the wide range of ADL performances for the community-dwelling elderly.

The purpose of this study is to investigate the item-level measurement qualities of the K-PAM and the novel validation properties

SUBJECTS AND METHODS

One hundred and thirty-eight community-dwelling elderlies (32 male and 106 female, mean age 75.99 ± 7.23) were recruited from local health promotion centers in Dong-gu, Daejeon city, Korea between November 12, 2013 and November 27, 2013. The participants were selected when agreeing to the study procedure and asked to fill out the K-PAM in relation to their basic ADL. Written informed consent was obtained from all participated elderlies prior to administering the questionnaire, which was approved by the Institutional Review Board at Woosong University, College of Health and Welfare. The questionnaire consists of a total of 155 items, comprising 44 items to capture functional activities, 21 items to capture upper extremity function, 24 items to capture hand function, 16 items to capture walking, and 51 items to capture basic ADLs. For the cross cultural adaptation process of the translation, one physical therapist who has 21 years of clinical experience in US and Korea performs conceptual reviews, suitability for cultural context of Korea. With the procedure, a total of 100 items of the original English version was excluded from the translation due to unsuitable for the cultural difference. A pilot test was given to a total of 10 participants and followed by clinician’s review for proof-reading. All items were scored on a four-point scale: 1-severely impaired, 2-moderately impaired, 3-mildly impaired and 4-not impaired or no opportunity was given.

The items of ADL were analyzed with Winstep Rasch analysis computer program using the rating scale model. Person separation reliability (analogous to Cronbach’s alpha) was used for internal consistency of the ADL measure of the K-PAM. Construct validity was examined by fit statistics and item-person map from the Rasch analysis.

RESULTS

As a conventional psychometric statistical test to examine whether the translated version of the ADL items adequately measure what they were originally intended to measure (i.e., basic ADL), person separation reliability (i.e., Cronbach’s alpha) was used. It was 0.93 prior to the deletion of 16 misfit items and 0.94 following the deletion. In addition, fit statistics generated by Rasch analysis were examined to inspect the item-level analysis. Using the criterion of misfit as mean square value >1.4 or <0.6, a total of 16 of 51 items were misfit while 35 items were fit to the model. Furthermore, the hierarchical structure of the item difficulty calibration matched person ability measures fairly well on elderlies with moderate and low ability, while high ceiling effects were observed with nearly half of participated elderlies with high ability in item-person map (Fig. 1). Figure 1 showed the skewed to right distribution of person and a relative normal distribution of item. The measures were ranged between −2 and 5 logits for person and between −2 and 2 for item difficulty. In addition, the items of swallowing foods, using spoon, using spoon for soup were rated as the easiest activity task, while the items of tying shoelaces, putting winter boots and drying lower back were rated as the most difficult activity task, while the items of swallowing foods, using spoon, using spoon for soup were rated as the easiest activity task.

DISCUSSION

As an initial means to determine the ADL items of the K-PAM, translated version of it showed an excellent internal consistency determined by person separation reliability. The Rasch analysis of 51 ADL items revealed that 16 items were misfit (i.e., >1.4 or <0.6 for fit statistics). Furthermore, the item-person map analysis showed that 35 items were well targeted the elderlies with moderate and low ability despite the high ceilings.

The person separation reliability showed excellent internal consistency ranged from 0.93 to 0.94. This is analogous to Cronbach’s alpha that shows how well the set of assessment items measure a single construct. Similarly, dimensionality of the ADL items was examined by the Rasch fit statistics. Fit statistics is goodness-of-fit statistics that can be used to perform confirmatory construct validity of items. In the present study, a total of 16 items were misfit to the Rasch model despite

J. Phys. Ther. Sci. Vol. 30, No. 1, 2018 178
Fig. 1. Item-person map to determine construct validity of the KAM following the deletion of 16 misfit items.

The map represents the relationship between person ability (left side of the dotted line) and item difficulty measures (right side of the dotted line) in logits. Each “.” on the left side of the map represents one elderly and “#” represents three elderly, with those at the bottom of the scale representing elderlies of low ability and those at the top of the scale representing elderlies with high ability. The ADL items of the KAM at their average measure are listed to the right of the map, with the easiest items at the bottom of the map and hardest items at the top of the map. “M” to the left/right of the dotted line represents average item and person measure, respectively. The map is anchored the average person ability measure in order to compare item difficulties. The description of 36 items included in the Rasch analysis is listed on the far right of the map with item difficulty.
the excellent internal consistency. This would mean that those 16 items were measuring the different latent trait rather than ADL. In other words, the internal consistency as a conventional way of examining construct validity was unable to perform item-level analysis that IRT model commonly does. One reason can be postulated for such a high internal consistency measure although 16 items were misfit. This may be resulted from a large numbers of 51 ADL items since Cronbach’s alpha is a function of the number of items.

Since these items were originally created based on the activity dimension of International Classification of Functioning, Disability and Health (ICF) by World Health Organization, the conceptual framework and classification system for developing items used in the study applied to activities involving movement, moving around and daily life tasks as defined by the activity dimension of the ICF[7]. As a means to determine to visually inspect the construct validity of ADL items of the K-PAM, item-person map analysis was performed. The empirically derived evidence generated by the Rasch analysis supported principles of motor control theories. The conceptual difficulty hierarchy of the basic ADL items was well delineated by the map. After 16 item deletion, remained 35 ADL items of the K-PAM were placed in an empirically meaningful hierarchy from easy to hard. In the present study, tying shoelaces task was the most difficult. Higher motor control may be conceptually necessary to performing the task. For instance, to initiate tying shoelaces activity, it is necessary to plan to reach down shoelaces, hold/pull the laces, and followed by tying the knot. Thus, a coordinated series of functional movements is critical among these activities. In contrast to the difficult activity, swallowing activity, conceptually considered as an easy task, may require less controlled movements. Delva et al., in a study of functional decline in dementia, found similar hierarchical structures of ADLs in which bathing task would require more complex coordination in comparison to other ADL tasks such as feeding. Overall, the hierarchy of item difficulty matched person ability measures fairly well except for the high ceilings.

In addition to the validity, there are high ceilings with nearly 20 percent (i.e., 27 of 138 elderlies) of participated elderlies. These high ceilings are of concern. First, why are there such a lot of elderlies at the ceiling and if 51 items are not enough to target the wide range of ability, what can be done to measure the ceilings? This may be a result of the skewed distribution of person. That is, since many participated elderlies were recruited from several health promotion centers, there may be a higher chance of remaining physically active. While originally designed to measure frail elderlies who reside in community, those items appeared to be perceived as easy tasks by the elderlies with high ability. Future studies should include various levels of functional status of the community-dwelling elderly. As previously stated, cross-cultural adapted Korean version of the activity measure is sufficient for establishing construct validity and stable psychometric properties confirmed by person separation reliability and fit statistics. These results can later be a meaningful evidence for creating a Korean version of web-based measurement system with a large item pool measuring the activity dimension.

Funding
This research was funded by the Woosong University Research Fund.

Conflict of interest
The author declares that there is no conflict of interest.

REFERENCES
1) Collard RM, Boter H, Schoevers RA, et al.: Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc, 2012, 60: 1487–1492. [Medline] [CrossRef]
2) Cesari M, Vellas B, Hse FC, et al. LIFE Study Group: A physical activity intervention to treat the frailty syndrome in older persons-results from the LIFE-P study. J Gerontol A Biol Sci Med Sci, 2015, 70: 216–222. [Medline] [CrossRef]
3) Bousaziz W, Lang PO, Schmitt E, et al.: Health benefits of multicomponent training programmes in seniors: a systematic review. Int J Clin Pract, 2016, 70: 520–536. [Medline] [CrossRef]
4) McHorney CA: Generic health measurement: past accomplishments and a measurement paradigm for the 21st century. Ann Intern Med, 1997, 127: 743–750. [Medline] [CrossRef]
5) Velozo CA, Peterson EW: Developing meaningful Fear of Falling Measures for community dwelling elderly. Am J Phys Med Rehabil, 2001, 80: 662–673. [Medline] [CrossRef]
6) Bong TG, Fox CM: Applying the Rasch model: fundamental measurement in the human science. Mahwah: Lawrence Erlbaum Associates Publishers, 2001, pp 23–31.
7) Velozo CA, Wang Y, Lehman L, et al.: Utilizing Rasch measurement models to develop a computer adaptive self-report of walking, climbing, and running. Disabil Rehabil, 2008, 30: 458–467. [Medline] [CrossRef]
8) Choi B: Measurement precision for Oswestry Back Pain Disability Questionnaire versus a web-based computer adaptive testing for measuring back pain. J Back Musculoskeletal Rehabil, 2015, 28: 145–152. [Medline] [CrossRef]
9) Taherbhai HM, Young MJ: Pre-equating: a simulation study based on a large scale assessment model. J Appl Meas, 2004, 5: 301–318. [Medline]
10) Wright BD, Masters GN: Rating scale analysis. Chicago: MESA Press, 1982, pp 11–17.
11) Fisher AG: The assessment of IADL motor skills: an application of many-faceted Rasch analysis. Am J Occup Ther, 1993, 47: 319–329. [Medline] [CrossRef]