Self-Reported Mobility and Instrumental Activities of Daily Living: Test–Retest Reliability and Criterion Validity

Ching-Yi Wang, Ming-Hsia Hu, Hui-Ya Chen, and Ren-Hau Li

To determine the test–retest reliability and criterion validity of self-reported function in mobility and instrumental activities of daily living (IADL) in older adults, a convenience sample of 70 subjects (72.9 ± 6.6 yr, 34 male) was split into able and disabled groups based on baseline assessment and into consistently able, consistently disabled, and inconsistent based on repeat assessments over 2 weeks. The criterion validities of the self-reported measures of mobility domain and IADL-physical subdomain were assessed with concurrent baseline measures of 4 mobility performances, and that of the self-reported measure of IADL-cognitive subdomain, with the Mini-Mental State Examination. Test–retest reliability was moderate for the mobility, IADL-physical, and IADL-cognitive subdomains (κ = .51–.66). Those who reported being able at baseline also performed better on physical- and cognitive-performance tests. Those with variable performance between test occasions tended to report inconsistently on repeat measures in mobility and IADL-cognitive, suggesting fluctuations in physical and cognitive performance.

Keywords: elderly, community-dwelling, aging, physical activity

Maintaining the ability to live independently is one of the top priorities for older people, as well as for health care professionals who care for older adults. A reliable, valid, and easy-to-use tool is required to monitor changes in function and identify individuals at an early stage of physical- or cognitive-function decline to provide effective early intervention or prevention programs.

Physical disability in older adults is usually assessed by self-reported measures of function in the domains of mobility, instrumental activities of daily living (IADL), and basic activities of daily living. The development of physical disability is thought to follow a hierarchical pattern (Barberger-Gateau, Rainville, Letenneur, & Dartigues, 2000; Chen et al., 2010). Older adults disabled with respect to mobility or IADL functions are in danger of developing disability in basic activities of daily living (Guralnik, Ferrucci, Simonsick, Salive, & Wallace, 1995; Harris, Kovar, Suzman, Kleinman, & Feldman, 1989), which usually results in the need for long-term care. Hence, those at earlier stages of disability (such as mobility and
Self-Reported Measure of Function (IADL) might respond better to early intervention and, from a preventive-medicine point of view, are therefore important to be identified early.

Compared with performance tests, which are usually considered gold-standard tests for physical capability, the self-reported measure of function has the advantages of low cost, ease of administration, and time economy, in addition to requiring no special space or equipment. Past studies have supported the validity of using a self-reported measure of function at one assessment to identify older adults with poor physical performance and health status (Fried, Bandeen-Roche, Chaves, & Johnson, 2000; Fried, Young, Rubin, Bandeen-Roche, & WHAS II Collaborative Research Group, 2001; Wang, Olson, Gleeson, & Protas, 2005; Wang, Sheu, & Protas, 2007). However, use of self-reported measures on two occasions to follow changes in physical and cognitive functions in older adults has not been reported. In addition, past studies have identified both physical and cognitive subdomains in the IADL domain (Barberger-Gateau et al., 1992; Ng, Niti, Chiam, & Kua, 2006). Whether the self-reported measure of function in the IADL-physical and IADL-cognitive subdomains could reflect older adults’ performance on physical and cognitive function, respectively, has not been tested in community-dwelling older adults. Thus, the ability of a self-reported measure of mobility and IADL function to reflect performance of physical and cognitive functions needs to be ascertained before the self-reported measure can be confidently recommended for use in monitoring changes and identifying decline in physical and cognitive function.

The retest reliability of self-reported measures of mobility and IADL functions has been found to be fair to good in community-based older adults, specifically among Dutch men (Hoeymans, Wouters, van den Bos, & Kromhout, 1997) and in a cohort of African Americans (Andresen, Malmstrom, Miller, Miller, & Wolinsky, 2005; Miller, Andresen, Malmstrom, Miller, & Wolinsky, 2006). The test–retest reliability of self-reported measures of mobility and IADL functions has not been reported in an Asian population. Older adults living in cities with different layouts or those with different cultures and lifestyles may encounter different challenges in their living activities related to maintaining independence in their community. Thus, the test–retest reliability needs to be established in the population of interest.

This study aimed to examine the test–retest reliability and criterion validity of a self-reported measure of function in the mobility domain and IADL-physical and IADL-cognitive subdomains in a group of community-dwelling older Taiwanese adults. The three specific goals were to (a) determine the test–retest reliability of the self-reported mobility and IADL status, determined by two face-to-face interviews separated by 2 weeks; (b) test the criterion validity of self-reported function in the mobility domain and IADL-physical and -cognitive subdomains based on baseline assessment, to reflect physical and cognitive function, respectively; and (c) test the criterion validity of self-reported function in the mobility domain and IADL-physical and -cognitive subdomains based on repeat assessments, to reflect the stability of physical and cognitive function, respectively. Participants were grouped into able versus disabled groups based on the baseline assessment and as consistently able, consistently disabled, or inconsistent based on the repeat assessments. By comparing the group mean differences in physical and cognitive performance taken at baseline, it was possible to examine the criterion validity of self-reported status of the mobility domain and the IADL-physical and -cognitive subdomains at baseline and repeat assessments.
We hypothesized that (a) the self-reported measure of mobility and IADL function would show moderate test–retest reliability similar to those reported previously in Western countries; (b) the mobility domain and IADL-physical subdomain would reflect performance on physical function, whereas the IADL-cognitive subdomain would reflect performance on cognitive function; and (c) older adults with either better or worse function would tend to report consistently at the repeat assessments, whereas those with intermediate function might experience fluctuations in function over time and therefore would tend to report inconsistently.

Method

Participants

Older adults participating in activities at community centers were approached by the research physical therapist to determine their eligibility to participate in this study. The inclusion criteria were being 60 years of age or more, living independently in the community, and being able to follow instructions and complete the assessments in this study. The exclusion criteria were lesions of the central and peripheral nervous system (such as stroke, head injury, Guillain-Barré syndrome, Parkinson’s disease, or multiple sclerosis) or severe musculoskeletal disorders that would interfere with physical-performance tests and a score lower than the suggested cutoff score based on age and years of education on the Chinese version of the Mini-Mental State Examination (Chiu, Lee, Chung, & Kwong, 1994; Katzman et al., 1988).

A priori power analysis was conducted based on an effect size for grip strength developed from a previous study (Wang et al., 2007). The calculated effect size of .83 for grip strength was used to determine the sample size; a minimum of 26 participants each in the able and disabled groups was needed to achieve a power of .80 and Type I error of .05 for two-tailed t test. Seventy-three older adults who met the inclusion criteria were invited by the research physical therapist to participate in this study. Three older adults with stroke and severe diabetes (which would interfere with their physical-performance tests) were excluded from this study. All participants signed an informed-consent form, reviewed and approved by an institutional review board, before participation.

Procedure and Materials

All participants were assessed at baseline and 2 weeks later by a research physical therapist who was blinded to the purposes of this study. At baseline assessment, participants underwent a face-to-face interview to collect their demographic (age, gender, years of education) and general health information, which consisted of self-perceived health status (healthier, same, or less healthy than other people their age) and chronic diseases diagnosed by a physician (hypertension, heart disease, diabetes, arthritis, lung disease, vision problems, hearing problems, cancer). They also received the following assessments at baseline: self-reported function in the mobility and IADL domains and physical- and cognitive-performance tests.

In the retest, only those who reported no adverse events during the past 2 weeks that would jeopardize their physical and cognitive function, such as serious injuries
or sickness, loss of loved ones, or financial crisis, were assessed for self-reported function in the mobility and IADL domains. At the retest, the rater had no access to the baseline assessment results.

**Self-Reported Measure of Function.** The self-reported measure of function consisted of the mobility domain and IADL-physical and -cognitive subdomains adopted from a previous study (Barberger-Gateau et al., 2000; Ng et al., 2006). There were three items in the mobility domain (Barberger-Gateau et al., 2000). In the IADL domain, there were three items in the IADL-cognitive subdomain, and another five items composed the IADL-physical subdomain (Ng et al., 2006). Items in each domain are listed in Table 1.

Self-reported function was assessed with standardized questions: “Without help from another person or special equipment, by yourself, are you able to perform the following tasks?” The three possible responses were “able,” “need help,” or “unable.” For any item they reported they had no need to do, we followed up by asking, “When you need to, can you do it independently, with help, or are you unable to?”

An item was rated independent if the participants reported that they were able to perform it independently; otherwise, it was rated dependent for those who reported “need help” or “unable” to perform. A domain was rated able if all items in that

### Table 1  Percentage of Disability, Indicated by Self-Reported Dependence, at Baseline and Retest Assessments and the Test–Retest Reliability of Mobility and Instrumental Activities of Daily Living (IADL) Functions (N = 70)

<table>
<thead>
<tr>
<th></th>
<th>Disability, %</th>
<th>Test–Retest Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Retest</td>
</tr>
<tr>
<td><strong>Mobility domain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>walking for 400 m</td>
<td>10.0</td>
<td>5.7</td>
</tr>
<tr>
<td>walking up one flight of stairs</td>
<td>14.3</td>
<td>7.1</td>
</tr>
<tr>
<td>lifting and carrying heavy object</td>
<td>20.0</td>
<td>17.1</td>
</tr>
<tr>
<td><strong>IADL-physical subdomain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shopping</td>
<td>2.9</td>
<td>2.9</td>
</tr>
<tr>
<td>mode of transportation</td>
<td>15.7</td>
<td>12.9</td>
</tr>
<tr>
<td>food preparation</td>
<td>15.7</td>
<td>20.0</td>
</tr>
<tr>
<td>housekeeping</td>
<td>1.4</td>
<td>1.4</td>
</tr>
<tr>
<td>laundry</td>
<td>32.9</td>
<td>22.9</td>
</tr>
<tr>
<td><strong>IADL-cognitive subdomain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ability to use telephone</td>
<td>5.7</td>
<td>1.4</td>
</tr>
<tr>
<td>responsibility for own medications</td>
<td>4.3</td>
<td>7.1</td>
</tr>
<tr>
<td>ability to handle finances</td>
<td>20.0</td>
<td>24.3</td>
</tr>
<tr>
<td><strong>IADL-total domain</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>55.7</td>
<td>52.9</td>
</tr>
</tbody>
</table>
domain were rated independent; otherwise, it was rated disabled. Older adults rated able at both the baseline and retest assessments were grouped as consistently able, those rated disabled at both contacts were grouped as consistently disabled, and those rated differently at the two contacts were grouped as inconsistent.

**Physical and Cognitive Function.** Physical function was measured by assessing the following performance tasks: grip strength, timed chair stands (TCS), timed up-and-go (TUG), and fastest gait speed (FGS). During the performance assessments, sufficient rest was provided when required. Grip strength was measured for the dominant hand using a Jamar hand dynamometer (Sammons Preston Rolyan, Bolingbrook, IL) set at the second handle position and following the standard testing procedure reported by the American Society of Hand Therapists (Mathiowetz, Weber, Volland, & Kashman, 1984). Participants were required to sit with the arm positioned by the side of the upper trunk and the elbow flexed to 90°. They were directed to exert maximal grip within 5 s. The best value of two trials was used. The TCS test measures the time (in seconds) required to stand up from a chair (46 cm high) five times as quickly and safely as possible (Guralnik et al., 1994). The test was performed if the participant could stand up once successfully from the chair with the arms crossed against the chest. The timing began with the participant sitting on the chair with the arms crossed against the chest and the verbal signal “Ready and go.” The timing was stopped when the participant stood upright for the fifth time. The TUG test records the time (in seconds) required for a participant to stand up from a chair (46 cm high), walk 3 m forward as quickly and safely as possible, turn around, return to the chair, turn around, and sit down (Podsiadlo & Richardson, 1991). Timing of the test began after the word *go* and stopped when the participant sat back down on the chair. FGS was assessed by recording the time (in seconds) required to walk a 15.24-m distance at the fastest speed, with participants starting from a standing posture. The participants walked twice consecutively, and the best performance was used for calculation of gait speed and for data analyses. The reliability of the TCS, TUG, and FGS has been found to be excellent (intraclass correlation coefficient \[2, 1\] = .89–.95; Wang, Sheu, & Protas, 2009), and that for grip strength, also excellent (ICC ≥ .92; Wang & Chen, 2010), in groups of community-dwelling older Asian adults.

Cognitive function was assessed by a Chinese version of the Mini-Mental State Examination (Chiu et al., 1994), which consists of 11 items with 30 responses and has a score range of 0–30.

**Statistical Analysis**

All data analysis was performed using SPSS 17.0. Descriptive statistics of demographic information and general health information were reported. The percentage of dependence in each item and domain (mobility, IADL-physical, IADL-cognitive) at baseline and retest was also reported.

The retest reliability of each item and domain (mobility, IADL-physical, IADL-cognitive) was assessed using kappa statistics (κ). Kappa values .00–.20 are considered slight, .21–.40 fair, .41–.60 moderate, .61–.80 substantial, and .81–1.0 almost perfect (Smith et al., 1990). Because a low percentage of physical disability can cause imbalanced marginal totals and can affect kappa values even with a relatively high value of observed agreement (Hoeymans et al., 1997), the
observed agreement was also reported (Cicchetti & Feinstein, 1990; Feinstein & Cicchetti, 1990; Hoeymans et al., 1997).

The age and sex distributions between the able and disabled groups based on self-reported mobility domain and IADL-physical and -cognitive subdomains were examined by \( t \) test and chi-square test, respectively, and no significant differences were found. The criterion validities of the self-reported function in the mobility domain and IADL-physical and -cognitive subdomains at baseline assessment, to reflect physical and cognitive performances, were examined by independent \( t \) tests with a significance level of .01. The effect-size index (\( d \)—for comparison of two groups) was also determined (Portney & Watkins, 2009) and reported.

The differences in age, \( F(2, 67) = 0.39, p = .68 \), and sex distribution, \( \chi^2(2) = 4.24, p = .12 \), among the consistently able, inconsistent, and consistently disabled groups, based on repeat self-reported mobility function, were not significant. The criterion validities of the self-reported mobility function on repeat measures to reflect physical performance were examined by one-way ANOVA with a significance level of .01. The effect-size index (\( f \)—for comparison of three groups) was also determined (Portney & Watkins, 2009) and reported.

Among the three groups classified by repeat self-reported IADL-physical function, sex distribution was found to be similar, \( \chi^2(2) = .51, p = .78 \), but a significant group difference in age was found, \( F(2, 67) = 5.60, p = .006 \). Thus, the criterion validities of the self-reported IADL-physical function on repeat measures to reflect physical performance were examined by separate analysis of covariance (ANCOVA), adjusted for age, with a significance level of \( p < .01 \).

Furthermore, among the three groups classified by self-reported IADL-cognitive function, the age, \( F(2, 67) = 2.46, p = .09 \), was found not to be significant, but sex distribution, \( \chi^2(2) = 8.62, p = .013 \), was significant. The criterion validity of the self-reported function in the IADL-cognitive subdomain on repeat measures to reflect cognitive performance was examined by an ANCOVA adjusted for sex with a significance level of \( p < .01 \). When the global ANOVA or ANCOVA showed a significant difference, a post hoc analysis was performed.

**Results**

A total of 70 community-dwelling older adults (72.9 ± 6.6 years, range 60–90; 34 men, 36 women) who met the inclusion and exclusion criteria were enrolled in this study. Few (11.4%) perceived themselves as less healthy than their peers. On average, they reported 3.2 (± 1.6) comorbidities. The mean interval before the second test was 14.6 (± 2.6) days. Most of our participants had an elementary school education or were literate (60.0%), 11.4% were illiterate, 22.9% had a junior to senior high school education, and 5.7% had a college or higher level of education.

The percentage of dependence in each item and domain is reported in Table 1. In the mobility domain, lifting or carrying a heavy object was the task that showed the highest percentage of dependence, followed by walking up stairs and walking 400 m. The proportion of agreement in repeated self-report was high in the mobility domain (88.6%), IADL-physical subdomain (75.7%), and IADL-cognitive subdomain (80.0%). The kappa values in the mobility domain (\( \kappa = 0.66 \)) and IADL-physical (\( \kappa = .51 \)) and IADL-cognitive (\( \kappa = .52 \)) subdomains were moderate. The kappa values for the items in the mobility domain (\( \kappa = .48–.71 \)) and IADL-physical (\( \kappa = .47–1.00 \)), and IADL-cognitive (\( \kappa = .39–.63 \)) subdomains were fair to moderate.
There was a significant group difference in performance on TCS, $t(19.4) = -2.66, p = .015, d = .84$; TUG, $t(19.1) = -3.92, p = .001, d = 1.24$; and FGS, $t(68) = 3.81, p < .001, d = 1.04$, between the able and disabled groups based on baseline assessment in the mobility domain (Table 2).

Among the consistently able, consistently disabled, and inconsistent groups, based on repeat measurement of the mobility domain, the group mean difference was found to be significant on TCS, $F(2, 67) = 8.18, p = .001, f = .49$; TUG, $F(2, 67) = -17.29, p < .001, f = .72$; and FGS, $F(2, 67) = 11.84, p < .001, f = .59$ (Table 3). The results of post hoc analysis indicated that the consistently able group was significantly different from the consistently disabled group ($p = .002, d = 1.26$), from the consistently disabled group ($p < .001, d = 1.86$) and the inconsistent group ($p = .033, d = 1.02$) on TUG, and from the consistently able group ($p < .001, d = 1.62$) and inconsistent group ($p = .019, d = 1.35$) on FGS. Participants in the disabled and consistently disabled groups in the mobility domain demonstrated significantly worse physical performance on the TCS, TUG, and FGS than the other groups (Table 2 and 3). A similar trend was found in the IADL-physical subdomain based on baseline (Table 2) and repeat assessment (Table 3); however, this trend did not reach statistical significance.

Older adults in the able group, $t(68) = 3.11, p = .003, d = 0.81$, and the consistently able group, $F(2, 66) = 8.08, p = .001, f = .49$, based on IADL-cognitive subdomain at baseline and repeat assessment, respectively, showed significantly better cognition than the inconsistent and consistently disabled groups (Table 2 and 4). The mean scores for physical and cognitive performance of the inconsistent group were variable between those of the consistently able and consistently disabled groups, regardless of whether the mobility domain or IADL subdomains were used.

**Discussion**

This study aimed to examine the test–retest reliability and criterion validity of self-reported measures of function in the mobility domain and IADL-physical and -cognitive subdomains in a group of community-dwelling older Asian adults. There were three important findings from the results. First, the test–retest reliability was found to be moderate for the mobility domain ($\kappa = .66$) and the IADL-physical ($\kappa = .51$) and -cognitive subdomains ($\kappa = .52$), whereas it was fair to moderate on individual items in the three domains but with high total agreement (>81.4%) over a 2-week interval. A few items with kappa statistics lower than .5 because of the low percentage of dependence in these tasks (such as walking up one flight of stairs, shopping, food preparation, ability to use the telephone, and responsibility for own medication) caused imbalanced marginal totals and affected kappa values but not the observed agreement (% of total agreement).

The percentage of the sample dependent at retest seemed to be less than that at baseline in the mobility domain but not in the IADL-physical subdomain. In the mobility domain, the percentage agreement for the “able” response was consistently higher than the percentage agreement for the “disabled” response across all items. In other words, participants who gave negative responses (able) at the first test were more likely to give the same response at the retest, whereas those gave positive responses (disabled) at the first test were more likely to change their responses at
### Table 2  Criterion Validity of Self-Reported Measures of Function in the Mobility Domain and IADL-Physical and -Cognitive Subdomains in the Baseline Able and Disabled Groups, \( M (SD) \)

<table>
<thead>
<tr>
<th>Mobility Domain</th>
<th>IADL-Physical Subdomain</th>
<th>IADL-Cognitive Subdomain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline measure</td>
<td>Able ( n = 52 )</td>
<td>Disabled ( n = 18 )</td>
</tr>
<tr>
<td>Grip strength (kg)</td>
<td>26.3 (7.5)</td>
<td>23.3 (9.0)</td>
</tr>
<tr>
<td>Timed chair stand (s)(^a)</td>
<td>8.7 (2.6)</td>
<td>12.4 (5.8)</td>
</tr>
<tr>
<td>Timed up-and-go (s)(^a)</td>
<td>8.6 (2.1)</td>
<td>13.4 (5.0)</td>
</tr>
<tr>
<td>Fastest gait speed (m/s)(^a)</td>
<td>1.5 (0.3)</td>
<td>1.1 (0.4)</td>
</tr>
<tr>
<td>C-MMSE(^b)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Note. IADL = instrumental activities of daily living; C-MMSE: Chinese version of the Mini-Mental State Examination.

\(^a\)Significant group mean difference in the mobility domain \((p < .01)\). \(^b\)Significant group mean difference in the IADL-cognitive subdomain \((p < .01)\).

### Table 3  Criterion Validity of Self-Reported Measures of Function in the Mobility Domain and IADL-Physical Subdomain to Reflect Physical Performance Among the Consistently Able, Inconsistent, and Consistently Disabled Groups Based on Repeat Assessments

<table>
<thead>
<tr>
<th>Mobility Domain</th>
<th>IADL-Physical Subdomain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline measure</td>
<td>Consistently able ( n = 51 )</td>
</tr>
<tr>
<td>Grip strength (kg)</td>
<td>26.4 (7.5)</td>
</tr>
<tr>
<td>Timed chair stand (s)(^a)</td>
<td>8.6 (2.6)(^b)</td>
</tr>
<tr>
<td>Timed up-and-go (s)(^a)</td>
<td>8.6 (2.1)(^c)</td>
</tr>
<tr>
<td>Fastest gait speed (m/s)(^a)</td>
<td>1.5 (0.3)</td>
</tr>
</tbody>
</table>

Note. IADL = instrumental activities of daily living.

\(^a\)Significant group mean difference \((p < .01)\). \(^b\)Significantly different from the consistently disabled group. \(^c\)Significantly different from the other two groups. \(^d\)Significantly different from the other two groups.
This phenomenon was not observed in the IADL-physical domain. Across the items in the IADL-physical domain, similar numbers of participants changed their responses between able and disabled at the retest. Thus, the percentages of dependence at both baseline and retest assessments were close. It is possible that factors other than physical performance affect older adults’ perception of their ability to perform the activities in the IADL-physical subdomain.

Our results of test–retest reliability in an Asian population are similar to previous findings from Western countries with a retest interval of 2 weeks (Hoeymans et al., 1997) but less favorable than those obtained with a retest interval shorter than 2 weeks (Andresen et al., 2005; Fried et al., 1996; Tager, Swanson, & Satariano, 1998). Different issues need to be considered as to the test interval of repeat measurements for assessment of test–retest reliability and for responsiveness to changes. In this study, the test interval chosen was 2 weeks because it was not long enough for real change to occur and not short enough for participants to recall previous answers. To avoid including individuals who encountered significant adverse events that could cause a real change in their physical and cognitive function during the 2-week interval, we asked participants if they had suffered from serious injuries or sickness, the loss of loved ones, or financial crises during the 2 weeks before the retest. No participants reported such adverse events, so none were excluded for this reason. In addition, we used only one rater, who used the standardized questionnaire and followed standardized protocol to administer all the self-reported measures and performance tests. Despite the efforts to maximize the test–retest reliability, our results indicated that those whose physical and cognitive functions were in the intermediate level tended to give inconsistent responses when prompted with a self-reported questionnaire. This inconsistency may be the result of fluctuations in function, which might explain the moderate test–retest reliability we found.

Second, consistent with previous studies, our results supported the criterion validity of self-reported measure of function in the mobility domain based on baseline assessment to identify older adults with worse physical performance (Fried et al., 2001; Wang et al., 2005; Wang et al., 2007). Individuals reporting disability in the mobility domain demonstrated significantly poorer physical performance on the TCS, TUG, and FGS than the able group and than the reference values reported in previous studies (Bischoff et al., 2003; Bohannon, 2006a,b; Wang, Olson, Yeh,

### Table 4

#### Criterion Validity of Self-Reported Measure in the IADL-Cognitive Subdomain to Reflect Cognitive Performance Among the Consistently Able, Inconsistent, and Consistently Disabled Groups During Repeat Assessments

<table>
<thead>
<tr>
<th>IADL-Cognitive Subdomain</th>
<th>Consistently able $(n = 42)$</th>
<th>Inconsistent $(n = 14)$</th>
<th>Consistently disabled $(n = 14)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-MMSE</td>
<td>26.8 (3.0)$^a$</td>
<td>23.1 (4.2)</td>
<td>22.9 (3.8)</td>
</tr>
</tbody>
</table>

*Note.* IADL = instrumental activities of daily living; C-MMSE = Chinese version of Mini-Mental State Examination. Significant group mean difference ($p < .01$).

$^a$Significantly different from the other two groups.
The TCS, TUG, and gait-speed performance in older adults have clinically important meanings—they are significant indicators of dependency (Bischoff et al., 2003), mobility disability (Wong, Wong, Pang, Azizah, & Dass, 2003), and future disablement (Wang, Yeh, & Hu, 2011). Our results also support the criterion validity of the IADL-cognitive subdomain on baseline assessment to identify community-dwelling older adults with poor cognitive function (Barberger-Gateau et al., 1992; Ng et al., 2006).

Finally, an original and unique finding in this study is the demonstration of the criterion validity of repeat self-reported measures of function. Our results showed that participants with physical and cognitive functions at the better and worse ends of the scale tended to report consistently, whereas participants who had an intermediate level of function tended to report inconsistently between test and retest, regardless of whether the mobility domain or IADL-cognitive subdomain was used, suggesting that this group of individuals experienced fluctuations in these areas.

The criterion validity of the IADL-physical subdomain to reflect physical performance was not supported using either baseline or repeat assessments. In post hoc analysis, we calculated the effect size and found it to be small across all four performance tests (eta-squared = .006−.030), thus indicating that the IADL-physical subdomain was not as sensitive as the mobility domain in reflecting older adults’ physical performance. The generalization of the results of our study is limited by the use of a convenience sample. Future studies to ascertain the test–retest reliability with a shorter interval and the validity of self-reported measures of function in mobility and IADL subdomains separately for men and women using a representative sample are warranted. Nevertheless, to the best of our knowledge, this is the first study to report the test–retest reliability of a self-reported measure of function in a group of older Asian adults. This is also the first study to explore the criterion validity of using a repeat self-reported measure of function to identify those at an intermediate level of function (those reported inconsistently) in physical and cognitive performance.

Conclusions

The results of the current study support the test–retest reliability (moderate) and criterion validity of self-reported measure of function in mobility and IADL subdomains based on baseline assessment and repeat assessments. The results of baseline assessment can be used to identify those with worse physical and cognitive performance, whereas results from repeat assessments can be used to identify individuals who may be in an unstable status of physical and cognitive function and may experience functional decline soon in the future (the inconsistent group). For identifying those with a decline in physical function, the mobility domain is useful. For identifying individuals with decline in cognitive function, the IADL-cognitive subdomain is helpful.

Acknowledgments

This study was supported by a research grant from the National Science Council of Taiwan (NSC97-2314-B-040-005). The authors report no conflicts of interest.
References


