Relationship between self-reported function and disability and balance performance measures in the elderly

Itshak Melzer, PhD, PT;* Ilan Kurz, MSc, PT; Orly Sarid, PhD; Alan M. Jette, PhD
Departments of 1Physical Therapy, Faculty of Health Sciences, and 2Social Work, Faculty of Humanities and Social Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel; 3Health and Disability Research Institute, Boston University, Boston, MA

Abstract—This study evaluated the reliability and validity of the Hebrew version of the Late-Life Function and Disability Instrument (LLFDI). Fifty-five older adults (mean age 79.7 +/- 5.2) participated. We calculated test-retest reliability with intraclass correlation coefficients (ICCs). Partial correlations determined the construct validity with a balance measure (Berg Balance Scale [BBS]) and a mobility measure (Timed Up and Go [TUG] test). We examined known-group validity by comparing the scores of cane and noncane users. Test-retest ICCs ranged from good to excellent (0.77–0.90) for the function component and fair to good for the disability component (0.63–0.83), except for the disability management role subscale (0.46). BBS and TUG were associated with LLFDI overall function (r = 0.48, p < 0.001 and r = –0.52, p < 0.001, respectively). TUG and BBS were weakly associated with disability limitations (r = –0.26 and 0.32, respectively) and disability frequency (r = –0.16 and 0.24, respectively). Cane users showed significantly lower function scores than noncane users. We demonstrated that the Hebrew version of the LLFDI reliably and validly assesses older adults’ function and disability. The LLFDI is recommended as an outcome instrument in studies in which older adults’ function and disability are outcomes of interest.

Key words: aging, balance, disability instrument, gait, Hebrew, late-life function, mobility, physical disability, physical functioning, rehabilitation.

INTRODUCTION

Self-reported capability in physical functioning and disability has long been considered an important focus of research on older persons. Current measures have been criticized for conceptual confusion, lack of sensitivity to change, poor reproducibility, poor validity, and inability to capture a wide range of upper- and lower-limb functioning [1–2]. Clinical and self-reported measures of function should demonstrate sensitivity to relevant change and remain invariable when function does not change. Also, efforts to evaluate the effectiveness of clinical trials and community-based interventions designed to affect late-life function and disability have been hindered significantly by current instrumentation limitations. Several studies have shown poor to moderate associations between performance-based measures and self-reported functional status in older subjects [3–8]. Furthermore, Daltroy et al. found that people may recalibrate their self-assessments based on recent health problems [9], which can raise

Abbreviations: BBS = Berg Balance Scale, BI = Barthel Index, EAMQ = Environmental Analysis of Mobility Questionnaire, ICC = intraclass correlation coefficient, LLFDI = Late-Life Function and Disability Instrument, PF-10 = 10-item Physical Functioning (subscale), TUG = Timed Up and Go (test).
*Address all correspondence to Itshak Melzer, PhD, PT; Schwartz Movement Analysis and Rehabilitation Laboratory, Department of Physical Therapy, Recanati School for Community Health Professions, Faculty of Health Sciences, Ben-Gurion University of the Negev, PO Box 653, Beer-Sheva 84105, Israel; +972-8-647-7727; fax: +972-8-647-7683. Email: itzikm@bgu.ac.il
DOI: 10.1682/JRRD.2006.10.0133
questions about the validity of self-reported measures of physical function and disability in older persons. Nevertheless, self-reported measures remain the most commonly used instruments in studies involving older adults because of their low cost and practicality.

A need exists for investigating whether balance and gait function in older persons is associated with self-reported physical functioning and disability since balance control plays a fundamental role in various activities of daily living, especially those that include elements of independent standing and gait. Age-related deterioration of the balance-control system can cause balance impairment and limited mobility and thereby cause disability that may contribute to falls. Deteriorating balance function in older people has been correlated with an increased risk for falls [10], decreased quality of life [11], and functional decline in physical and social activities [12].

The present study investigated (1) the test-retest reliability of the Hebrew version of the Late-Life Function and Disability Instrument (LLFDI) [13–14], (2) the construct validity of the function and disability components of the LLFDI through correlation of scales with performance-based clinical measures of balance and gait, and (3) known-group validity to assess whether LLFDI scores could distinguish between function and disability levels in the physical functioning of older adults who used a cane compared with older adults who walked independently.

METHODS

Study Population

Fifty-five nondisabled older adults (69–91 years old) were recruited from community dwelling centers in the Beer-Sheva, Israel, region through flyers and advertisements. All subjects provided informed consent in accordance with procedures approved by the Soroka University Medical Center Institutional Review Board (Helsinki Ethical Committee Approval 3794). The inclusion criteria included ability to ambulate, Mini-Mental State Examination score >24, no major neurological or orthopedic disorders, Berg Balance Scale (BBS) score >45, and no serious visual impairment or blindness.

Study Design

A content expert experienced in translation and cultural adaptation translated the LLFDI from English to Hebrew. To ensure that all items were culturally appropriate, rehabilitation experts reviewed and revised drafts and provided feedback on the appropriateness of the individual items. The Hebrew version was then translated back to English by an experienced translator and then revised and approved by Dr. Jette, the developer of the LLFDI.

Late-Life Function and Disability Instrument

The LLFDI was developed as a comprehensive assessment of function and disability in community-dwelling older adults [13–14]. Consistent with Naki’s disablement framework, the LLFDI contains items that represent functional limitations (inability to perform discrete physical tasks encountered in daily routines) and disability (inability to participate in major life tasks and social roles) [13–14]. It was designed to respond to meaningful changes in function and disability. The function component evaluates self-reported difficulty in the performance of 32 physical activities in three dimensions: upper limb, basic lower limb, and advanced lower limb [14]. Subscales are each scored on a 0–100 scale, with higher scores indicating higher levels of functioning. The disability component evaluates self-reported limitations in life activities and frequency of participation in 16 major life tasks [13]. The limitation (instrumental and management roles) and frequency (social and personal roles) dimensions of the disability component are each scored on a 0–100 scale, with higher scores indicating higher levels of functioning.

The function component of the LLFDI has shown extremely high test-retest reliability (intraclass correlation coefficient [ICC] = 0.91 to 0.98) and the disability component has shown moderate to high test-retest reliability (ICC = 0.68 to 0.82) [13–14]. The disability component has shown concurrent validity by correlating with the 10-item Physical Functioning (PF-10) subscale of the Medical Outcomes Study 36-item Short-Form Health Survey [13] and the London Handicap Scale [15].

Balance Function and Mobility Performance Tests

Balance and gait were quantified with the BBS [16] and the Timed Up and Go (TUG) test [17]. In the BBS, we scored the participants on 14 tasks graded on a 0–4 scale (maximum 56) to evaluate balance under different conditions; higher scores indicate higher levels of balance function [16]. In the TUG, participants sat in a chair placed 3 m from a wall. They were instructed to rise from the chair, walk at their normal pace to the wall, turn
around, return to the chair, and sit down. The researcher uses a stopwatch to time this task, and lower scores indicate higher levels of functioning. All subjects performed the examinations without any assistive device.

According to a logistic regression model, a BBS score between 45 and 53 corresponds to individuals with a 20 to 75 percent predicted probability of falling, with high interrater reliability (ICC = 0.98) [18]. High correlations also exist between the BBS and the Balance Subscale developed by Tinetti (r = 0.91), the Barthel Index (BI) mobility subscale (r = 0.67), and the TUG (r = −0.76). The fall-prevention guidelines of the American Geriatrics Society, British Geriatrics Society, and American Academy of Orthopaedic Surgeons recommend that all older persons reporting a single fall be examined with the TUG [19]. The TUG was found to correlate with other measurements such as gait speed (r = −0.61) and the BI (r = −0.78).

To assess test-retest reliability, two experienced investigators administered the LLFDI to participants in two data-collection sessions separated by 10–14 days. The examiners had no knowledge of each other’s results. After completing these questionnaires at the first data-collection session, participants completed performance-based examinations with the BBS [16] and TUG [17].

We assessed validity of the LLFDI by correlating each LLFDI subscale with the BBS and TUG performance tests. We hypothesized that the function component of the LLFDI would demonstrate convergent validity through strong positive associations with both performance tests and that the strongest positive association would be with the lower-limb function domains. In addition, we hypothesized that the balance and gait test scores would demonstrate discriminant validity by demonstrating weaker associations with the disability component of the LLFDI.

We assessed known-group validity by evaluating whether LLFDI scale scores distinguished between function and disability levels in the physical functioning levels of older adults who used a cane versus walked independently. We hypothesized that the LLFDI function component would discriminate better between cane and noncane users than the LLFDI disability component.

**Statistical Analysis**

We assessed intertester reliability using a two-way random model for ICC single measure. Test-retest reliability compared the results of the first and second LLFDI administrations. ICCs were determined for both the function and disability components of the two test sessions. The following guidelines were used for interpretation of ICCs: <0.40 represents poor reliability, 0.40 to 0.75 represents fair-to-good reliability, and >0.75 represents excellent reliability [20].

Associations among all variables (BBS, TUG, and LLFDI function and disability components) were assessed by Pearson correlation. Because TUG and BBS summary scores were significantly correlated with age (r = −0.34, p < 0.001 and r = 0.32, p = 0.001, respectively), we ran partial correlations on all variables while controlling for age. Statistical significance was accepted at p < 0.004 (0.05/12) with a testwise correction.

An independent t-test was performed that compared the LLFDI subscale scores of cane and noncane users. Statistical significance was accepted at p < 0.05. All data were analyzed with SPSS software (SPSS Inc, Chicago, Illinois).

**RESULTS**

The participants’ demographic background characteristics are shown in Table 1, along with a summary of their balance and gait performance test results.

**Test-Retest Reliability**

Good to excellent test-retest reliability was found for the LLFDI function component (ICC = 0.77 to 0.90) (Table 2). Overall and advanced lower-limb function were the most reliable (ICC = 0.90 and 0.86, respectively). Test-retest reliability of the LLFDI disability component was fair to good (ICC = 0.63 to 0.83), except for the management role, which showed fair reliability but was statistically significant (ICC = 0.46, p < 0.001).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>79.7 ± 5.2</td>
</tr>
<tr>
<td>Mini-Mental State Examination</td>
<td>29.0 ± 0.9</td>
</tr>
<tr>
<td>Medications (No.)</td>
<td>4.1 ± 2.5</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>66.7 ± 11.7</td>
</tr>
<tr>
<td>Male/Female (No.)</td>
<td>13/42</td>
</tr>
<tr>
<td>Berg Balance Scale</td>
<td>52.8 ± 2.9</td>
</tr>
<tr>
<td>Timed Up and Go Test (s)</td>
<td>7.5 ± 2.4</td>
</tr>
</tbody>
</table>

Table 1. Summary of characteristics of study participants (n = 55). Data shown as mean ± standard deviation unless otherwise indicated.
Construct Validity

The moderate and significant correlations between the TUG and LLFDI overall function ($r = -0.52$, $p < 0.001$), basic lower-limb function ($r = -0.49$, $p < 0.001$), and advanced lower-limb function ($r = -0.49$, $p < 0.001$) are shown in Table 3. Moderately strong correlations were found between the BBS and LLFDI overall function ($r = 0.48$, $p < 0.001$), basic lower-limb function ($r = 0.51$, $p < 0.001$), and advanced lower-limb function ($r = 0.46$, $p < 0.001$). As hypothesized, the TUG and BBS were less strongly associated with upper-limb function ($r = -0.34$, $p < 0.01$, and $r = 0.30$, $p < 0.02$, respectively).

As hypothesized, weak and not statistically significant correlations were found between the TUG, BBS, and limitation subscores of the LLFDI disability component ($r = -0.26$ and 0.32, respectively) and between the TUG, BBS, and frequency subscores of the LLFDI disability component ($r = -0.16$ and 0.24, respectively), indicating the discriminant validity of LLFDI (Table 3).

The overall function component of the LLFDI was more strongly correlated with the limitation subscale than with the frequency subscale of the disability component. For example, correlation between the overall function component and the disability limitation score was fair ($r = 0.43$, $p = 0.001$), slightly lower for the instrumental subscore ($r = 0.39$, $p = 0.004$), and weaker and not statistically significant for the management subscore ($r = 0.24$, $p = 0.08$). The overall function component was weakly and not significantly correlated to the frequency score ($r = 0.26$, $p = 0.057$) and personal subscore ($r = 0.13$, $p = 0.34$); however, it was significantly correlated to the social subscore ($r = 0.39$, $p = 0.004$).

Known-Group Validity

As hypothesized, cane users had significantly lower scores in all LLFDI function components ($p < 0.001$), and the limitation subscore of the LLFDI disability component was significantly lower in the cane compared with noncane users ($p < 0.001$). However, as expected, the frequency subscale of the LLFDI disability component was only marginally significant (Table 4).

DISCUSSION

The present study showed high test-retest reliability of the function component of the Hebrew version of the LLFDI within three dimensions (ICC = 0.77 to 0.90) and good test-retest reliability for the disability component (ICC = 0.63 to 0.83), apart from the management role (ICC = 0.46). These findings are very similar to Haley et al. [13] and Jette et al. [14], suggesting that the reliability of the Hebrew version is similar to the English version. As with the English version, the management role subscore in the Hebrew version was low primarily because of the small number of scale items ($n = 4$).
These reliability findings are consistent with the test-retest correlations of function and disability measures of other self-reported and performance-based measures when applied in nondisabled older women [21–22] and frail older people [23]. Wolinsky et al. found that the magnitudes of reliability of physical performance and physiological assessments were equal to or less than the ICCs found in the present study [24]. For example, the ICCs in Wolinsky et al. were 0.81 for grip strength, 0.72 for chair stands, 0.56 for gait speed, 0.60 for one-leg stand, 0.52 for semitandem stand, 0.58 for tandem stand with eyes closed, and 0.27 for tandem stand with eyes open. Physiological assessments in their study included systolic and diastolic blood pressure, height, weight, body fat, and peak expiratory flow. Except for blood pressure (ICC = 0.51 and 0.55 for systolic and diastolic, respectively), the physiological assessments had ICCs > 0.89 [24].

The LLFDI demonstrated substantial construct validity with functional performance. The function component of the LLFDI demonstrated substantial associations with the TUG, and these correlations were strongest for lower-limb function ($r = -0.49$ to $-0.52$, $p < 0.001$). Hoeymans et al. showed lower associations between gait speed (8 ft walk) and self-reported mobility ($r = 0.17$ to 0.32) and weaker associations between the speed of a 400 m walk and self-reported difficulty in walking and lifting tasks ($r = -0.27$ to $-0.37$) [8]. Suzuki et al. did not find associations between maximum gait speed and PF-10 score ($r = 0.28$) [5].

Balance performance (as measured by the BBS) in the present study was significantly associated with the lower-limb function components of the LLFDI ($r = 0.46$ to 0.51, $p < 0.001$). Suzuki et al. found similar associations between stair-climbing performance and PF-10 score ($r = 0.48$) [5]. Shumway-Cook et al. found that the Environmental Analysis of Mobility Questionnaire (EAMQ) summary encounter score was positively correlated with the Short Physical Performance Battery ($r = 0.71$, $p < 0.001$) and inversely related to avoidance ($r = 0.77$, $p < 0.001$) [25]. Similarly, the BBS and the EAMQ summary encounter scores were strongly positively correlated ($r = 0.73$, $p < 0.001$) and the BBS and the EAMQ avoidance scores were inversely related ($r = 0.72$, $p < 0.001$). In contrast, Simonsick et al. found that the Short Physical Performance Battery score was weakly associated with self-reported mobility tasks in nondisabled older persons ($r = 0.21$ to 0.29) [4]. In the same study, repeated chair-rise performance was significantly associated with PF-10 score ($r = 0.58$). These findings suggest that the LLFDI better assesses daily physical tasks that require balance control (defined as a task-specific multijoint skill that includes the interaction of several physiological systems), while the PF-10 better assesses repeated chair-rise performance, which is related to quadriceps muscle strength (defined as the capacity to produce torque around a certain joint).

As expected, the associations between the LLFDI disability component and performance-based measures of balance and gait were weaker than those for the LLFDI

### Table 4.

Known-group validity of Hebrew version of Late-Life Function and Disability Instrument (LLFDI) ($n = 55$). Data shown as mean ± standard deviation.

<table>
<thead>
<tr>
<th>LLFDI</th>
<th>Total Sample</th>
<th>Cane Users</th>
<th>Noncane Users</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Function Component</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Function</td>
<td>64.5 ± 10.4</td>
<td>54.2 ± 6.8</td>
<td>66.4 ± 9.8</td>
<td>0.001</td>
</tr>
<tr>
<td>Upper Limb</td>
<td>79.0 ± 11.5</td>
<td>69.2 ± 9.0</td>
<td>80.9 ± 11.0</td>
<td>0.004</td>
</tr>
<tr>
<td>Basic Lower Limb</td>
<td>74.8 ± 13.7</td>
<td>61.6 ± 10.3</td>
<td>77.3 ± 12.9</td>
<td>0.001</td>
</tr>
<tr>
<td>Advanced Lower Limb</td>
<td>59.6 ± 16.3</td>
<td>43.9 ± 12.0</td>
<td>61.3 ± 15.9</td>
<td>0.003</td>
</tr>
<tr>
<td><strong>Disability Component</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Limitation</td>
<td>74.8 ± 13.4</td>
<td>58.6 ± 8.9</td>
<td>76.9 ± 12.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Instrumental Role</td>
<td>76.3 ± 14.4</td>
<td>56.9 ± 9.0</td>
<td>78.2 ± 13.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Management Role</td>
<td>86.7 ± 10.9</td>
<td>72.8 ± 18.2</td>
<td>88.2 ± 10.1</td>
<td>0.001</td>
</tr>
<tr>
<td>Overall Frequency</td>
<td>51.5 ± 4.8</td>
<td>48.6 ± 4.2</td>
<td>52.1 ± 4.7</td>
<td>0.047</td>
</tr>
<tr>
<td>Social Role</td>
<td>47.7 ± 6.9</td>
<td>43.9 ± 5.6</td>
<td>48.7 ± 6.5</td>
<td>0.046</td>
</tr>
<tr>
<td>Personal Role</td>
<td>56.3 ± 8.6</td>
<td>52.6 ± 6.6</td>
<td>57.0 ± 8.8</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Note: p-value compared mean ± standard deviation in two groups based on t-test. Scores ranged from 0–100, with higher scores indicating higher functioning levels.
function component. Low correlations were found between the TUG and BBS and the limitations or frequency with which older adults performed their daily life activities, as presented in the LLFDI disability component.

Our data agree with part of the data presented by Sayers et al. [26]. They found that the LLFDI function components demonstrate significant associations with performance-based measures of function. Sayers et al. found a significant association between the Short Physical Performance Battery and self-paced 400 m walk and the LLFDI disability limitation \( (r = 0.37 \text{ to } 0.44, \ p < 0.001) \) and disability frequency \( (r = 0.16 \text{ to } 0.20, \ p < 0.004) \) results [26]. These investigators suggested that the strength of the associations between the LLFDI and performance-based measures of function may better represent physical performance than previously used instruments. The performance measures were discrete physical tasks (without environmental or behavioral influence), which may partially explain why the limitation and frequency dimensions were not significantly associated with these tasks. In fact, Jette et al. suggest that the ability of older persons to perform life’s tasks (the limitation dimension) and their regularity in participating in life’s tasks (the frequency dimension) were related to each other but not to performance-based measures [13]. In the present study, the older persons actually performed functional tasks that they felt they were capable of performing \( (r = 0.43) \); however, the frequency of performing those tasks was unrelated to their physical functioning capabilities \( (r = 0.26) \). These results suggest that although older persons did not perceive a limitation in performing certain tasks, they performed particular tasks less often.

The LLFDI function component demonstrated known-group validity or the capacity to discriminate among older persons with different levels of functioning (i.e., cane and noncane users), which supports its validity. These results are consistent with previous findings that self-reported difficulty in function identifies distinct levels of functional ability [14,27].

Not surprisingly, these findings show that self-reported upper-limb function was not related to balance function and mobility. But, self-reported lower-limb function (in LLFDI) identifies functional limitations associated with balance function and mobility (BBS and TUG) in older persons. However, the disability component (limitation and frequency) was not revealed through the balance and mobility performance tests. Sayers et al. suggested that physical disability is not simply the inability to perform physical tasks but is also influenced by environmental and behavioral factors [26]. These findings also imply that, at each level of functional ability, a wide range of limitations and frequency of performance may exist. Verbrugge and Jette claimed that instruments that have inconsistently operationalized disability and have failed to consider the substantial behavioral components associated with physical disability limit the measurement of self-reported disability [28]. These aspects of function and disability are represented in the LLFDI components and subscales.

CONCLUSIONS

The present study indicated that, similar to the original English LLFDI, the Hebrew version of the LLFDI is a highly reliable and valid measure of function and disability in older adults. These data contribute to existing information on the performance of the LLFDI and support its reliability and validity in the geriatric population. It is recommended as an outcome measure in studies with older adults in which function and disability are important outcomes. Further research must determine whether the LLFDI is more sensitive than existing instruments and is responsive to change from relevant clinical and community-based interventions.

ACKNOWLEDGMENTS

This material is the result of work supported with resources and the use of facilities at the Multidisciplinary Center of Aging in the Faculty of Health Sciences at Ben-Gurion University of the Negev, Beer-Sheva, Israel.

The authors have declared that no competing interests exist.

REFERENCES


Submitted for publication October 19, 2006. Accepted in revised form May 9, 2007.