Measures of Physical Performance and Risk for Progressive and Catastrophic Disability: Results From the Women’s Health and Aging Study

Graziano Onder,1,2 Brenda W. J. H. Penninx,1 Luigi Ferrucci,3,4 Linda P. Fried,5 Jack M. Guralnik,6 and Marco Pahor1

1Sticht Center on Aging, Wake Forest University of the Health Sciences, Winston–Salem, North Carolina.
2Department of Geriatrics, Catholic University of the Sacred Heart, Rome, Italy.
3Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, Baltimore, Maryland.
4Laboratory of Clinical Epidemiology, INRCA Geriatric Department, Florence, Italy.
5Departments of Medicine and Epidemiology, The Johns Hopkins Medical Institutions, Baltimore, Maryland.
6Epidemiology, Demography, and Biometry Program, National Institute on Aging, Bethesda, Maryland.

Background. Physical performance measures can predict incident disability, but little research has assessed and compared how these measures predict progressive and rapid-onset (catastrophic) disability. The authors evaluated the ability of upper and lower extremity performance measures to predict progressive and catastrophic disability in activities of daily living (ADL), mobility, and upper extremity function.

Methods. The incidence of progressive and catastrophic disability was assessed semiannually during a 3-year period in 884 women participating in the Women’s Health and Aging Study I. Four-meter walking speed, balance, and chair stands tests were used to evaluate lower extremity function. The putting-on-blouse test, the Purdue pegboard test, and grip strength were used to assess upper extremity function. Summary performance scores (SPS) for the lower and upper extremities were calculated. Among participants in whom disability developed, those who reported no difficulty in the previous year were defined as cases of catastrophic disability, and those who previously reported little or some difficulty were considered to be cases of progressive disability. Cox proportional hazard regression analyses were used to evaluate the association of performance measures and time to incident disability. The predictive ability of performance measures was compared using receiver-operator characteristic curves.

Results. All lower and upper extremity measures, with the exception of grip strength, significantly predicted the onset of progressive ADL disability, but only walking speed was significantly associated with the onset of catastrophic ADL disability. The chair stands test, walking speed, and the lower extremity SPS were significantly associated with the onset of both progressive and catastrophic mobility disability. Only lower extremity individual tests and SPS significantly predicted the onset of both progressive and catastrophic upper extremity disability. The receiver-operator characteristic curves for ADL and mobility disability showed that all performance measures evaluated had a greater predictive ability for progressive than for catastrophic incident disability. This finding was not consistent for upper extremity disability. The areas under the curve for walking speed and lower extremity SPS were very similar, ranging from 0.58 to 0.81 and from 0.57 to 0.85, and the predictive ability of these two measures was the greatest for all disability outcomes assessed.

Conclusion. Physical performance measures of lower extremity and, in particular, walking speed and lower extremity SPS are valuable tools to predict different forms of disability, especially those with a progressive onset.

Prevention of physical disability in older persons is a primary goal of geriatric medicine. Many predictors of disability have been identified, including demographic characteristics, specific chronic conditions, and health behaviors (1,2). In many cases, disability has been considered a simple, static condition, however, particularly among physically frail older adults, it is a highly dynamic process with considerable diversity (3). The disability process is often the result of a progressive breakdown of the homeostatic equilibrium, described as a complex sequence of events in which multiple chronic diseases play an important role (4). However, some older persons may become disabled suddenly, as a consequence of a catastrophic medical event, such as stroke or hip fracture, without showing any previous sign of functional decline (5–7).

Standardized tests of physical performance, such as walking speed, balance tests, and grip strength, have been developed to assess a person’s ability to perform various movements of the upper and lower extremities that are required to accomplish common daily activities (8–11). Based on these tests, summary scores were created that can assess performance abilities along a broad spectrum of functioning (12–15). These physical performance tests and scores have proven valid and reliable (10,16,17), and they are now widely used in geriatric research because of their sensitivity to change over time (15) and their predictive validity for important outcomes such as self-rated health, institutionalization, hospitalization, falls, mortality, and onset of disability in diverse populations (16,18–26). However, little research has focused on how physical performance measures predict progressive versus rapid-onset types of disability. In addition, it has not been determined whether the predictive ability of various performance tests varies depending on the type of disability outcome (e.g., activities of daily living...
[ADL], mobility, or upper extremity function). In this study, using a sample population specifically selected to investigate changes in functional status over time, we assessed the ability of upper and lower extremity performance tests and scores to predict progressive and rapid (catastrophic) onset of ADL, mobility, and upper extremity disability.

**Methods**

We used data from the Women’s Health and Aging Study I, a 3-year longitudinal study conducted by the Johns Hopkins University and sponsored by the U. S. National Institute on Aging. Overall, 1002 women were randomly selected from residents of the eastern half of Baltimore City who reported difficulty in two or more of four functional domains (mobility and exercise tolerance, upper extremity function, basic self-care, and higher functioning tasks of independent living) and scored more than 17 on the Mini Mental State Examination (27). These women represented the approximately one third of the most disabled older women living in the community. A detailed description of the study design, data collection methods, and health and functional status of the study population is reported elsewhere (28,29).

**Performance Measures**

Trained examiners assessed lower extremity function using standardized measures, including 4-m walking speed (faster of two walks done at usual walking pace), the chair stands test, and the balance test. Upper extremity function was assessed using the putting-on-blouse test, the Purdue pegboard test, and the grip strength of the dominant hand (best of three trials). These tests are described elsewhere (29). As previously reported (15), for each individual task, participants unable to complete the test and worst performers were recoded as follows: walking speed, 0.09 m/second; grip strength, 5 kg; chair stands, 32.1 seconds; putting on blouse test, 233 seconds; Purdue pegboard test, 58.3 seconds.

**Summary Performance Scores**

To calculate summary performance scores (SPS), we rescaled individual measures to values ranging from 0 to 1, using the following formulas (higher scores signify better performance):

- Walking speed: $1 - \left( \frac{9}{\text{speed in cm/s}} \right)$
- Chair stands test: $1 - \left( \frac{\text{time in seconds}}{32.1} \right)$
- Standing balance test: $\left( \frac{\text{time in seconds}}{30} \right)$
- Putting-on-blouse test: $1 - \left( \frac{\text{time in seconds}}{233} \right)$
- Purdue pegboard test: $1 - \left( \frac{\text{time in seconds}}{58.3} \right)$
- Grip strength test: $1 - \left( \frac{1}{\text{grip strength in kg}} \right)$

We calculated SPS for lower extremities (range, 0–2.71) and upper extremities (range, 0–2.44) by adding the rescaled scores for the lower and upper extremity tests (15).

**Disability Outcomes**

Patients were examined in their homes at baseline and every 6 months thereafter for 3 years. At baseline and at the six semiannual follow-up assessments, the presence of disability in various domains of functioning was assessed. For the current study, in line with findings of a previous report (30), we assessed three outcomes: (a) disability performing ADLs, including bathing, dressing, eating, transferring from the bed to a chair and using the toilet; (b) disability walking across a room (mobility disability); and (c) disability lifting 4.5 kg (upper extremity disability). The level of disability was measured by reports of ‘‘no,’’ ‘‘a little,’’ ‘‘some,’’ ‘‘a lot of difficulty,’’ or ‘‘inability to perform a specific activity.’’ For this analysis, disability was considered present when participants reported a lot of difficulty or inability to perform a specific activity.

In line with findings of previous studies, to define progressive and catastrophic disability, we referred to the two assessments before the onset of disability (6,7). If participants reported a lot of difficulty or inability to perform a specific activity between follow-ups 2 and 6 and they had little or some difficulty at one or both of the semiannual assessments before the onset of disability (covering the previous year), we classified them as having progressive disability. If they reported no difficulty at both of these assessments, we classified the event as catastrophic disability. If information on disability at one of the two assessments before the onset of disability was missing, we based the disability status on one assessment. We excluded from the analyses any participants with missing data on both assessments before the onset of disability (ADL disability, $n = 4$; mobility disability, $n = 4$; upper extremity disability, $n = 6$). We excluded from the analyses any participants who reported a lot of difficulty or inability at baseline or at follow-up 1 (i.e., after 6 months).

**Data Analysis**

For 118 participants of the initial sample of 1002, we did not collect data on one or more performance measures at baseline. In the remaining sample of 884 participants, we performed separate Cox proportional hazard regression analyses to evaluate the association of each performance measure and SPS on time to onset of incident progressive and catastrophic disability versus no disability. We excluded from the analyses any participants who reported a specific disability at baseline or at the first 6-month follow-up. We considered the first report of disability at follow-up as the event of new disability, regardless of disability status reported in subsequent follow-up interviews. Those surviving with no evidence of new disability were censored at 3 years, those dying with no evidence of new disability were censored at the time of their deaths, and those lost to follow-up were censored at their last interview. To permit direct comparison among performance measures and scores, relative risks (RRs) for the onset of events were presented per standard deviation (SD) increase. We adjusted the analyses for age and race.

In additional analyses, we constructed receiver-operator characteristic (ROC) curves to compare the relative predictive ability of performance measures and scores. We calculated areas under the curves (AUC) and compared these areas with the method of Hanley and McNeil (31). We performed analyses using SPSS, version 10.1 (SPSS, Chicago, IL).

**Results**

The mean age of the 884 participants was 78.7 years (SD, 8), 72% were white, and the mean Mini-Mental State
The proportion of incident disability that is progressive and catastrophic is shown. Disability was considered present when participants reported a lot of difficulty or inability to perform that specific activity. Only participants free of disability at baseline and follow-up examination 1 were included in these analyses.

Examination score was 26.5 (SD, 3). Cardiovascular diseases and osteoarthritis represented the most frequent chronic conditions. Table 1 shows other characteristics of the study population, including the baseline means of the performance measures and scores.

Of 458 participants who did not report ADL disability at baseline and the first follow-up, 221 (48%) became disabled in ADL during the 3-year follow-up period. Of these 221 participants, 149 (67%) reported a lesser degree of disability in the previous year (progressive ADL disability), whereas disability developed in 72 (33%) who had not previously reported difficulty in ADLs (catastrophic ADL disability) (Figure 1). Similarly, mobility disability developed in 175 of the 684 (26%) women free of mobility disability at baseline and follow-up examination 1 during the 3-year follow-up period. We identified progressive mobility disability in 96 (55%) and catastrophic mobility disability in 79 (45%) participants. Finally, upper extremity disability developed in 202 of the 421 (48%) participants without upper extremity disability at baseline and follow-up examination 1 during the study period: We classified 110 (54%) cases as progressive and 92 (46%) as catastrophic upper extremity disability.

Table 2 shows the RRs of disability according to the performance tests and scores, after they were rescaled by their SDs. All lower and upper extremity performance tests and scores, with the exception of grip strength, significantly predicted the onset of progressive ADL disability, whereas only walking speed (RR, 0.72 per 0.31 m/second; 95% confidence interval [CI], 0.53 to 0.99) was significantly associated with the onset of catastrophic ADL disability. All lower extremity tests, the putting-on-blouse and both lower and upper extremity SPS, were significant predictors of progressive mobility disability, whereas only the chair stands test (RR, 1.39 per 8.4 seconds; 95% CI, 1.09 to 1.77), the walking speed (RR, 0.57 per 0.31 m/second; 95% CI, 0.41 to 0.80) and the lower extremity SPS (RR, 0.58 per 0.69 points in score; 95% CI, 0.41 to 0.82) were significantly associated with the onset of catastrophic mobility disability. Only lower extremity tests and SPS significantly predicted the onset of both progressive and catastrophic upper extremity disability. Among upper extremity measures, grip strength significantly predicted only the onset of progressive upper extremity disability (RR, 0.67 per 5.9 kg; 95% CI, 0.54 to 0.84), and the SPS predicted only the onset of catastrophic upper extremity disability (RR, 0.80 per 0.49 points in score; 95% CI, 0.65 to 0.99), whereas the putting-on-blouse test and the Purdue pegboard test were not associated with either of these two outcomes.

Table 3 presents areas under ROC curves, which assessed the ability of performance tests and scores to predict different disability outcomes. For ADL and mobility disability, all performance tests and scores evaluated had a greater predictive ability for progressive rather than catastrophic incident disability, whereas this finding was not
consistent for upper extremity disability. The predictive ability of walking speed and lower extremity SPS was very similar, and the AUCs of these two measures were among the greatest for all the disability outcomes assessed. For ADL progressive disability, the lower extremity SPS had the greatest area under the ROC curve (i.e., the highest predictive ability; AUC, 0.68; standard error [SE], 0.03), and it significantly differed from those of the balance test (AUC, 0.63; SE, 0.03; \( p \) vs lower extremity SPS < .05) and grip strength (AUC, 0.61; SE, 0.03; \( p \) vs lower extremity SPS < .05). The walking speed test was the best predictor of ADL catastrophic disability (AUC, 0.60; SE, 0.04), but none of the AUCs of the measures evaluated significantly differed from the one of the lower extremity SPS (AUC, 0.57; SE, 0.04). For progressive mobility disability, the AUC of the lower extremity SPS (AUC, 0.85; SE, 0.02) was significantly greater compared with those of all other measures, with the only exception of walking speed (AUC, 0.81; SE, 0.03). For catastrophic mobility disability, the AUC of the walking speed and the lower extremity SPS were the greatest (AUC, 0.65 and 0.63, respectively), but they did not significantly differ from those of the other measures. Finally, compared with lower extremity SPS, neither grip strength nor the upper extremity SPS had a significantly greater predictive ability for progressive upper extremity disability (AUC, 0.62 and 0.60, respectively), whereas, surprisingly, the AUC of lower extremity SPS was the highest for catastrophic upper extremity disability (AUC, 0.68; SE, 0.02), despite the fact that this result was statistically significant only for grip strength.

**DISCUSSION**

We found that physical performance measures of lower extremity function significantly predict the onset of progressive ADL, mobility, and upper extremity disability. In contrast, the upper extremity measures were much less consistently associated with the onset of disability in these tasks. For ADL and mobility disability, all the performance measures assessed had a higher predictive ability for progressive than for catastrophic outcomes. The predictive

---

**Table 2. Relative Risks of Incident Progressive and Catastrophic Disability for Lower and Upper Extremity Performance Measures**

<table>
<thead>
<tr>
<th>Performance Measures</th>
<th>Lower Extremities SPS</th>
<th>Upper Extremities SPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events ( (n = 189) )</td>
<td>Events ( (n = 72) )</td>
<td>Events ( (n = 110) )</td>
</tr>
<tr>
<td><strong>Progressive</strong></td>
<td><strong>Catastrophic</strong></td>
<td><strong>Progressive</strong></td>
</tr>
<tr>
<td>Chair stands (per 8.4 s)</td>
<td>1.54 (1.29–1.83)</td>
<td>1.14 (0.51–1.86)</td>
</tr>
<tr>
<td>Walking speed (per 0.31 m/s)</td>
<td>0.65 (0.52–0.82)</td>
<td>0.72 (0.53–0.99)</td>
</tr>
<tr>
<td><strong>Lower Extremities SPS</strong></td>
<td><strong>Upper Extremities SPS</strong></td>
<td></td>
</tr>
</tbody>
</table>

---

**Table 3. Area Under ROC Curves to Compare the Relative Predictive Ability of Performance Measures for Incident Progressive and Catastrophic Disability**

<table>
<thead>
<tr>
<th>Performance Measures</th>
<th>ADLs ( (n = 458) )</th>
<th>Mobility ( (n = 684) )</th>
<th>Upper Extremity ( (n = 421) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events ( (n = 199) )</td>
<td>Events ( (n = 72) )</td>
<td>Events ( (n = 96) )</td>
<td>Events ( (n = 79) )</td>
</tr>
<tr>
<td><strong>Progressive</strong></td>
<td><strong>Catastrophic</strong></td>
<td><strong>Progressive</strong></td>
<td><strong>Catastrophic</strong></td>
</tr>
</tbody>
</table>

---

* ROC = receiver-operator characteristic; ADLs = activities of daily living; SPS = summary performance score.

\( p \) vs lower extremity SPS < .05.

\( p \) vs lower extremity SPS < .01.

\( p \) vs lower extremity SPS < .001.
ability of walking speed and lower extremity SPS was very similar for various disability outcomes, and it was greater compared with other physical performance measures.

Overall, the incidence of disabilities in this sample was more elevated compared with a general population (32). This finding may reflect the fact that the Women’s Health and Aging Study I includes only women with functional limitations, a population that will likely develop physical disability (28). Indeed, in such a population, pathologic events may not be appropriately counteracted by compensatory mechanisms, and consequently many events may be followed by a significant change in functional status.

As shown by the ROC curve analyses, the ability of performance measures to predict incident ADL and mobility disability seemed higher for progressive than for catastrophic outcomes. This finding is in line with the concept that progressive disability is the result of a steady downward trend in functional abilities, often related to chronic health conditions, which ultimately leads to loss of independence. Therefore, cases of progressive disability are easily predicted by performance measures.

However, individual performance tests and SPS were also able to significantly predict incident catastrophic outcomes (particularly mobility and upper extremity disability). This finding may reflect the fact that women in the Women’s Health and Aging Study I, already had difficulty in at least two functional domains, and, in this vulnerable population, not all cases of catastrophic disability were related to a single and sudden event, but they were often consequences of rapid deterioration of general health, related to the presence of multiple predisposing factors, which can be identified by performance tests.

Overall, lower extremity measures, particularly walking speed and SPS, showed a greater predictive ability than upper extremity measures for incident disability outcomes including upper extremity disability. This finding suggests that lower extremity performance measures, and in particular walking speed and the summary score, may represent general measures of health and physical performance and not just specific indicators of localized poor function. Conversely, grip strength was a very specific predictor of future incident progressive upper extremity disability, suggesting that this measure may be valid only in its domain specificity. In addition, previous studies, lower extremity tests showed a greater sensitivity to change and a higher reliability over upper extremity tasks (10,15). These results suggest that lower extremity measures are preferable measures for studies of physical function.

In line with a previous observation (16), we also found that assessing walking speed alone is almost as good as performing the full battery of performance tests for the prediction of incident disability. This finding supports the routine measurement of walking speed in older persons in the clinical setting, where an objective measure of lower extremity functioning would be useful but too little time exists to perform a full performance battery. In addition, despite the fact that the 6-minute walking test has been proved to be a valid and reliable measure of physical function, the 4 m walking speed test offers several advantages: it is easier to administer, is less time consuming, and it is not limited by cardiorespiratory or peripheral vascular disease (33). Walking speed assessment may also be an efficient tool as the first step in screening many older persons to identify those at increased disability risk and to target disability prevention interventions (16,34). In the research setting, both the walking speed and the SPS are based on the three lower extremity tests may be valuable measures. However, the SPS may be more sensitive to change over time and reliability (15,17). In addition, in walking speed alone, the SPS explores a wider spectrum of functioning, which may be of value in capturing multiple risk factors in the pathway from disease to disability (34). Indeed, specific diseases and impairments may affect specific aspects of lower extremity function, which may then determine the characteristics of a person’s disability.

A potential limitation of this study is that it includes only women. However, it has been shown that the proportion of persons with incident progressive and catastrophic disability does not differ significantly between men and women (6,7), and performance measures have been proven to be similarly associated with incident disability events in both sexes (12). In addition, these findings are based on a population of persons who had some functional limitation at baseline, and therefore should be confirmed in healthier populations of older adults.

Conclusion

This study provides solid evidence that physical performance measures of the lower extremity predict different forms of disability, especially those with a progressive onset. In particular, the walking speed and lower extremity SPS could represent useful and preferable measures to identify persons at high risk for incident disability, which can be a target of disability prevention interventions and could serve as outcome measures for studies on physical function in older persons.

Acknowledgments

The Women’s Health and Aging Study was supported by the National Institute on Aging (contract N01AG12112). The work of Dr. Onder was supported by the Wake Forest Claude D. Pepper Older Americans Independence Center (NIA grant IP30 AG 21332-01).

Address correspondence to Graziano Onder, MD, PhD, Centro Medicina dell’Invecchiamento, Università Cattolica del Sacro Cuore - Policlinico A. Gemelli, L.go Francesco Vito 1, 00168 Roma, Italy. E-mail: graziano_onder@rm.unicatt.it

References

5. Ferrucci L, Guralnik JM, Pahor M, Corti MC, Havlík RJ. Hospital diagnosis, medical charges, and nursing home admissions in the year when older persons become severely disabled. JAMA. 1997;277:728–734.