LOWER-EXTREMITY FUNCTION IN PERSONS OVER THE AGE OF 70 YEARS AS A PREDICTOR OF SUBSEQUENT DISABILITY

JACK M. GURALNIK, M.D., PH.D., LUIGI FERRUCCI, M.D., PH.D., ELEANOR M. SIMONSICK, PH.D., MARCEL E. SALIVE, M.D., M.P.H., AND ROBERT B. WALLACE, M.D.

Abstract Background. Functional assessment is an important part of the evaluation of elderly persons. We conducted this study to determine whether objective measures of physical function can predict subsequent disability in older persons.

Methods. This prospective cohort study included men and women 71 years of age or older who were living in the community, who reported no disability in the activities of daily living, and who reported that they were able to walk one-half mile (0.8 km) and climb stairs without assistance. The subjects completed a short battery of physical-performance tests and participated in a follow-up interview four years later. The tests included an assessment of standing balance, a timed 8-ft (2.4-m) walk at a normal pace, and a timed test of five repetitions of rising from a chair and sitting down.

Results. Among the 1122 subjects who were not disabled at base line and who participated in the four-year follow-up, lower scores on the base-line performance tests were associated with a statistically significant, graduated increase in the frequency of disability in the activities of daily living and mobility-related disability at follow-up. After adjustment for age, sex, and the presence of chronic disease, those with the lowest scores on the performance tests were 4.2 to 4.9 times as likely to have disability at four years as those with the highest performance scores, and those with intermediate performance scores were 1.6 to 1.8 times as likely to have disability.

Conclusions. Among nondisabled older persons living in the community, objective measures of lower-extremity function were highly predictive of subsequent disability. Measures of physical performance may identify older persons with a preclinical stage of disability who may benefit from interventions to prevent the development of frank disability. (N Engl J Med 1995;332:556-61.)

I

N recognition of the importance of both extending life and increasing the number of years during which people are free of disability, the national health objectives for the year 2000 included as an overarching goal an increase in years of healthy life, with a full range of functional capacity at each stage of life. Disability in the older population, and the attendant need for informal and formal care, will increasingly affect older people, their families, and the health care system as the population continues to age.

The addition of functional evaluation to the traditional clinical examination provides information that is critical in the comprehensive assessment of elderly persons. Disability can be identified accurately through responses to a wide variety of questions about the ability to perform activities ranging from basic self-care to household activities and more strenuous tasks. Increasingly, functional status has also been characterized through the use of measures of physical performance, which are objective tests of subjects’ performance of standardized tasks, evaluated according to predetermined criteria that may include counting repetitions or timing the activity. These measures have been shown to predict outcomes such as falls, institutionalization, and death.

Nearly all studies involving the use of performance measures have been of groups that included disabled persons. In these populations, performance measures predict adverse outcomes such as institutionalization and death because of their ability to identify subjects with substantial reductions in function. What has not been clearly demonstrated is whether measures of physical performance can provide useful prognostic information about older persons who report little or no disability. In the analyses presented here, we tested the hypothesis that performance measures capture information on the range of functioning in people who are not currently disabled and that such information can be used to predict the subsequent onset of disability. The measures we used assessed general lower-extremity function, and the types of disability we evaluated involve lower-extremity function and have a substantial effect on the ability of older persons to remain independent.

Methods

Study Population

The data for this report were collected as part of the Established Populations for Epidemiologic Studies of the Elderly, a series of collaborative, longitudinal studies of aging initiated and funded by the Epidemiology, Demography, and Biometry Program of the National Institute on Aging. The eligible population consisted of all persons 65 years of age or older who lived in Iowa and Washington counties, Iowa. In this rural area the vast majority of residents live in 16 small towns. In 1981 and 1982, 3673 persons (80 percent of those eligible) participated in a comprehensive interview, previously described in detail. Nearly all the participants were white, and the education and income levels of this cohort were higher than in the other communities in the Established Populations for Epidemiologic Studies of the Elderly (East Boston, Massachusetts; New Haven, Connecticut; and five counties in north central North Carolina). Follow-up interviews were conducted annually for 7 years and again at 10 years.

At the sixth annual follow-up interview, conducted in 1988 and considered the base line for the analyses presented here, a questionnaire similar to that used in 1981 and 1982 was administered and physical-performance measures were added to the protocol for the first time. At that time, interviews were conducted with 2547 of the 2711 subjects (94.0 percent) not known to have died since 1982. To select a cohort of people without disability for these analyses, we ex-
cluded the following: 278 people living in institutions; 111 living at home who were unable to participate in the interview because of cognitive or physical impairment, for whom a proxy respondent had to be interviewed; and 70 who reported that they had a disability in the activities of daily living (defined below) or had mobility-related disability (defined as the inability to walk a half mile [0.8 km] or climb stairs without assistance); and 98 who reported no disability but were unable to complete the performance tests described below. Data on disability or performance measures were missing for 69 subjects (2.7 percent), most of whom had to be interviewed by telephone because they had left the geographic area.

Among the 1363 persons who reported no disability and completed the performance tests, 208 (15.3 percent) died during the next four years. Of those not known to have died, 33 (2.9 percent) were lost to follow-up, leaving 1122 for whom data on disability status at the four-year follow-up interview were available; these served as the primary subjects of this report. As compared with these subjects, those who were interviewed in 1988 but not included in the final analyses had a higher mean age (81.5 vs. 77.1 years), had a similar sex distribution (66.1 vs. 64.9 percent female), and were less likely to have 12 or more years of education (43.4 vs. 36.9 percent). Disability in the activities of daily living was present in 26.4 percent of those excluded from the analyses.

To evaluate a separate hypothesis that better scores on measures of physical performance are associated with improvement in the degree of disability, we performed additional analyses of data on 359 subjects who reported disability in mobility but no disability in the activities of daily living at baseline. The criteria for inclusion and exclusion were otherwise the same as those listed above.

Chronic Conditions at Base Line

The presence of chronic conditions — defined as a history of heart attack, stroke, cancer, or hip fracture — was ascertained from the subjects’ reports. The conditions considered were those that had previously been found to be associated with loss of mobility13 and for which data had been collected at all follow-up evaluations. Subjects were considered to have a history of heart attack, stroke, cancer, or hip fracture only if they reported having spent one or more nights in the hospital for the condition.

Measures of Physical Performance

Lower-extremity function was assessed by measures of standing balance, walking speed, and ability to rise from a chair that have previously been described in detail.9 Assessments were carried out in the subjects’ homes by interviewers who had been specially trained in the protocols. (The performance-measure protocol, data-entry form, and videotape used to train interviewers are available from the authors on request.) Correlations between observers of more than 0.93 for walking speed9 and test–retest correlations of more than 0.89 for walking speed,7 0.73 for repeated rising from a chair,16 and 0.97 for balance17 have been reported for these measures.

On the basis of the performance of more than 5000 persons at three sites in the Established Populations for Epidemiologic Studies of the Elderly, we created five performance scores (from 0 to 4) for each test, with a score of 0 representing the inability to complete the test and 4 the highest level of performance.7 The results of the tests were moderately correlated, with Spearman correlation coefficients for scores on the various tests ranging from 0.39 to 0.48. The subjects included in these analyses had scores ranging from 1 to 4 on each test.

For tests of standing balance, the subjects were asked to attempt to maintain their feet in the side-by-side, semi-tandem (heel of one foot beside the big toe of the other foot), and tandem (heel of one foot directly in front of the other foot) positions for 10 seconds each. The subjects were given a score of 1 if they could hold a side-by-side standing position for 10 seconds but were unable to hold a semi-tandem position for 10 seconds, a score of 2 if they could hold a semi-tandem position for 10 seconds but were unable to hold a full tandem position for more than 2 seconds, a score of 3 if they could stand in the full tandem position for 3 to 9 seconds, and a score of 4 if they could stand in the full tandem position for 10 seconds.

An 8-ft (2.4-m) walk at the subjects’ normal pace was timed, and the participants were scored according to quartiles for the length of time required. The time of the faster of two walks was used for scoring, as follows: ≥5.7 seconds, a score of 1; 4.1 to 3.6 seconds, a score of 2; 3.2 to 2.0 seconds, a score of 3, and ≤1.3 seconds, a score of 4.

Subjects were asked to fold their arms across their chests and to stand up from a sitting position once; if they successfully rose from the chair, they were asked to stand up and sit down five times as quickly as possible. Quartiles for the length of time required for this measure were used for scoring, as follows: ≥16.7 seconds, a score of 1; 15.7 to 16.6 seconds, a score of 2; 11.2 to 13.6 seconds, a score of 3; and ≤11.1 seconds, a score of 4.

A summary performance score was created by adding the scores for the tests of standing balance, walking, and repeatedly rising from a chair. The validity of this scale has been demonstrated in analyses showing a gradient of risk of admission to a nursing home and mortality along the full range of the scale.21 The subjects included in these analyses had summary scores ranging from 3 to 12. Among the 1122 subjects with follow-up data at four years, the base-line summary performance scores were distributed as follows: 0.09 percent had a score of 3, 1.3 percent a score of 4, 2.4 percent a score of 3, 6.3 percent a score of 6, 9.2 percent a score of 7, 15.4 percent a score of 8, 18.8 percent a score of 9, 18.8 percent a score of 10, 17.1 percent a score of 11, and 10.6 percent a score of 12.

Disability Status at Follow-up

The presence of disability at follow-up was assessed on the basis of reports of the subjects or their proxies. Because the performance tests at base line measured lower-extremity function, definitions of disability related to lower-extremity function were used. A three-level hierarchical scale based on the questions used to exclude those with disability at base line was used to classify the subjects at follow-up as having no disability, having mobility-related disability only (the inability to walk a half mile or climb stairs without help),13,14 or having a disability in the activities of daily living (the inability to perform one or more of the following basic activities without the help of another person: moving from a bed to a chair, using the toilet, bathing, and walking across a small room9,20) and mobility-related disability. Less than 0.7 percent of the subjects did not fit into one of these categories because they reported having a disability in the activities of daily living and no mobility-related disability; these persons were classified as having a disability in the activities of daily living.

Subjects’ reports of their degree of disability have generally been found to be highly reliable and valid.21,22 Evaluation of the reliability of the measures used here by repeated testing after three weeks showed agreement of 89 percent for measures of mobility and more than 90 percent for activities of daily living.23 Subjects’ reports of their ability to move from a bed to a chair, use the toilet, walk, and climb stairs have been shown to be valid in comparisons with direct, standardized observations of their performance on these tasks,24-27 with agreement rates ranging from 66 to 97 percent.25-27 The mortality rates in subgroups of this cohort defined by the disability categories appear to validate the use of our hierarchical scale as a predictor of mortality.28 Similarly, the validity of the measures of function in the activities of daily living29 and of mobility30 in predicting admission to a nursing home has been demonstrated by studies of the Established Populations for Epidemiologic Studies of the Elderly cohort.

Statistical Analysis

Disability at four years was analyzed according to base-line scores on the individual tests of lower-extremity function and the summary performance score. The Mantel–Haenszel chi-square test was used to test the linear association between performance on the tests and disability at follow-up. Multiple logistic-regression analysis was used to assess the independent association of the summary performance score with disability status at follow-up, after adjustment for age, sex, and the number of chronic conditions.

Results

Among the 1122 subjects who were not disabled at base line and were followed for four years, 212 (18.9 percent) had mobility-related disability and 112 (10.0
percent) had disability in the activities of daily living at follow-up. Table 1 shows disability status at four years according to scores on the individual tests of lower-extremity function. Those with better performance (higher scores) had significantly less disability in mobility and disability in the activities of daily living at four years.

The association of the summary performance score at base line with disability status four years later is shown in Figure 1. There is a clear gradient in the risk of mobility-related disability and disability in the activities of daily living across the full range of performance scores; higher scores, indicating better functional status at base line, are associated with a lower risk of subsequent disability.

Table 2 shows the results of two separate logistic-regression models used to estimate the relative risk of disability for two levels of performance, as compared with the highest level of performance, with adjustment for age, sex, and the number of chronic conditions. The subjects with summary performance scores of 4, 5, or 6 were 4.2 to 4.9 times more likely to have disability in the activities of daily living or mobility-related disability at four years than those with scores of 10, 11, or 12, and those with performance scores of 7, 8, or 9 were 1.6 to 1.8 times more likely to become disabled. Very similar results were found in analyses that assessed the risk of disability after one year of follow-up in 1276 subjects (data not shown). At one year, the relative risk of disability in the activities of daily living was 3.7 (95 percent confidence interval, 2.0 to 16.6) for those with performance scores of 4 to 6, and 2.1 (95 percent confidence interval, 0.8 to 5.6) for those with performance scores of 7 to 9; for mobility-related disability the relative risks were 5.0 (95 percent confidence interval, 3.0 to 8.2) and 2.1 (95 percent confidence interval, 1.4 to 3.1), respectively. Individual measures of performance were also significant predictors of disability in adjusted models. For example, the relative risks of mobility-related disability at four years for subjects with the poorest scores on the tests of walking, rising from a chair, and balance, as compared with those with the best scores, were 4.8 (95 percent confidence interval, 2.4 to 9.6), 4.1 (95 percent confidence interval, 2.3 to 7.2), and 1.9 (95 percent confidence interval, 1.2 to 2.9), respectively.

Analyses were conducted that further restricted the base-line population by excluding subjects who reported needing help in other activities of daily living (dressing or eating) or in doing heavy housework or who said they had any difficulty with lifting weights of more than 10 lb (4.5 kg); pushing or pulling a large object; stooping, crouching, or kneeling; extending their arms above shoulder level; or writing or handling small objects (these items are derived from Nagi30). After these exclusions, there remained 400 subjects with four years of follow-up. Although the number of subjects with disability at follow-up was small in this subgroup, relations between lower-extremity function as measured in the tests and subsequent disability were similar to those in the full cohort. At four years, the relative risk of disability in the activities of daily living was 7.1 (95 percent confidence interval, 2.4 to 20.9) for subjects with summary performance scores of 4 to 6, as compared with those with scores of 10 to 12, and 1.3 (95 percent confidence interval, 0.6 to 2.9) for those with performance scores of 7 to 9; the relative risks of mobility-related disability were 4.2 (95 percent confidence interval, 1.6 to 16.0) and 1.3 (95 percent confidence interval, 0.6 to 2.9), respectively.

Table 1. Presence and Type of Disability at Four Years of Follow-up, According to Scores on Base-Line Measures of Performance.

<table>
<thead>
<tr>
<th>Test and Score*</th>
<th>Status at Four Years</th>
<th>( \text{No disability} )</th>
<th>( \text{Mobility-related disability} )</th>
<th>( \text{Disability in ADL} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>percent of subjects</td>
<td>( \text{No disability} )</td>
<td>( \text{Mobility-related disability} )</td>
<td>( \text{Disability in ADL} )</td>
</tr>
<tr>
<td>8-ft (2.4-m) walk</td>
<td></td>
<td>35.6</td>
<td>31.1</td>
<td>33.3</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>59.9</td>
<td>24.6</td>
<td>15.5</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>71.8</td>
<td>19.9</td>
<td>8.3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>80.4</td>
<td>13.5</td>
<td>6.2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>59.7</td>
<td>25.3</td>
<td>15.1</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>67.0</td>
<td>21.9</td>
<td>11.1</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>74.9</td>
<td>16.4</td>
<td>8.7</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>79.3</td>
<td>13.9</td>
<td>6.8</td>
</tr>
<tr>
<td>Standing balance</td>
<td></td>
<td>37.7</td>
<td>31.2</td>
<td>31.2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>53.4</td>
<td>30.5</td>
<td>16.0</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>67.8</td>
<td>20.9</td>
<td>11.3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>77.7</td>
<td>15.4</td>
<td>6.2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>( \text{No disability} )</td>
<td>( \text{Mobility-related disability} )</td>
<td>( \text{Disability in ADL} )</td>
</tr>
</tbody>
</table>

*Scores ranged from 1 to 4, with 1 indicating the poorest performance, and 4 the best performance. \( P<0.001 \) for the association between the scores on each performance measure and subsequent disability status, by the chi-square test.

†ADL denotes activities of daily living.

Figure 1. Disability Status at Four Years According to the Base-Line Summary Performance Scores among 1121 Subjects with No Disability at Base Line. Higher scores indicate better performance on the tests and thus better functional status. One person with a score of 3 has been excluded. \( P<0.001 \) for the association between performance scores and disability status, by the chi-square test. ADL denotes activities of daily living.
The results of this study are compatible with the report of Seeman and colleagues\(^\text{16}\) that in a highly functional but more diverse older population, poorer results on tests of physical performance at base line were associated with greater decline in scores on the performance measures three years later. A limitation of our study is that we did not validate the disability-outcome measures by direct observation of the subjects in this population. A further limitation is that we included only the subjects who survived for four years. Nonetheless, the association of performance with disability was very similar at one and four years, indicating that lower-extremity function predicts the onset of disability in both the short and the long term.

A number of risk factors for disability have been reported, including specific chronic diseases, health-related behavioral factors, and sociodemographic characteristics.\(^\text{15,31-33}\) Although these factors can be used to identify persons who are at increased risk of disability, the characterization of a functional state that predicts disability could also be useful in studying the process by which people become disabled and in developing strategies for intervention. The World Health Organization and others have proposed models to explain the functional consequences of disease.\(^\text{30,34,35}\) Although they differ in certain aspects, these models portray a progression from disease through various stages of functional change that are compatible with the path from decreased lower-extremity function to disability that we used as the framework for this study. Fried et al.\(^\text{36}\) have hypothesized the existence of a stage of preclinical disability in which there is a decrease in functional ability, a need to use compensatory strategies, and a greater risk of functional decline and of the onset of disability. The demonstration that poorer scores on tests of physical performance in subjects without disability predict the subsequent development of disability provides evidence that this state of preclinical disability can be identified.

Lower-extremity function probably predicts the subsequent development of disability in large part because it reflects the effects of chronic disease, coexisting conditions, and physiologic decline that have not yet caused frank disability. Although the presence of certain diseases at base line was controlled for in our analysis...

### Table 2. Adjusted Relative Risk of Disability at Four Years of Follow-up, According to the Summary Performance Score at Base Line.

<table>
<thead>
<tr>
<th>Summary Performance Score</th>
<th>No. of Subjects†</th>
<th>Disability in ADL</th>
<th>No. with Disability</th>
<th>RR (95% CI)</th>
<th>Disability in Mobility</th>
<th>No. with Disability</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4–6</td>
<td>112</td>
<td>32</td>
<td>4.2 (2.3–7.7)</td>
<td>70</td>
<td>4.9 (3.1–7.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7–9</td>
<td>487</td>
<td>50</td>
<td>1.6 (1.0–2.6)</td>
<td>159</td>
<td>1.8 (1.3–2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10–12</td>
<td>522</td>
<td>29</td>
<td>1.0</td>
<td>94</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The relative risks have been adjusted for age, sex, and number of chronic conditions. ADL denotes activities of daily living. RR relative risk, and CI confidence interval. Higher scores indicate better performance on the tests and thus better functional status at base line. Subjects with scores of 10 to 12 serve as the reference group.

†One person with a summary performance score of 3 has been excluded.

‡Includes both those with mobility-related disability only and those with mobility-related disability and disability in the activities of daily living.

**DISCUSSION**

Our study provides strong evidence that measures of lower-extremity function in a nondisabled population predict the subsequent onset of disability. The performance measures showed a gradient of risk among our nondisabled subjects and could be used to identify subgroups at very low and high risk of disability. The results were similar when the analyses were further limited to subjects with a very high level of functioning, who reported that they had no difficulty or needed no help with a larger group of activities than was used initially to select the study population.

The Iowa population we studied is not representative of all older persons, and our findings should be replicated in nonrural, racially mixed populations. The results of this study are compatible with the report of Seeman and colleagues\(^\text{16}\) that in a highly functional but more diverse older population, poorer results on tests of physical performance at base line were associated with greater decline in scores on the performance measures three years later. A limitation of our study is that we did not validate the disability-outcome measures by direct observation of the subjects in this population. A further limitation is that we included only the subjects who survived for four years. Nonetheless, the associations...
yes, our data set did not include information either on an extensive list of conditions or on the severity of the diseases that were included. It is conceivable that if comprehensive data on existing diseases, their severity, and coexisting conditions were entered into the models, the performance score would no longer be an independent predictor of disability. However, disuse unrelated to disease status, a lower level of fitness, increased susceptibility to injury in those with impaired function, and a host of other modifiers of the relation between disease and disability make it unlikely that disease status alone would predict subsequent disability. Again, the tests of physical performance might capture information about many of these factors, including aspects of well-being, such as motivation, that are hard to measure.

Alternative explanations for the findings must also be considered. Subjects with poorer scores on the performance tests may already have had mild disability that was not assessed by the questionnaire used to select the study population. In this case, the measures of performance would have reflected the presence of mild disability, which in itself would place subjects at greater risk for subsequent serious disability. Another explanation is that those with poorer performance scores may have stated inaccurately that they were not disabled at base line and correctly reported their disability at follow-up. These two explanations are unlikely to explain our findings completely, however, since it was also shown that the performance measures predicted disability in analyses that included only the subjects who reported no disability at base line in a long list of functional areas.

Primary prevention of disease is the ultimate method of preventing disability. There is also a great need for prevention among persons who have chronic diseases that could lead to disability and among those whose mild reductions in function could progress to severe disability. These results have implications for clinical trials of interventions to reduce the onset and progression of disability. A number of studies have shown that interventions may improve characteristics such as strength, gait, and balance. Our study provides evidence that subjects with higher levels of functioning in these areas have a lower incidence of subsequent disability. This situation is analogous to others in which an intervention is shown to affect an intermediate end point, such as blood pressure, and epidemiologic studies have shown that the intermediate end point is related to an outcome of ultimate importance, such as the incidence of stroke. The only way truly to evaluate the efficacy of the intervention is through a randomized clinical trial in which the outcome measure is the end point of ultimate concern. In addition to providing evidence that intermediate end points are related to the incidence of disability, our results suggest that measures of performance may be valuable in identifying persons who are currently not disabled but are at increased risk for subsequent disability and therefore good candidates for a trial of an intervention to prevent disability.

We also found that in those with moderate disability — in this case, mobility-related disability — measures of lower-extremity function were related to the worsening or improvement of disability. Trials of interventions designed to improve performance in a group such as this could lead not only to the prevention of severe disability but also to the promotion of recovery from disability.

We are indebted to Yong Lee and Geane Schubert for their important contributions in the analysis of the data.

REFERENCES