## New and worsening cconditions and change in physical and cognitive performance during weekly interventions over six months: The Women's Health and Aging Study

Guralnik, J.M.; Ferrucci, L.; Penninx, B.W.J.H.; Kasper, J.D.; Leveille, S.; Bandeen-Roche, K.; Fried, L.P.

***published in***
Journals of Gerontology. Series A : Biological Sciences & Medical Sciences
1999

***DOI (link to publisher)***
10.1093/gerona/54.8.M410

***document version***
Publisher's PDF, also known as Version of record

Journal of Gerontology: MEDICAL SCIENCES
1999, Vol. 54A, No. 8, M410–M422



# New and Worsening Conditions and Change in Physical and Cognitive Performance During Weekly Evaluations Over 6 Months: The Women's Health and Aging Study

Jack M. Guralnik,[1] Luigi Ferrucci,[1,2] Brenda W.J.H. Penninx,[1,3] Judith D. Kasper,[4] Suzanne G. Leveille,[1] Karen Bandeen-Roche,[5] and Linda P. Fried[6]

[1]Epidemiology, Demography, and Biometry Program, National Institute on Aging, Bethesda, Maryland.
[2]Geriatric Department, "I Fraticini," National Research Institute (INRCA), Florence, Italy.
[3]EMGO Institute, Vrije University, Amsterdam, The Netherlands.
Departments of [4]Health Policy and Management, [5]Biostatistics, and [6]Medicine and Epidemiology, The Johns Hopkins Medical Institutions, Baltimore, Maryland.

***Background.*** Despite the large burden of chronic disease in older persons, especially those with disability, little research has documented changes in symptomatology over short periods of time. Additionally, although it has been demonstrated that medical conditions strongly affect functional level, short-term worsening in condition status has not been investigated for its impact on functional change.

***Methods.*** In a substudy of the Women's Health and Aging Study, 102 women with mild to severe disability received weekly home visits over a 6-month period. Each week they were queried as to the onset of 14 acute, generally self-limited conditions and the worsening or new diagnosis of 12 chronic conditions (condition reporting). They also received a battery of physical and cognitive performance tests.

***Results.*** There was a high rate of condition reporting over 24 weekly interviews. Nearly all women reported acute and episodic conditions; the average number of weeks of reporting one or more conditions was 11.8 per woman. For chronic conditions, the average number of weeks of reporting worsening of one or more conditions was 5.2 per woman. Multiple reports of onset or worsening of specific conditions were common, especially for arthritis of the hands, hips, knees, or feet; urinary problems; dizziness or unsteadiness on feet; and back pain. The total number of condition reports and number of weeks of condition reporting were generally not associated with an individual's slope of change in performance tests. For specific conditions, there were generally small and nonsignificant changes in performance in those who reported onset or worsening after 3 or more weeks of not reporting this.

***Conclusions.*** Older disabled women frequently report the onset or worsening of acute and chronic conditions. In weekly observations, these conditions are not related to changes in physical and cognitive performance measures. Possible reasons for this are that (a) condition reporting may not be valid, (b) changes or severity of conditions were of insufficient magnitude to affect functioning, or (c) performance measures are not sensitive to the kinds of changes in chronic and acute conditions that affect people from week to week. We concluded that performance measures are not useful in monitoring modest, short-term changes in health status, but may still be valuable for assessing more major changes in health and functioning over time.

THERE is a high prevalence of disease in the older population, and the co-occurrence of conditions, or comorbidity, is common. Women report more comorbidity than men. Employing a list of nine common diseases in a nationally representative cohort, it was found that two or more conditions were present in 61% of women and 47% of men age 70–79, and 70% of women and 53% of men age 80 years and older (1). In nearly 4,000 women living in the community who were screened for eligibility for the Women's Health and Aging Study, the average number of diseases, selected from a list of 14, was three (2). Only 5% of this representative sample of women aged 65 years and older had no conditions, 81% had two or more conditions, and 18% had five or more conditions. Older people also receive a large amount of medical care for their problems. In the National Ambulatory Medical Care Survey it was found that persons 65 to 74 years old had an average of 4.9 outpatient visits per year, and persons 75 years and older had an average of 5.9 visits per year (3).

Despite the clear documentation of high disease prevalence and medical care use in older people, little work has been done that documents how symptoms of disease vary over time, from month to month or week to week. Older persons commonly report that they have good weeks and bad weeks, and clinical observations reveal that the course of symptoms in diseases such as arthritis may be extremely variable (4). However, a systematic evaluation of reported worsening of conditions over a period of time has not been done. Furthermore, there is little documentation of how acute illnesses and episodic events, such as periods of fever or back pain, affect the older population. These kinds of conditions often are self-treated and do not come to the attention of the medical care system unless they result in a cascade of events, which in turn results in more serious consequences.

The presence of chronic disease has a large impact on functional status and disability, as demonstrated in both cross-sectional (1) and prospective (5,6) studies. The onset of major new diseases, such as hip fracture, stroke, and myocardial infarction, can have a particularly deleterious effect on functional abilities. However, it has not been demonstrated whether, at a lesser level of disease severity, illness worsening or particularly "bad weeks," so common in anecdotal reports from clinical practice, have an impact on the ability to function.

Using data from a sample of older women who were evaluated on a weekly basis over 6 months, this study describes the rate of reported chronic disease worsening and acute or episodic illnesses. It analyzes factors related to level of reporting and then evaluates the association of condition reporting with change in functional status, assessed using both physical and cognitive performance measures.

## Methods

### *Study Population*

This report uses data from the Weekly Disability Substudy of the Women's Health and Aging Study (WHAS). The WHAS is a study of the causes and course of physical disability in the one-third most disabled women living in the community. It is sponsored by the Epidemiology, Demography, and Biometry Program of the National Institute on Aging and is being conducted by The Johns Hopkins Medical Institutions. The sampling strategy, eligibility criteria, and assessment protocols have been described in detail previously (7). Briefly, an age-stratified random sample, with oversampling of women age 85 and older, was selected from Health Care Financing Administration Medicare files for 12 contiguous postal ZIP code areas in the eastern half of Baltimore city and a small part of Baltimore County. The sample, composed of four replicates, totaled 6,521 women, of whom 5,316 were eligible for screening and 4,137 participated in screening.

In those undergoing screening, study eligibility was based on disability status and cognitive functioning. Disability was assessed by self-report of degree of difficulty in performing 15 activities that were stratified into four domains, based on previous research showing clustering of difficulty in certain tasks (8): mobility/exercise tolerance (walking one-quarter mile, walking up 10 steps without resting, getting in and out of bed or chairs, doing heavy housework); upper extremity abilities (raising arms up over head, using fingers to grasp or handle, lifting and carrying 10 pounds); higher functioning tasks of independent living (using the telephone, doing light housework, preparing meals, shopping for personal items); and basic self-care (bathing, dressing, eating, using the toilet). Cognitive function was assessed with the Mini-Mental State Examination [MMSE; (9)]. Overall, 1,409 women met a priori criteria for study entry, reporting difficulty or needing help from another person in tasks in two or more domains and scoring 18 or higher on the MMSE. Of these, 1,002 agreed to participate in the baseline evaluation, which included separate in-home visits for a comprehensive interview, physical examination, and blood testing.

The Weekly Disability Substudy was designed to select approximately equal numbers of women in nine groups, defined by age (65–74, 75–84, and 85+ years) and three levels of disability (difficulty in two, three, or four domains). Women entering the second and third replicates of the main study were consecutively recruited until all nine cells were filled. Weekly substudy visits began one to two weeks after the baseline assessments were completed. Of the 102 women for whom data are presented, 34 were in each age group, 32 had disability in two domains, 33 in three domains, and 37 in four domains. Participants were interviewed and examined in their homes by specially trained interviewers on the same day and at the same time each week for 24 consecutive weeks, with assessments taking approximately one-half hour. Of 2,448 possible visits, 2,279 visits were completed (93.1%). There were an average of 1.7 missed visits per person; the number of missed visits ranged from 0 to 14 (one woman missed 14 visits, one missed 11, and the remainder missed 10 or fewer visits). The mean number of missed visits increased with age (0.8 in those 65–74, 1.7 in those 75–84, and 2.4 in those age 85+) and severity of disability (1.3, 1.7, and 2.0 in those with two, three, and four domains of disability, respectively).

### *Assessments*

*Baseline characteristics.*—A number of participant characteristics used in these analyses come from the initial comprehensive baseline assessment. The presence of 17 chronic conditions was ascertained using algorithms that utilized information from the interview, examination, x-rays, medications, physicians' reports, and medical record reviews (10). Definite and possible cases were summed to represent total baseline conditions, and the six most common baseline conditions were considered separately. Depressive symptoms were assessed using the Geriatric Depression Scale [GDS; (11,12)], and generally accepted cutpoints were applied (13). Participants were asked to rate their present health as excellent, very good, good, fair, or poor. Life quality was evaluated with the Perceived Quality of Life Scale (14). Participants were asked to rate 20 items on a scale of 1 to 10, with 10 being the highest perceived quality of life. Those with an average score of 0–3 were considered to have low quality of life compared to those with higher average scores. Four items from the anxiety subscale of the Hopkins Symptom Checklist were utilized, with a high level of anxiety defined as report of anxiety symptoms on two or more items (15). Two questions on personal mastery were adapted from Pearlin and Schooler (16). Women who disagreed strongly or disagreed somewhat were compared to those who agreed strongly or agreed somewhat to the statements, "I can do just about anything I really set my mind to" and "I often feel helpless in dealing with problems of life."

*Condition reporting at weekly visits.*—Chronic diseases and acute or episodic conditions were evaluated separately. To assess chronic disease status each week, participants were first asked if any of their old health problems had been worse in the past week. They were then shown a card with a list of chronic diseases and asked which had been worse. They were also asked if they had learned from a doctor that they had any new diseases from among the list. Reports of newly diagnosed chronic diseases were uncommon and were combined with worsened existing chronic diseases because they often represent long-term problems that are diagnosed when their worsen-

ing has prompted a visit to the physician. Participants were shown a card with a list of acute or episodic conditions and asked if anything on the list had been a problem in the past week. The term "condition reporting" is used throughout to indicate these specific definitions.

*Physical and cognitive performance measures at weekly visits.*—All physical performance measures were demonstrated by the examiner each week. Hand grip strength was measured using a JAMAR hand dynamometer (Model BK-7498, Fred Sammons, Inc., Burr Ridge, IL) in a sitting position with the wrist in a neutral position and the elbow flexed 90°. Usual gait velocity was evaluated on a 4-meter course. Participants were instructed to walk at their usual pace, as if they were walking down the street to go to the store. They began with both feet at the starting line and began walking after a verbal command, with timing started when the command was given. Subjects could use a walking aid but not the help of another person. Two tests were administered that had high correlation ($r = .96$), and the mean is used here. A single rapid walk was also done in a similar way. Participants were instructed to walk at a rapid pace, as fast as they could. Measurement was made of the time required to rise from a chair and return to the sitting position five times. Manual dexterity was tested using a pegboard test, with pins residing in a well and 10 holes arranged vertically. Participants were asked to individually place pins in the holes as fast as possible. The faster time of trials done with the left and right hands is presented here. Participants were timed picking up a key and opening a lock and putting on and buttoning a blouse. Cognitive functioning was assessed each week using two sets of items from the MMSE, delayed recall of three words and spelling a five-letter word backwards (9). To avoid using the same words each week, a series of eight sets of words were used three times over the 24 weeks. This approach was approved by the originator of the MMSE, who also reviewed the substitutions for the standard words (Personal communication, M. Folstein, 1993).

*Statistical Analysis*

To account for missing interviews when estimating number of reports of specific conditions over 24 weeks and number of weeks in which any condition was reported, condition rates for weeks in which participants were seen were multiplied by 24. For chronic and acute conditions separately, mean number of weeks of reporting of one or more conditions were calculated for each stratum of variables representing baseline characteristics. The age-adjusted association of baseline characteristics with number of weeks of condition reporting was examined by treating each non-missing week as an opportunity for condition reporting or not reporting and using logistic analysis to evaluate the odds ratio of reporting. Because each participant's 24 possible weeks of reporting are not independent, the generalized estimating equations approach (17) with exchangeable working correlation was used to account for within-person associations in computing regression estimates and calculating their standard errors, using the SAS GENMOD procedure (18).

For individual performance measures, each participant's mean and standard deviation over 24 weeks were calculated. The medians of these values for all participants are presented. The intraclass correlation coefficients [ICC; (19)] for performance measures done one week apart are presented for four sets of weeks. The first three sets, weeks 5 and 6, weeks 12 and 13, and weeks 19 and 20 were arbitrarily chosen to evaluate trends in the ICC over the course of the study. The fourth set of weeks represents the first occurrence in weeks 5 through 20 in which participants had no chronic or acute condition reports, and therefore represents a period of no reported change in medical status.

The association of condition reporting with change in performance was assessed by testing two hypotheses, represented schematically in Figure 1. Hypothesis 1 was that decline in performance over 24 weeks is directly related to amount of condition reporting (total number of weekly reports) during that period, after adjusting for baseline performance. Analyses were done separately for each performance test and for chronic and acute condition reporting. For each subject a linear regression was performed that regressed performance on time in weeks, with the slope and intercept for performance retained for further analysis. The hypothesis that amount of condition reporting is associated with slope of performance was tested using linear regression models in which the independent variables were condition reports and the intercept of each subject's regression of performance on time and the dependent variable was the slope of each individual's regression of performance on time.

The second approach to evaluating the impact of condition reporting on performance change was done by identifying the first 4-week period (index period) in which a participant made no condition report for a specific condition for 3 weeks and made a report on the 4th week (Figure 1). Those included in these analyses were required to have this pattern of condition reporting plus performance testing for at least two of the first 3 weeks and the 4th week and were thus a subset of individuals with any reporting of the condition. Hypothesis 2 was that there was a significant difference between the average performance in weeks 1 to 3 and performance in week 4.

If week 4 condition reporting coincided with not completing the performance test, these analyses could be biased. To evaluate whether the number of women not doing the performance test in week 4 was greater than expected, a comparison 4-week period was identified. This was an interval in which the individual did not report the condition in any of the 4 weeks and was selected to be the first 4-week period meeting this requirement after the index period or, if no such 4-week period was present, a 4-week period prior to the index period. Using the McNemar test, the percentage who did not do the performance test in week 4 was compared for the index period and the comparison 4-week interval.

## Results

*Rates of Condition Reporting*

The individual chronic diseases and acute or episodic conditions that were assessed on a weekly basis are listed in Table 1. The chronic diseases most often reported to have worsened were arthritis of the upper and lower extremities. Only 46.1% of women reported no worsening of arthritis in the hips, knees, or feet, and 55.9% reported no worsening of arthritis in the hands, arms, or shoulders. Nearly 20% of women reported worsening lower extremity arthritis five or more times. In 8 of

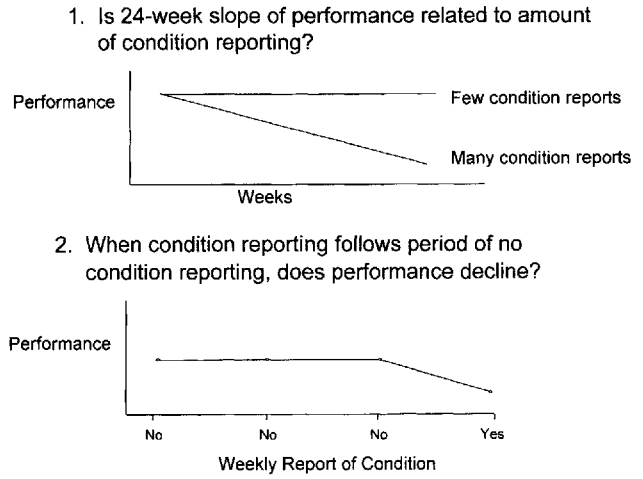


Figure 1. Schematic illustration of methods by which the association of condition reporting with change in performance was assessed.

the 12 categories of chronic diseases assessed, more than 5% of women reported worsening or new onset at least once in the 24 weeks. Overall, an average of 7.4 chronic condition reports were made for one or more conditions over the 24 weeks. Reports of acute or episodic conditions were more common than chronic condition reports. During the 24 weekly visits, over half of participants reported back pain, having no energy or being very tired, being dizzy or unsteady on the feet, or having leg swelling. These conditions plus urinary problems were reported by many women more than five times, and the number of weeks reported ranged from 0 to 20 or more for these five conditions. Over the 24 weeks, an average of 21.4 acute condition reports per woman were made for one or more conditions.

The data above refer to total number of condition reports, but it is also of interest to examine the number of weeks in which any condition report was made. The average number of weeks in which at least one chronic disease report was made was 5.2, and the average number of weeks in which acute or episodic conditions were reported was 11.8, with the range from 0 to 24 weeks. Figure 2 illustrates the distribution of number of weeks any chronic or acute condition report was made. Nineteen percent of women had no weeks of chronic condition reporting, another 22.5% reported chronic conditions on 1 or 2 weeks, but a substantial proportion reported chronic conditions on 5 or more weeks. Many weeks of acute condition reporting was common, and 14.7% reported at least one acute or episodic condition on 23 or 24 visits.

*Factors Associated With Condition Reporting*

Table 2 shows the distributions of baseline characteristics in the weekly substudy participants and their association with number of weeks in which conditions were reported. Worsening chronic conditions were reported more often in those with five or more baseline conditions, a greater level of disability, a GDS score of 10–13 (but not a score above 13), and arthritis. Those with high perceived quality of life were half as likely to report worsening chronic conditions. A model that adjusted for these variables and age demonstrated that only high level of disability, GDS score of 10–13, fair or poor self-rated health, and arthritis were independent predictors of reporting chronic disease worsening.

In age-adjusted analyses, having five or more baseline conditions, a GDS score of 10 or greater, fair or poor self-rated health, higher level of anxiety, feelings of helplessness, and history of stroke predicted more acute or episodic conditions. Less frequent acute condition reporting was observed in those with high perceived quality of life and a perception of mastery. After simultaneous adjustment for all these variables and age, only the measure of feelings of helplessness was a significant predictor of acute conditions. Importantly, age was not a significant predictor of reporting either chronic or acute conditions. Neither education (results not shown) nor race was significantly associated with chronic or acute condition reporting.

*Weekly Performance Tests*

For each weekly substudy participant, the mean performance was calculated for all completed performance tests. Table 3 shows the medians of these means for the entire sample. The ICC is also shown for tests done one week apart at four times during the 24-week study. Very high test–retest reliability was seen for grip strength, usual and fast-paced walks, repeated chair stands, and the pegboard test. Moderate to high levels of reliability were present for the lock-and-key test, putting on a blouse, and the two cognitive tests. Between weeks 5/6 and weeks 19/20, there was no clear improvement in reliability for tests with the highest ICCs, but modest improvement in reliability was seen for measures with lower ICCs. Reliability was not substantially better for 2-week intervals in which there was no chronic or acute condition reporting, theoretically the ideal weeks to test reliability.

For each participant, the slope of change in performance over 24 weeks was estimated for each performance test. Previous research demonstrated that for one of the tests, usual gait speed, there was evidence of some learning effect in certain age and disability subgroups but that a straight line adequately represented change over this observation period (20). To simplify these analyses, change in performance over time will be considered to be linear. The distribution of the slopes for each of the performance tests is illustrated in Figure 3, including the mean (asterisk), median (line inside box), 25th and 75th percentiles (box), and 5th and 95th percentiles (ends of dotted lines). The units under each box plot represent the slope in units per week change, with the units referenced in parentheses for each specific test. For example, grip strength slopes are in kg/week. For all tests the mean and median are quite similar and are close to 0. There is a distribution of slopes around 0, with evidence of both improvement and decline in function.

In estimating the background noise in the performance measures, which would interfere with ability to assess change over time, it is useful to have a measure of goodness-of-fit of the regressions. Estimating an $r^2$ is of no value, because regression lines with a slope of 0 will have an $r^2$ close to 0. For each performance test, to estimate how far, on average, a participant's individual test results were from her regression line, the absolute values of regression residuals were summed and then divided by the number of tests done, providing a measure of average residual for that participant. To illustrate the variation for the entire sample, a median value of these residuals is shown in

Table 1. Participant Report of Worsening or Onset of Chronic Conditions and Occurrence of Acute/Episodic Conditions Over 24 Weekly Visits*

| | Number of Reports (%) | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 1–2 | 3–4 | 5+ | Range | Mean |
| Chronic Conditions | | | | | | |
| Arthritis hips/knees/feet | 46.1% | 25.5% | 8.8% | 19.6% | 19 | 2.5 |
| Arthritis hands/arms/shoulders | 55.9 | 17.6 | 10.8 | 15.7 | 20 | 2.0 |
| Lung disease (emphysema/asthma) | 82.4 | 10.8 | 2.9 | 3.9 | 16 | 0.6 |
| Eye disease (cataracts/glucoma/other) | 84.3 | 9.8 | 2.0 | 3.9 | 9 | 0.5 |
| Poor circulation in legs | 80.4 | 15.7 | 2.0 | 2.0 | 6 | 0.4 |
| High blood pressure (hypertension) | 79.4 | 18.6 | 2.0 | 0.0 | 3 | 0.3 |
| Heart disease (heart attack/CHF/other) | 86.3 | 11.8 | 1.0 | 1.0 | 8 | 0.3 |
| Depression/other emotional problems | 89.2 | 8.8 | 1.0 | 1.0 | 7 | 0.2 |
| Parkinson's disease | 96.1 | 2.0 | 1.0 | 1.0 | 11 | 0.2 |
| Disc disease | 95.1 | 1.0 | 3.9 | 0.0 | 4 | 0.2 |
| Diabetes | 96.1 | 2.0 | 2.0 | 0.0 | 4 | 0.1 |
| Stroke | 97.1 | 2.0 | 1.0 | 0.0 | 3 | 0.1 |
| Acute/Episodic Conditions | | | | | | |
| Back pain | 30.4 | 26.5 | 6.9 | 36.3 | 23 | 5.2 |
| No energy or very tired | 22.5 | 40.2 | 9.8 | 27.5 | 20 | 3.7 |
| Dizzy or unsteady on feet | 37.3 | 25.5 | 6.9 | 30.4 | 24 | 4.0 |
| Leg swelling | 49.0 | 17.6 | 4.9 | 28.4 | 23 | 3.5 |
| Urinary (infection/incontinence) | 56.9 | 24.5 | 2.9 | 15.7 | 22 | 2.2 |
| Vomiting/nausea/upset stomach | 63.7 | 24.5 | 7.8 | 3.9 | 8 | 0.8 |
| Fever | 83.3 | 12.7 | 2.9 | 1.0 | 5 | 0.3 |
| Flu | 80.4 | 19.6 | 0.0 | 0.0 | 2 | 0.3 |
| Injury | 79.4 | 20.6 | 0.0 | 0.0 | 2 | 0.2 |
| Memory problems or confusion | 93.1 | 3.9 | 2.0 | 0.0 | 6 | 0.2 |
| Pneumonia | 95.1 | 3.9 | 1.0 | 0.0 | 3 | 0.1 |
| Fainting | 94.1 | 5.9 | 0.0 | 0.0 | 2 | 0.1 |
| Medication side effects | 95.1 | 4.9 | 0.0 | 0.0 | 2 | 0.1 |
| Surgery | 98.0 | 2.0 | 0.0 | 0.0 | 2 | <0.1 |

*Uses imputed rates for missing weeks.
*Source:* Women's Health and Aging Study, Weekly Disability Substudy.

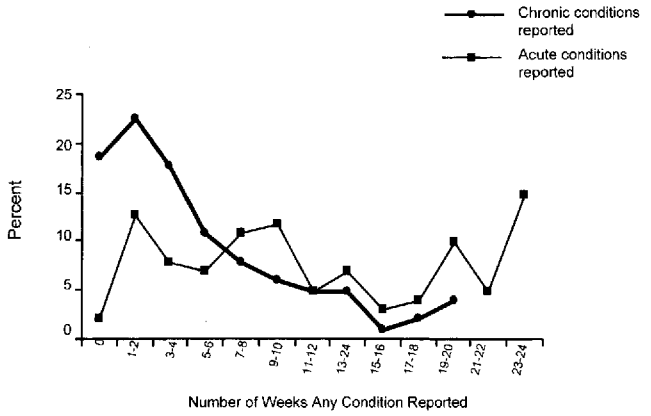


Figure 2. Distributions of number of weeks any chronic or acute condition was reported.

the right-hand column of Figure 3. The relationship of average residuals to the mean performance score gives some idea of how much variability is present over 24 weeks. For example, the average grip strength test was 1.09 kg from the regression line, a modest deviation in a test with a mean of 20.5 kg (Table 3). The average residual in the grip strength test represents about 5% of the mean. This percentage ranged from 6%–15% for the gait speed tests, repeated chair stands, pegboard test, and spelling backwards. The percentage ranged from 16%–20% for the blouse test and delayed memory and was nearly 33% for opening lock with a key.

*Condition Reporting and Change in Performance*

The first approach to test the association of condition reporting with change in performance (Figure 1, top schematic) yielded generally negative results in a large number of models that were tested. Models were created in which the independent variable was total number of acute condition reports, number of weeks of reporting acute conditions, total number of chronic condition reports, or number of weeks of reporting chronic conditions. Dependent variables that were evaluated were the slopes of performance from each individual's regression of performance on time in weeks, with separate models done for each of the performance tests. All models were adjusted for the intercept of each woman's individual regression of performance on time in weeks. No association was seen for acute condition reporting and slope of change in any of the performance tests. No association was seen for chronic condition reporting and measures of lower extremity functioning (normal and fast gait speed, repeated chair stands), grip strength, or cognitive tests.

Table 2. Association of Participant Characteristics With Number of Weeks of Reports of Worsening/New Chronic Conditions and Acute/Episodic Conditions

| | | Chronic Conditions | | | Acute Conditions | | |
|---|---|---|---|---|---|---|---|
| | | | Odds Ratio for Condition Reporting | | | Odds Ratio for Condition Reporting | |
| Characteristics | *n* | Mean Number of Weeks Reported† | Age-Adjusted | Fully Adjusted‡ | Mean Number of Weeks Reported† | Age-Adjusted | Fully Adjusted‡ |
| Age | | | | | | | |
| 65–74 | 34 | 5.1 | 1.0 | | 9.5 | 1.0 | |
| 75–84 | 34 | 5.5 | 1.1 | | 13.3 | 1.9 | |
| 85+ | 34 | 4.8 | 0.9 | | 12.4 | 1.6 | |
| Race | | | | | | | |
| White | 67 | 5.0 | 1.0 | | 12.2 | 1.0 | |
| Black | 35 | 5.5 | 1.1 | | 10.9 | 0.8 | |
| Number of baseline conditions | | | | | | | |
| 0–2 | 29 | 3.8 | 1.0 | 1.0 | 8.7 | 1.0 | 1.0 |
| 3–4 | 43 | 4.3 | 1.2 | 0.8 | 12.3 | 1.8 | 1.1 |
| 5+ | 30 | 7.8 | 2.6** | 1.4 | 14.0 | 2.4* | 1.2 |
| Number of disability domains | | | | | | | |
| 2 | 32 | 3.0 | 1.0 | 1.0 | 10.3 | 1.0 | |
| 3 | 33 | 4.2 | 1.5 | 1.1 | 11.2 | 1.2 | |
| 4 | 37 | 7.9 | 3.4** | 1.9* | 13.5 | 1.7 | |
| MMSE score | | | | | | | |
| 18–24 | 31 | 4.9 | 1.0 | | 13.5 | 1.0 | |
| 25–27 | 25 | 5.6 | 1.2 | | 11.0 | 0.7 | |
| 28–30 | 46 | 5.1 | 1.0 | | 11.0 | 0.8 | |
| Geriatric Depresison Scale | | | | | | | |
| <10 | 60 | 4.0 | 1.0 | 1.0 | 10.2 | 1.0 | 1.0 |
| 10–13 | 21 | 7.2 | 2.1* | 1.8* | 13.2 | 1.9* | 1.3 |
| ≥14 | 21 | 6.4 | 1.8 | 1.1 | 14.9 | 2.5** | 0.9 |
| Self-rated health | | | | | | | |
| Excellent/very good/good | 42 | 2.8 | 1.0 | 1.0 | 8.9 | 1.0 | 1.0 |
| Fair/poor | 58 | 6.8 | 3.1** | 2.5** | 13.8 | 2.2** | 1.7 |
| Perceived Quality of Life Scale | | | | | | | |
| Low | 59 | 6.2 | 1.0 | 1.0 | 13.9 | 1.0 | 1.0 |
| High | 43 | 3.7 | 0.5** | 1.0 | 8.8 | 0.4** | 0.7 |
| Anxiety | | | | | | | |
| Low | 60 | 5.0 | 1.0 | | 10.4 | 1.0 | 1.0 |
| High | 42 | 5.4 | 1.1 | | 13.7 | 1.7* | 1.2 |
| Mastery—Can do anything | | | | | | | |
| No | 31 | 6.6 | 1.0 | | 14.3 | 1.0 | 1.0 |
| Yes | 71 | 4.5 | 0.6 | | 10.6 | 0.6* | 0.9 |
| Mastery—Helpless | | | | | | | |
| No | 61 | 4.8 | 1.0 | | 9.5 | 1.0 | 1.0 |
| Yes | 41 | 5.7 | 1.2 | | 15.0 | 2.8** | 2.3** |
| Heart disease | | | | | | | |
| No | 49 | 4.5 | 1.0 | | 10.8 | 1.0 | |
| Yes | 53 | 5.8 | 1.4 | | 12.7 | 1.4 | |
| Arthritis | | | | | | | |
| No | 43 | 3.1 | 1.0 | 1.0 | 11.6 | 1.0 | |
| Yes | 59 | 6.6 | 2.5** | 2.1** | 11.9 | 1.0 | |
| Diabetes | | | | | | | |
| No | 86 | 5.4 | 1.0 | | 12.0 | 1.0 | |
| Yes | 16 | 3.8 | 0.6 | | 10.4 | 0.9 | |
| Stroke | | | | | | | |
| No | 88 | 4.8 | 1.0 | | 11.1 | 1.0 | 1.0 |
| Yes | 14 | 7.3 | 1.7 | | 15.8 | 2.4* | 1.2 |

*Continued on next page*

Table 2. Association of Participant Characteristics With Number of Weeks of Reports of Worsening/New Chronic Conditions and Acute/Episodic Conditions (Continued)

| | | Chronic Conditions | | | Acute Conditions | | |
|---|---|---|---|---|---|---|---|
| | | | Odds Ratio for Condition Reporting | | | Odds Ratio for Condition Reporting | |
| Characteristics | *n* | Mean Number of Weeks Reported† | Age-Adjusted | Fully Adjusted‡ | Mean Number of Weeks Reported† | Age-Adjusted | Fully Adjusted‡ |
| Hip fracture | | | | | | | |
| No | 94 | 5.3 | 1.0 | | 11.6 | 1.0 | |
| Yes | 8 | 4.0 | 0.7 | | 13.0 | 1.1 | |
| Hypertension | | | | | | | |
| No | 45 | 4.6 | 1.0 | | 10.5 | 1.0 | |
| Yes | 57 | 5.6 | 1.3 | | 12.8 | 1.5 | |

†Number of conditions adjusted for number of interviews missed.
‡Adjusted for age and all variables shown.
*$p < .05$; **$p < .01$.
*Source:* Women's Health and Aging Study, Weekly Disability Substudy.

Table 3. Physical and Cognitive Performance Measures: Means and Standard Deviations Over 24 Weeks and Intraclass Corelation Coefficients (ICC) of Assessments One Week Apart

| | Medians for All Participants of: | | ICC | | | |
|---|---|---|---|---|---|---|
| | | | Weeks (No Exclusions) | | | Weeks With No Reported Conditions Only |
| Measures | Individual's 24-Week Means | *SD* | 5/6 | 12/13 | 19/20 | |
| Grip strength (Kg) | 20.5 | 1.3 | .92 | .93 | .95 | .92 |
| Normal gait speed (cm/sec) | 67.8 | 10.3 | .88 | .89 | .88 | .87 |
| Fast gait speed (cm/sec) | 92.0 | 13.8 | .89 | .83 | .76 | .87 |
| Repeated chair stand (sec) | 16.2 | 1.9 | .76 | .84 | .89 | .88 |
| Pegboard test (sec) | 29.2 | 2.7 | .89 | .85 | .87 | .75 |
| Opening lock with key (sec) | 6.7 | 3.4 | .61 | .55 | .81 | .60 |
| Putting on/buttoning blouse (sec) | 79.7 | 23.0 | .56 | .74 | .76 | .72 |
| Spelling backwards (number correct) | 4.4 | 0.8 | .63 | .67 | .70 | .69 |
| Delayed memory (number correct) | 2.6 | 0.6 | .55 | .45 | .61 | .70 |

*Source:* Women's Health and Aging Study, Weekly Disability Substudy.

However, both number of chronic condition reports and number of weeks of reporting chronic conditions were significantly associated ($p < .05$) with three hand performance tests (putting on blouse, opening lock with key, and pegboard test), with greater condition reporting related to a more positive slope, indicating longer time to perform the task over the 24 weeks (data not shown).

Table 4 shows the results of analyses that evaluated how condition reports in the 4th week of a 4-week interval might be related to performance decline (Figure 1, bottom schematic). This was done for conditions whose relationship to the performance being tested is very direct (arthritis of hands and upper extremity performance; arthritis of hips, knees, or feet and lower extremity performance) and for more general conditions (dizziness, back problems, and flu-like symptoms, defined as report of fever, influenza, vomiting, or nausea) that might be expected to affect many aspects of performance. It was also done for any report of acute or chronic conditions. In general, there were small performance declines that were not statistically significant, although performance on some tests actually improved in the fourth week. In those with flu-like symptoms, dizziness, and worsening of one or more chronic conditions, chair stand time increased by about 1 second in the fourth week ($p < .05$). A very small group of women had a pattern of no condition reports for 3 weeks, followed by three or more chronic or acute condition reports in week 4. Even in this group there were modest declines in performance seen only in fast gait speed and chair rise time (Table 4). All analyses were also done comparing the single performance test in the third week of the interval to the week 4 performance test, and similar results were found (data not shown). Additionally, results were consistent when analyses were rerun starting after week 5 to eliminate potential for opposing actions of a learning effect and performance worsening that could occur during the early part of the study (data not shown).

One explanation for the modest or absent declines in function in the fourth week of the intervals of interest is that those most

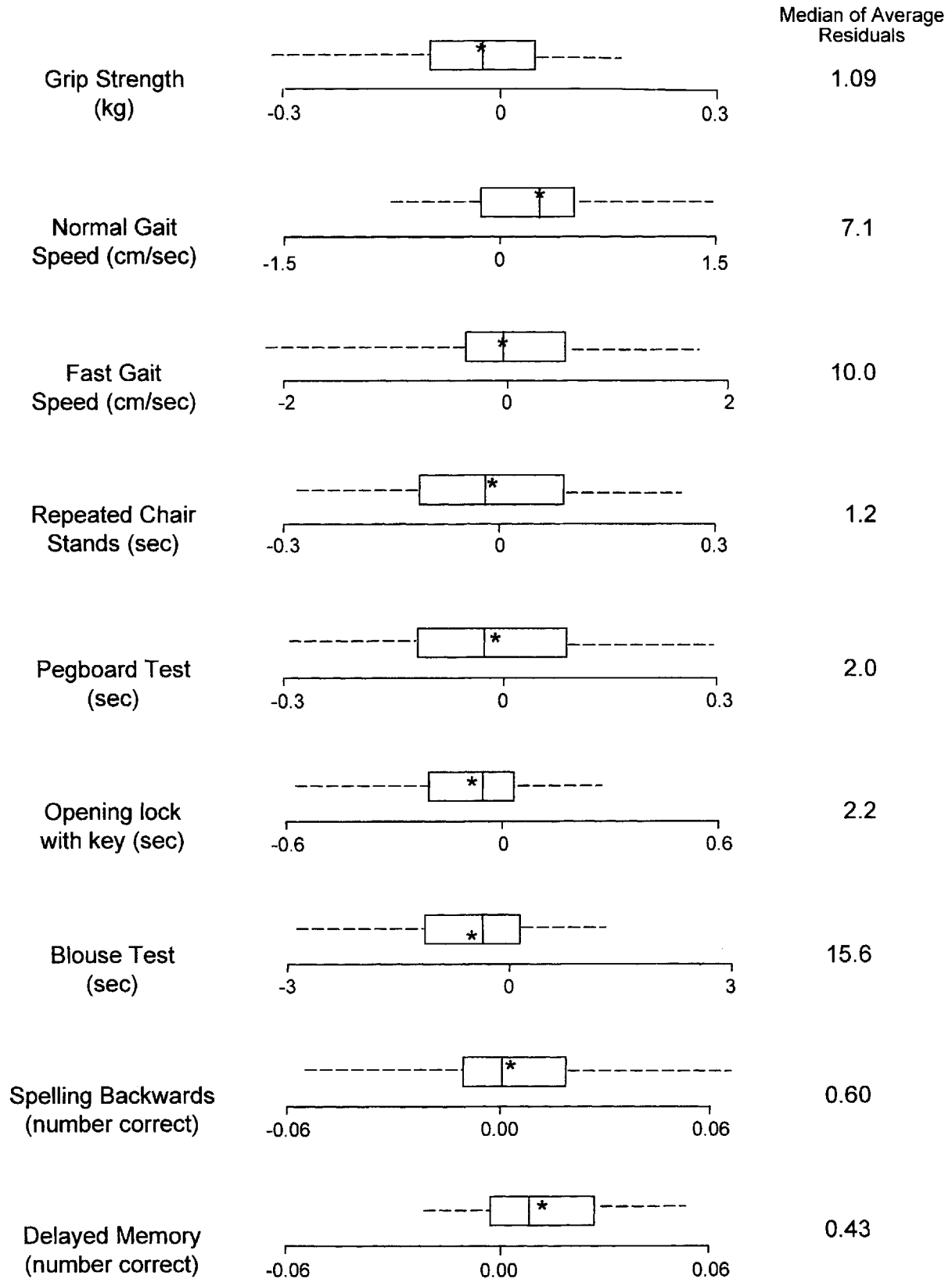


Figure 3. Distribution of slopes (units/week) of performance tests over 24 weeks and medians of all participants' mean residuals. Mean residual is the average distance of an individual's measurements from regression line. Distribution of slopes is shown as mean (asterisk), median (line inside box), 25th and 75th percentiles (box), and 5th and 95th percentiles (ends of dotted lines).

affected were unable to do the test. For example, for those 38 women who had a pattern of no reports of worsening of hand osteoarthritis for 3 weeks and a fourth week with a report, 31 women did the grip strength test in week 4 and seven (18.4%) did not (Table 4). Rates of missing performance tests were compared to the last week in another 4-week interval (control week 4) in which the condition was not reported in any week. For women with worsening of osteoarthritis of the hands, the rate of not performing the grip strength test was twice that for the control week 4, but other hand performance tests were actually missing at a slightly lower rate (Table 4). For women reporting flu-like symptoms and dizziness, the rate of not doing performance tests at week 4 was higher than in control week 4 for several tests, but these differences generally did not reach statistical significance due to the small sample size.

## Discussion

### *Condition Reporting*

In this study of disabled older women, the amount of condition reporting over a 6-month period was quite substantial. New onset or worsening of one or more of a list of 12 chronic conditions was reported, on average, in 5 of the 24 weeks. One or more acute or episodic conditions, from a list of 14 conditions, was reported on average in 12 of the 24 weeks. These data indicate that the ebb and flow of disease in older persons is a dynamic process and that they frequently suffer exacerbations and new or recurring symptoms.

Other research has also found the high level of condition and symptom reporting we saw in this study. In a small sample of persons aged 62 years and older who were asked about symptoms the previous day, more than 60% reported fatigue, weakness, or unsteadiness on the feet as well as pain, with one quarter reporting respiratory symptoms, including colds and fever (21). Studies using health diaries have been particularly effective in capturing symptom reporting in daily life (22). Compatible with our results of very high reporting of worsening arthritis, a representative sample of persons 65 years and older reported musculoskeletal symptoms in one-fifth of the days in which a diary was maintained (23). Further, over half of the days when any symptoms were reported were days that included musculoskeletal symptoms. The term "iceberg of morbidity" has been used to describe the large number of symptoms that people experience in their daily lives (23,24). In our study, the chronic diseases reported were already diagnosed by a physician, but the worsening of these conditions and the reported acute and episodic conditions may, in large part, not have come to medical attention.

### *Factors Associated With Condition Reporting*

We also investigated whether certain participant characteristics were associated with level of condition reporting. In age-adjusted analyses, predictors of number of weeks of chronic condition reports included disease status (greater number of diseases and arthritis), higher level of disability, depression, fair or poor self-rated health, and low perceived quality of life. These measures probably represent different aspects of how ill these women were or how ill they perceived themselves to be. The worsening symptoms of these conditions did not begin the day the weekly substudy began, and measures such as baseline self-rated health probably reflect a valid interpretation of an ongoing process rather than a traditional risk factor for the onset of an incident problem. It is notable that neither age nor cognitive status was related to reporting of worsening of chronic conditions.

In contrast to chronic conditions, acute and episodic condition reporting was related more strongly to psychological factors, including, in the age-adjusted analyses, depression, anxiety, and two aspects of mastery. The multivariate analyses in this cohort of only 102 women had limited statistical power, but it is interesting that the only variable to be significant in this model for acute conditions was the report on the mastery question that "I often feel helpless in dealing with problems of life." As in the chronic conditions, the acute and episodic conditions could have been bothersome at the time the baseline assessments were made, and an important question is whether the psychological variables represent the consequences of ongoing episodic disease. Repeated acute and episodic conditions, often felt to be out of an individual's control, can have a demoralizing effect. Alternative explanations for these findings are that the psychological variables are truly risk factors for acute and episodic conditions, or that they simply identify people more likely to report these conditions.

Previous research has investigated whether older people overreport illness. Comparisons of men and women are not possible in this study, but there is evidence that among people with the same chronic conditions diagnosed by physicians, women report more symptoms and general, rather than localized, symptoms (24). In a community-based study comparing condition reports of patients with confirmed physician diagnoses, those with a greater level of disability were more likely to overreport conditions (25), although there is considerable evidence of underreporting of conditions by survey respondents (26). Depression increases with age, and a higher proportion of those with depression meet the definition of hypochondriasis (27). However, the increase in depression with age can be attributed to the higher rate of actual chronic disease at older ages (28). Overall, evidence offers no support for an increase in hypochondriasis at advanced ages (29), and the overall increase in reporting of disease in older persons is probably related to real changes in health (30).

By measuring functioning each week over 24 weeks, this study provides a large amount of information about the stability and variability of the performance tests that were used. ICCs done for four different weekly intervals, including one done for weeks in which there was no condition reporting, were high and indicate excellent test–retest reliability. Randomly chosen pairs of weeks had ICCs no lower than weeks specifically chosen for no condition reporting, providing preliminary evidence that change in performance is not affected by the kinds of health status changes being evaluated. For all performance tests, the slopes of performance over time were close to zero for a large proportion of the sample, although some participants did have slopes indicating improvement or decline in function during the observation period. Floor and ceiling effects probably had a limiting effect on the slopes for certain participants, especially for the two cognitive tests, spelling five-letter words backwards and delayed memory, where many women did have the extreme values. There was, as expected, a certain amount of variability around these regression lines, although for most measures the

Table 4. Performance Measure Changes in Women Reporting Worsening or Onset of Conditions After 3 Weeks of Reporting No New Problems With Conditions

| Condition Report and Performance Test | Performance Data Weeks 1 to 4 (*n*) | | Performance Test Not Done (%) | | Mean Performance Score in Those With 4 Weeks of Data | | |
|---|---|---|---|---|---|---|---|
| | Weeks 1–3 | Week 4 | Week 4 | Control Week 4 | Weeks 1–3 | Week 4 | *p* |
| Arthritis, hands | | | | | | | |
| Grip strength (Kg) | 38 | 31 | 18.4 | 9.4 | 22.6 | 22.1 | 0.11 |
| Pegboard test (sec) | 40 | 38 | 5.0 | 8.8 | 29.3 | 29.6 | 0.45 |
| Lock/key (sec) | 41 | 39 | 4.9 | 11.4 | 6.7 | 7.2 | 0.63 |
| Blouse test (sec) | 36 | 33 | 8.3 | 12.9 | 76.8 | 71.6 | 0.06 |
| Arthritis, hips/knees/feet | | | | | | | |
| Normal gait speed (cm/sec) | 33 | 29 | 12.1 | 12.5 | 64.9 | 66.5 | 0.52 |
| Fast gait speed (cm/sec) | 26 | 23 | 11.5 | 12.9 | 88.7 | 85.6 | 0.54 |
| Repeated chair stands (sec) | 24 | 20 | 16.7 | 19.0 | 16.7 | 17.4 | 0.46 |
| Flu-like symptoms | | | | | | | |
| Normal gait speed (cm/sec) | 35 | 28 | 20.0 | 9.7 | 64.3 | 64.6 | 0.89 |
| Fast gait speed (cm/sec) | 33 | 27 | 18.2 | 6.9 | 85.4 | 93.0 | 0.40 |
| Repeated chair stands (sec) | 31 | 23 | 25.8 | 17.2 | 16.1 | 17.3 | 0.01 |
| Grip strength (Kg) | 38 | 32 | 15.8 | 8.8 | 21.4 | 21.3 | 0.78 |
| Blouse test (sec) | 37 | 31 | 16.2 | 14.7 | 86.1 | 89.7 | 0.64 |
| Spelling backwards (number correct) | 42 | 42 | 0.0 | 0.0 | 4.0 | 3.9 | 0.65 |
| Delayed memory (number correct) | 43 | 43 | 0.0 | 0.0 | 2.5 | 2.7 | 0.08 |
| Dizziness | | | | | | | |
| Normal gait speed (cm/sec) | 38 | 30 | 21.1 | 17.2 | 65.2 | 66.4 | 0.54 |
| Fast gait speed (cm/sec) | 33 | 23 | 30.3 | 13.6* | 84.4 | 85.5 | 0.72 |
| Repeated chair stands (sec) | 28 | 18 | 35.7 | 10.0* | 15.7 | 16.9 | 0.02 |
| Grip strength (Kg) | 44 | 39 | 11.4 | 8.3 | 21.4 | 21.0 | 0.15 |
| Blouse test (sec) | 40 | 32 | 20.0 | 6.5 | 76.9 | 81.9 | 0.39 |
| Spelling backwards (number correct) | 46 | 45 | 2.2 | 0.0 | 4.1 | 3.9 | 0.38 |
| Delayed memory (number correct) | 48 | 48 | 0.0 | 0.0 | 2.6 | 2.5 | 0.72 |
| Back Problems | | | | | | | |
| Normal gait speed (cm/sec) | 42 | 39 | 7.1 | 6.1 | 67.7 | 67.7 | 0.99 |
| Fast gait speed (cm/sec) | 40 | 38 | 5.0 | 6.5 | 92.0 | 89.5 | 0.30 |
| Repeated chair stands (sec) | 37 | 30 | 18.9 | 11.5 | 16.7 | 17.7 | 0.12 |
| Any Chronic Condition | | | | | | | |
| Normal gait seed (cm/sec) | 61 | 55 | 9.8 | 10.0 | 68.9 | 68.2 | 0.75 |
| Fast gait speed (cm/sec) | 54 | 46 | 14.8 | 11.1 | 92.1 | 91.0 | 0.69 |
| Repeated chair stands (sec) | 48 | 44 | 8.3 | 17.5 | 16.2 | 16.5 | 0.48 |
| Grip strength (Kg) | 65 | 61 | 6.2 | 7.1 | 21.5 | 21.4 | 0.76 |
| Blouse test (sec) | 58 | 56 | 3.4 | 14.3* | 84.0 | 82.9 | 0.75 |
| Spelling backwards (number correct) | 69 | 68 | 1.4 | 0.0 | 3.7 | 3.8 | 0.34 |
| Delayed memory (number correct) | 70 | 70 | 0.0 | 0.0 | 2.4 | 2.3 | 0.16 |
| Any Acute Condition | | | | | | | |
| Normal gait speed (cm/sec) | 59 | 50 | 15.3 | 12.5 | 69.7 | 67.6 | 0.28 |
| Fast gait speed (cm/sec) | 58 | 48 | 17.2 | 12.8 | 92.1 | 88.7 | 0.15 |
| Repeated chair stands (sec) | 54 | 44 | 18.5 | 14.3 | 15.6 | 16.3 | 0.03 |
| Grip strength (Kg) | 62 | 55 | 11.3 | 9.8 | 21.5 | 21.2 | 0.17 |
| Blouse test (sec) | 58 | 52 | 10.3 | 17.9 | 81.6 | 88.5 | 0.27 |
| Spelling backwards (number correct) | 68 | 66 | 2.9 | 0.0 | 4.0 | 3.9 | 0.52 |
| Delayed memory (number correct) | 68 | 68 | 0.0 | 0.0 | 2.4 | 2.4 | 0.96 |
| 3 or More Conditions | | | | | | | |
| Normal gait speed (cm/sec) | 6 | 5 | 16.7 | † | 78.4 | 78.4 | 0.99 |
| Fast gait speed (cm/sec) | 5 | 5 | 0.0 | | 110.8 | 97.9 | 0.16 |
| Repeated chair stands (sec) | 5 | 3 | 40.0 | | 13.6 | 14.7 | 0.64 |
| Grip strength (Kg) | 6 | 5 | 16.7 | | 27.8 | 38.4 | 0.57 |
| Blouse test (sec) | 6 | 5 | 16.7 | | 54.8 | 53.5 | 0.82 |
| Spelling backwards (number correct) | 6 | 6 | 0.0 | | 4.6 | 4.2 | 0.36 |
| Delayed memory (number correct) | 6 | 6 | 0.0 | | 2.8 | 2.8 | 0.36 |

*$p < .05$, comparing percentage not doing performance test to percentage not doing test in control weeks (control weeks indicate, in same subjects, percentage not doing performance test in 4th week of 4-week sequence in which no worsening or onset of conditions reported).

†Too few control weeks to do meaningful comparison.

*Source:* Women's Health and Aging Study, Weekly Disability Substudy.

average distance of observed values from the regression lines was modest, compared to the magnitude of the mean value for the measure (Figure 3). In one of the few studies with a design similar to ours, moderate variability in reporting of world views and religious beliefs occurred over 25 weeks (31). In that study it was argued that such variability was not simply random noise, as the factor structure found for components of the scales when studied in individuals over time was similar to that found across individuals.

*Condition Reporting and Performance*

Contrary to the original hypothesis, both analytic approaches shown in Figure 1 showed no clear and consistent relation of condition reports with change in functioning. There were two exceptions to the negative findings in this study. First, more reporting of chronic condition worsening over 24 weeks was associated with increasing performance time for three tests of hand function (putting on a blouse, opening a lock with a key, and completing the pegboard test). Second, the time to perform five chair stands increased in the week participants reported flu-like symptoms, dizziness, and any acute or episodic condition. The large number of statistical comparisons that were performed means that any significant findings must be interpreted with caution, but there was consistency in the relationships of several hand function tests with greater chronic disease reporting and of a single test (chair stand) with onset of several different acute or episodic conditions. This provides evidence that these findings may be due to more than chance alone. Compared to tests of walking and rising from a chair, which are routinely done, hand function tests were out of the ordinary for participants and may have been viewed more like competitive games. If so, motivation may have strongly influenced performance, and symptoms from worsening of chronic conditions might be reflected in declining motivation and performance. The chair stand test was the most vigorous of all tests performed, and as a result it may have been sensitive enough to pick up performance decrements in the week when a condition was reported.

There is strong evidence that performance measures of physical and cognitive functioning are powerful predictors of a variety of health outcomes such as institutionalization, disability, and mortality (32–37). In general, performance decrements themselves do not lead directly to these outcomes. Rather, the performance measures are predictive because they represent the overall burden and severity of disease and physiological decline. It would therefore seem that any change in health status should have an impact on performance. What, then, might explain our generally negative findings on the relationship of condition reporting and change in performance? Three alternative explanations should be considered.

First, condition reporting such as that elicited here may not be a valid indicator of disease onset or worsening. Previous research has demonstrated that there is a complex process involved in interpreting bodily change as an illness (38). In those with a high burden of disease, crossing the threshold for reporting the worsening of a condition may be influenced by a whole host of factors aside from actual pathological change or functional decline. Additionally, although this cohort certainly has considerable comorbidity, the study design may have encouraged condition reporting. Participants reported to interviewers whom they came to know quite well over the 6-month study. Furthermore, participants were invited to report their problems and were given a list of conditions to help with this report. Also, while certain conditions were categorized as acute and episodic, in some persons these can be long-term problems. For conditions such as back pain, fatigue, and dizziness it is difficult to differentiate whether multiple weekly reports represent independent episodes or a persistent underlying problem. Finally, worsening of conditions was queried for the past week, and if these conditions had abated by the time of the functional assessment, there might be no functional decline identified.

A second explanation for the negative findings is that the condition reporting was valid, and, as the results suggest, functioning was largely unaffected. First, the magnitude of change in severity of a condition that elicits a report of that condition might often be small. The functional consequences of more severe and important condition changes could be diluted by conditions that had minor impacts. Asking if participants sought care for their new or worsening conditions would have identified more severe problems (38), and it has been demonstrated that major hospitalizations and major illnesses over one year result in more restricted activity and bed days and poorer function (39). Our examiners also observed that many women suffering from more severe changes in their conditions tried hard to continue to perform their best. Those suffering an aggravation in hand osteoarthritis might report this but still work through this pain and maintain their usual performance on the hand function tests. We have come to expect and even laud athletes for maintaining performance in the face of health problems, but have perhaps failed to recognize more ordinary circumstances in which this occurs, such as in people with chronic conditions. In older persons it has been demonstrated that self-efficacy buffers the impact of diminished physical capacity on functional decline (40).

The third explanation for the findings is that performance measures are simply not sensitive to functional consequences of changes in chronic and acute conditions that affect people from week to week. These instruments are quite reliable, but it has been clearly articulated that reliability of an instrument does not guarantee that it will be sensitive to change (41). It is also important to note that sensitivity to change is specific to the kinds of changes being observed, and an instrument should not be globally classified as either sensitive or not sensitive to change. Women in our study with poor performance did not have this level of performance their entire lives, and it is likely that their decline to current levels was in many cases the result of major changes in health status. In a cohort followed for 3 years, greater declines in physical performance tests were seen in those who had incident health conditions and hospitalizations during follow-up (42). It thus seems likely that as the full cohort of women in the WHAS are followed over several years, performance decline will occur that can be related to more substantial changes in disease status than those assessed here.

The findings of this study do not support the hypothesis that performance measures respond to the kinds of changes in health status that older people report from week to week. While this is not a positive finding in terms of supporting superior measurement characteristics of performance testing, it may actually have a positive aspect as we utilize these measures to follow older populations. The activities of daily living (ADLs) are in-

sensitive to most changes in health status occurring in 90% of the older population, but when someone does change from nondisabled to disabled in ADLs this is a major transition that often reflects substantial worsening of disease. Analogously, consistent improvement or decline in performance might only be seen when an intervention effect or disease change is of large magnitude. In using performance testing in clinical trials, for example, we perhaps do not need to be concerned if the assessment weeks were good weeks or bad weeks for the participants. The measures are not sensitive to these kinds of perturbations but may change in important ways when more major transitions in health and functional status occur.

ACKNOWLEDGMENTS

This study was supported by contract N01-AG-1-2112 from the National Institute on Aging. The authors wish to thank Ase Sewall and James Hendrickson for their valuable contributions to this work.

Address correspondence to Dr. Jack M. Guralnik, Epidemiology, Demography, and Biometry Program, National Institute on Aging, Gateway Building, Suite 3C-309, 7201 Wisconsin Avenue, Bethesda, MD 20892-9205,

REFERENCES

1. Guralnik JM, LaCroix AZ, Everett DF, Kovar MG. *Aging in the Eighties: The Prevalence of Comorbidity and Its Association With Disability.* Hyattsville, MD: National Center For Health Statistics; 1989. Publication 170.
2. Fried LP, Bandeen-Roche K, Kasper JD, Guralnik JM. Association of comorbidity with disability in older women: The Women's Health and Aging Study. *J Clin Epidemiol.* 1998;52:29–37.
3. Woodwell DA. *National Ambulatory Medical Care Survey: 1995 Summary.* Hyattsville, MD: National Center For Health Statistics; 1997. Publication 286.
4. Felson DT. The course of osteoarthritis and the factors that affect it. *Rheum Dis Clin North Am.* 1993;19:607–615.
5. Guralnik JM, LaCroix AZ, Abbott RD, et al. Maintaining mobility in late life. I: Demographic characteristics and chronic conditions. *Am J Epidemiol.* 1993;137:845–857.
6. Boult C, Kane RL, Louis TA, Boult L, McCaffrey D. Chronic conditions that lead to functional limitation in the elderly. *J Gerontol Med Sci.* 1994;49:M28–M36.
7. Guralnik JM, Fried LP, Simonsick EM, Kaspar JP, Lafferty ME. The Women's Health and Aging Study: an introduction. In: Guralnik JM, Fried LP, Simonsick EM, Kaspar JP, Lafferty ME, eds. *The Women's Health and Aging Study: Health and Social Characteristics of Older Women With Disability.* Bethesda, MD: National Institute on Aging; 1995:1–9. NIH publication 95-4009. (Full text available at www.nih.gov/nia/edb/whasbook/title.htm)
8. Fried LP, Ettinger WH, Lind B, Newman AB, Gardin J. Physical disability in older adults: a physiological approach. Cardiovascular Health Study Research Group. *J Clin Epidemiol.* 1994;47:747–760.
9. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State": a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* 1975;12:189–198.
10. Fried LP, Kasper JD, Williamson JD, Skinner EA, Morris CD, Hochberg M. Disease ascertainment algorithms. In: Guralnik JM, Fried LP, Simonsick EM, Kaspar JP, Lafferty ME, eds. *The Women's Health and Aging Study: Health and Social Characteristics of Older Women with Disability.* Bethesda, MD: National Institute on Aging; 1995:Appendix E. NIH publication 95-4009.
11. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res.* 1982;17:37–49.
12. Burke WJ, Nitcher RI, Roccaforte WH, Wengel SP. A prospective evaluation of the Geriatric Depression Scale in an outpatient geriatric assessment center. *J Am Geriatr Soc.* 1992;40:1227–1230.
13. Lyness JM, Noel TK, Cox C, King DA, Conwell Y, Caine ED. Screening for depression in elderly primary care patients. A comparison of the Center for Epidemiologic Studies–Depression scale and the Geriatric Depression Scale. *Arch Intern Med.* 1997;157:449–454.
14. Merrill SS, Seeman TE, Kasl SV, Berkman LF. Gender differences in the comparison of self-reported disability and performance measures. *J Gerontol Med Sci.* 1997;52A:M19–M26.
15. Kasper JD, Simonsick EM. Mental health and general well being. In: Guralnik JM, Fried LP, Simonsick EM, Kasper JD, Lafferty ME, eds. *The Women's Health and Aging Study: Health and Social Characteristics of Older Women With Disability.* Bethesda, MD: National Institute on Aging; 1995:85–96. NIH publication 95-4009. (Full text available at www.nih.gov/nia/edb/whasbook/title.htm)
16. Pearlin LI, Schooler C. The structure of coping. *J Health Soc Behavior.* 1978;18:2–21.
17. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics.* 1988;44:1049–1060.
18. SAS Institute. SAS/STAT software: the GENMODE procedure, release 6.09, 6.07. Cary, NC: SAS Institute; 1993.
19. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psych Bull.* 1979;86:420–428.
20. Ferrucci L, Guralnik JM, Salive ME, et al. Effect of age and severity of disability on short-term variation in walking speed: The Women's Health and Aging Study. *J Clin Epidemiol.* 1996;49:1089–1096.
21. Brody EM, Kleban MH. Day-to-day mental and physical health symptoms of older people: a report on health logs. *Gerontologist.* 1983;23:75–85.
22. Verbrugge LM. Health diaries. *Med Care.* 1980;18:73–95.
23. Verbrugge LM, Ascione FJ. Exploring the iceberg: common symptoms and how people care for them. *Med Care.* 1987;25:539–569.
24. Kooiker SE. Exploring the iceberg of morbidity: a comparison of different survey methods for assessing the occurrence of everyday illness. *Soc Sci Med.* 1995;41:317–332.
25. Kriegsman D, Pennnix B, Van Eijk J, Boeke JP, Deeg D. Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly: a study on the accuracy of patients' self-reports and on determinants of inaccuracy. *J Clin Epidemiol.* 1996;49:1407–1417.
26. Edwards WS, Winn DM, Collins JG. Evaluation of 2-week doctor visit reporting in the national Health Interview Survey. *Vital Health Stat.* 1996;296:1–46.
27. Kramer-Ginsberg E, Greenwald BS, Aisen PS, Brod-Miller C. Hypochondriasis in the elderly depressed. *J Am Geriatr Soc.* 1989;37:507–510.
28. Blazer D, Burchett B, Service C, George LK. The association of age and depression among the elderly: an epidemiologic exploration. *J Gerontol Soc Sci.* 1991;46:S210–S215.
29. Barsky AJ, Frank CB, Cleary PD, Wyshak C, Klerman GL. The relationship between hypochondriasis and age. *Am J Psychiatry.* 1991;147:923–928.
30. Costa PT, McCrae RR. Hypochondriasis, neuroticism and aging: when are somatic complaints unfounded? *Am Psychol.* 1985;40:19–28.
31. Kim JE, Nesselroade JR, Featherman DL. The state component in self-reported worldviews and religious beliefs of older adults: the MacArthur Successful Aging Studies. *Psychol Aging.* 1996;11:396–407.
32. Gill TM, Williams CS, Richardson ED, Berkman LF, Tinetti ME. A predictive model for ADL dependence in community-living older adults based on a reduced set of cognitive status items. *J Am Geriatr Soc.* 1997;45:441–445.
33. Gill TM, Williams CS, Tinetti ME. Assessing risk for the onset of functional dependence among older adults: the role of physical performance. *J Am Geriatr Soc.* 1995;43:603–609.
34. Williams ME. Identifying the older person likely to require long-term care services. *J Am Geriatr Soc.* 1987;35:761–766.
35. Reuben DB, Siu AL, Kimpau S. The predictive validity of self-report and performance-based measures of function and health. *J Gerontol Med Sci.* 1992;47:M106–M110.
36. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol Med Sci.* 1994;49:M85–M94.
37. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *N Engl J Med.* 1995;332:556–561.

38. Haug ME, Musil CM, Warner CD, Morris DL. Elderly person's interpretation of a bodily change as an illness symptom. *J Aging Health.* 1997;9:529–552.
39. Wagner EH, LaCroix AZ, Grothaus LC, Hecht JA. Responsiveness of health status measures to change among older adults. *J Am Geriatr Soc.* 1993;41:241–248.
40. Mendes dLC, Seeman TE, Baker DI, Richardson ED, Tinetti ME. Self-efficacy, physical decline, and change in functioning in community-living elders: a prospective study. *J Gerontol Soc Sci.* 1996;51B:S183–S190.
41. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis.* 1987;40:171–178.
42. Seeman TE, Charpentier PA, Berkman LF, et al. Predicting changes in physical performance in a high-functioning elderly cohort: MacArthur Studies of Successful Aging. *J Gerontol Med Sci.* 1994;49:M97–M108.

*Received June 10, 1998*
*Accepted October 12, 1998*