Overestimation of Chronic Disability Among Elderly Persons

Thomas M. Gill, MD; Evelyne A. Gahbauer, MD, MPH

Background: Although there is no generally accepted definition for the term short-term disability, chronic disability has been defined as disability lasting or expected to last at least 90 days according to a protocol that was established by the National Long-Term Care Survey. We evaluated the validity of the established protocol and determined the accuracy of prevalence estimates of chronic disability among elderly persons in the United States.

Methods: Chronic disability was ascertained during a comprehensive assessment using the established protocol. Participants were subsequently classified as having chronic disability (the gold standard) based on the presence of disability during consecutive monthly interviews immediately before or after the comprehensive assessment.

Results: Of the 552 participants, 120 (21.7%) met criteria for chronic disability according to the established protocol. Of these, 30 (25.0%) and 39 (32.5%) did not meet criteria according to the gold standard under assumptions that were favorable and unfavorable (ie, stringent) to the established protocol, respectively. Conversely, of the 95 participants (17.2%) who met the gold standard criteria for chronic disability according to the favorable strategy and the 89 (16.1%) who met the criteria according to the stringent strategy, 5 (5.3%) and 8 (9.0%), respectively, did not meet criteria for chronic disability according to the established protocol. Relative to the established estimate of 7.0 million, our projections yielded about 2.0 million fewer chronically disabled elderly Americans in 1999.

Conclusion: Our results threaten the validity of the currently established protocol for ascertaining chronic disability and suggest that the burden of chronic disability among elderly Americans has been substantially overestimated.

Arch Intern Med. 2005;165:2625-2630

In a series of highly influential studies,1,2 Manton and colleagues, using data from the National Long-Term Care Survey, have demonstrated impressive reductions over the past 2 decades in the prevalence of chronic disability among elderly persons. Despite substantial growth in the elderly population (from 26.9 million in 1982 to 35.5 million in 1999), the number of chronically disabled Americans 65 years or older has decreased from 7.1 million to 7.0 million, leading to a relative reduction of nearly 25%. Although these important findings have been widely trumpeted in the medical literature4-6 and the lay press,7 relatively little attention has been given to how chronic disability was actually assessed. Previous research has shown that estimates of disability differ considerably across national surveys for a variety of reasons, including differences in the way that disability is assessed.8-10

In the National Long-Term Care Survey, a chronic disability has been defined as a disability lasting or expected to last at least 90 days.11,12 For a series of essential and instrumental activities of daily living, persons indicating that they need help (from another person or special equipment) are subsequently asked whether they have needed help for 3 months or longer, expect to need help for the next 3 months or longer, or will have needed help for 3 months or longer from beginning to end. Because it is uncertain whether elderly persons can accurately recall the exact onset of their disability, especially over an extended period of time, or can estimate the future duration of their disability, we decided to formally evaluate the validity of this protocol for ascertaining chronic disability.

To accomplish this objective, we used data from a unique longitudinal study that includes monthly assessments of disability in activities of daily living from a large cohort of elderly persons.11,12 We postulated that the protocol established by the National Long-Term Care Survey, which was implemented for the first time during a comprehensive assessment at 54 months, would lead to estimates of chronic disability that are inaccurate when compared with esti-
mates that were based on the presence of disability during consecutive monthly interviews immediately before and after the 54-month assessment (the gold standard).

**STUDY POPULATION**

Participants were members of the Precipitating Events Project, an ongoing longitudinal study of 754 community-living persons, 70 years or older, who were initially nondisabled in 4 essential activities of daily living—bathing, dressing, walking inside the house, and transferring from a chair. The assembly of the cohort has been described in detail elsewhere; the participation rate was 75.2%. Participants completed comprehensive assessments at 18-month intervals and monthly telephone interviews for the ascertainment of disability. The study protocol was approved by the Yale Human Investigation Committee, and all participants provided verbal informed consent.

The analytic sample for the current study included participants who completed the comprehensive assessment at 54 months. Of the 754 participants, 166 (22.0%) had died before the 54-month assessment, and 27 (3.6%) had dropped out of the study after a median follow-up period of 22 months. Of the remaining 561 participants, 3 (0.5%) refused to complete the 54-month assessment, and 6 (1.1%) had incomplete information on their disability status, leaving 552 participants in the analytic sample.

Compared with these participants, the 202 cohort members who were not included in the analytic sample were (at baseline) older (80.4 vs 77.7 years; P < .001), had more chronic conditions (2.3 vs 1.7; P < .001), and were more likely to be male (42.6% vs 32.8%; P = .01) and physically frail (59.9% vs 36.4%; P < .001). There were no significant baseline differences according to race or ethnicity, living situation, education, or cognitive status.

**DATA COLLECTION**

The research nurses who completed the 54-month assessments were kept blinded to the results of the monthly assessments. As described in 2 earlier reports, the comprehensive assessments and monthly telephone interviews were completed with a designated proxy for participants with significant cognitive impairment. At 54 months, 68 (12.3%) of the comprehensive assessments were completed by a designated proxy.

**54-MONTH ASSESSMENT**

During the 54-month assessment, data were collected on living situation; self-reported, physician-diagnosed chronic conditions; cognitive status as assessed by the Folstein Mini-Mental State Examination; and physical frailty, defined on the basis of slow gait speed as previously described. Chronic disability, defined as disability lasting at least 90 days (ie, 3 months), was ascertained for the first time during the 54-month assessment using the protocol established by the National Long-Term Care Survey, hereafter referred to simply as the established protocol. To our knowledge, alternative protocols for assessing chronic disability have not been reported. For each of the 4 essential activities of daily living, participants were asked, “At the present time, do you need help from another person to (complete the task)?” Participants who answered “No” were considered to be nondisabled for that task. Participants who answered “Yes” were subsequently asked up to 3 questions, shown in the following tabulation, to determine whether the disability was chronic.

**MONTHLY TELEPHONE INTERVIEWS**

Complete details regarding the monthly assessments of disability, including formal tests of reliability and accuracy, are provided elsewhere. During the monthly telephone interviews, participants were assessed for disability using standard questions that were identical to those used during the comprehensive assessments. For each of the 4 essential activities of daily living, we asked, “At the present time, do you need help from another person to (complete the task)?” Participants who needed help with any of the tasks were considered to be disabled. Participants were not asked about eating, toileting, or grooming. Disability in these activities of daily living is uncommon in the absence of disability in bathing, dressing, walking, or transferring from a chair; hence, these omissions will have little appreciable effect on our estimates of chronic disability. The reliability of our disability assessment was substantial (κ = 0.75) for reassessments completed within 48 hours and excellent (κ = 1.00) for reassessments performed the same day.

The accuracy of proxy reports, compared with reports from cognitively intact participants, was also found to be excellent, with κ = 1.00.

**STATISTICAL ANALYSIS**

The prevalence of chronic disability was determined on the basis of the 54-month assessment, and the number of chronically disabled activities of daily living was tabulated. Participants with chronic disability were subsequently classified according to the specific subtypes (shown in the tabulation in the “54-Month Assessment” subsection).

The overall objective of our analysis was to determine the validity of the established protocol for ascertaining chronic disability as implemented during the 54-month assessment. Data from the monthly telephone interviews were used to define the gold standard. Participants were classified as having chronic disability based on the presence of disability during consecutive monthly telephone interviews immediately before or after the 54-month assessment. To bolster the validity of the gold standard, we evaluated a subgroup of 186 participants who had no disability “at the present time” and found that only 2 (1.1%) reported disability “at any time during the last month.” To be conservative, we did not require that the subtype designation for the established protocol match that for the gold standard.
subsequently, under assumptions that were more stringent, that is, unfavorable. These assumptions differed depending on whether chronic disability was present or absent according to the established protocol, as summarized in Table 2 and described in detail in the next 2 subsections.

**PRESENCE OF CHRONIC DISABILITY ACCORDING TO ESTABLISHED PROTOCOL**

For the favorable strategy, the gold standard for chronic disability required disability during consecutive telephone interviews for intervals that could have potentially spanned 90 days or longer relative to the time of the 54-month assessment. For the stringent strategy, the gold standard required disability during consecutive telephone interviews for intervals of at least 90 days. To illustrate, for the first subtype of “previous 3 months” in Table 1, the gold standard required disability during months 51 to 54 for the favorable strategy and months 51 to 54 for the stringent strategy. For the second subtype of “next 3 months,” the gold standard required disability during months 55 to 57 for each of the 2 strategies. For the third subtype of “combination,” the gold standard required disability in consecutive months during any interval within 90 days of the 54-month assessment. This included months 52 to 54, 53 to 55, 54 to 56, or 55 to 57 for the favorable strategy and months 52 to 55, 53 to 56, or 54 to 57 for the stringent strategy.

**ABSENCE OF CHRONIC DISABILITY ACCORDING TO ESTABLISHED PROTOCOL**

In contrast, for this scenario the gold standard for chronic disability required disability during consecutive telephone interviews for intervals of at least 90 days relative to the time of the 54-month assessment for the favorable strategy and for intervals that could have potentially spanned 90 days or longer for the stringent strategy. To illustrate, in the last 2 columns of Table 1, the gold standard was achieved if disability was present during months 51 to 54, 52 to 55, 53 to 56, 54 to 57, or 55 to 58 for the favorable strategy and during months 51 to 54, 52 to 55, 53 to 56, or 54 to 57 for the stringent strategy. For both strategies, participants needed to report disability in 1 or more of the 4 essential activities of daily living during the 54-month assessment because this served as the reference point for the gold standard. Of the 432 participants who were non-disabled at 54 months, only 1 (0.2%) reported disability in the month immediately before and the month immediately following the 54-month assessment.

**ESTIMATES OF CHRONIC DISABILITY ACCORDING TO ESTABLISHED PROTOCOL AND GOLD STANDARD**

The validity of the established protocol for chronic disability was determined relative to the gold standard definitions for the favorable and stringent strategies, respectively. The 2 metrics of validity were sensitivity and specificity, which were defined, per convention, as the ability of the established protocol to correctly identify the presence (ie, true positive) and absence (ie, true negative) of chronic disability, respectively. These values, together with the prevalence of chronic disability from the 54-month assessment, were used to estimate the number of chronically disabled elderly Americans in 1999 according to the gold standard definitions, through a series of calculations (available on request) that required the use of positive and negative predictive values. According to the established protocol, the number of chronically disabled elderly Americans in 1999 was 7.0 million, as determined by the National Long-Term Care Survey.

Because our intent was to compare assessments at monthly intervals (ie, gold standard) with participants’ reports over these"
same intervals (i.e., established protocol), our estimates should not be affected appreciably by differences in the characteristics of the 2 study populations, the specific activities of daily living evaluated, or the methods used to define disability as long as the rates of chronic disability are comparable in our study and the National Long-Term Care Survey. Unlike sensitivity and specificity, positive and negative predictive values are highly dependent on the prevalence of the specific disorder.23 All analyses were performed using the statistical software SAS (version 8.2; SAS Institute, Cary, NC).

**RESULTS**

The characteristics of the analytic sample, at the time of the 54-month assessment, are shown in Table 3. The majority of participants were white women who did not live alone; about half were physically frail.

Of the 552 participants, 120 (21.7%) met criteria for chronic disability according to the established protocol. The mean (SD) number of chronically disabled activities of daily living evaluated, or the methods used to define disability as long as the rates of chronic disability are comparable in our study and the National Long-Term Care Survey. Unlike sensitivity and specificity, positive and negative predictive values are highly dependent on the prevalence of the specific disorder.23 All analyses were performed using the statistical software SAS (version 8.2; SAS Institute, Cary, NC).
tion of chronic disability prevalence but underestimation of disability occurrence, is the high rate of recovery among newly disabled elderly persons.\textsuperscript{12,29} Increasingly, disability has been viewed as a dynamic process, often characterized by multiple recurrent episodes rather than an enduring condition or progressive disorder.\textsuperscript{27,29} Our results indicate that many elderly persons cannot accurately estimate the duration of their disability episodes.

Despite differences in study populations and methods used to define disability, the established protocol yielded a rate of chronic disability in the current study (21.7%) that was comparable to that in the 1999 National Long-Term Care Survey (19.7%). Comparable rates enhance the likelihood that our projections are accurate.\textsuperscript{23} Over-sampling persons who were physically frail undoubtedly led to a higher rate of chronic disability, whereas our more narrowly focused definition of disability, which required personal assistance in at least 1 of 4 essential activities of daily living, likely led to a lower rate. Had we used a more expansive definition of disability that included instrumental activities of daily living or the need for special equipment, the validity of the established protocol may have been diminished further.\textsuperscript{8,10}

To address the modest asynchrony in time periods between the established protocol and monthly telephone interviews, we considered 2 gold standard definitions for chronic disability. Hence, our results provide upper and lower bounds for the number of chronically disabled elderly Americans. Because we could not account for the complex sample design of the National Long-Term Care Survey, these estimates should be interpreted cautiously. Our primary intent was to rigorously evaluate the validity of the currently established protocol for ascertaining chronic disability rather than to generate precise population-based estimates of chronic disability. Hence, our results provide upper and lower bounds for the number of chronically disabled elderly Americans.

Even though it is difficult, if not impossible, to exclude the possibility of measurement error, we believe that it is an unlikely explanation for our findings because our disability assessments were highly reliable, our rate of proxy reports was lower than that reported in the National Long-Term Care Survey,\textsuperscript{10} and the accuracy of our proxy reports was high. Although our participants are older than those in the National Long-Term Care Survey, there is no evidence that the accuracy of self-reported disability differs by age in the absence of cognitive impairment. Further reducing the possibility of bias, our 2 strategies for ascertaining chronic disability were completed independently of one another in the context of an ongoing longitudinal study. Our results cannot be attributed to underreporting of disability during the monthly telephone interviews because individuals who report no disability at the present time are unlikely to have had any disability during the past month.

In summary, the results of our study suggest that prevalence estimates of chronic disability have been spuriously inflated. Given the dynamic nature of disability, new strategies are needed to adequately capture the true burden of chronic disability among elderly persons.

### Table 4. Estimates of Chronic Disability According to Established Protocol and Gold Standard

<table>
<thead>
<tr>
<th>Chronic Disability According to Established Protocol</th>
<th>Favorable Strategy</th>
<th>Stringent Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present</td>
<td>Absent</td>
</tr>
<tr>
<td>PEP Study, No. of participants</td>
<td>Present</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Absent</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>95</td>
<td>457</td>
</tr>
<tr>
<td>United States,* No. of persons per million</td>
<td>Present</td>
<td>5,0</td>
</tr>
<tr>
<td></td>
<td>Absent</td>
<td>0,3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>5,3</td>
<td>30,2</td>
</tr>
</tbody>
</table>

Abbreviation: PEP, Precipitating Events Project.
*According to the established protocol used in the National Long-Term Care Survey, the number of chronically disabled elderly Americans in 1999 was 7.0 million. The total number of persons 65 years or older was 35.5 million.

Accepted for Publication: June 26, 2005.
Correspondence: Thomas M. Gill, MD, Dorothy Adler Geriatric Assessment Center, Yale University School of Medicine, 20 York St, New Haven, CT 06504 (gill@ynhh.org).
Financial Disclosure: None.
Funding/Support: The work for this study was funded by the National Institute on Aging, Bethesda, Md (grants R37AG17560 and R01AG022993), and was conducted at the Claude D. Pepper Older Americans Independence Center (grant P30AG21342), Yale University, New Haven, Conn. Dr Gill is the recipient of a Midcareer Investigator Award in Patient-Oriented Research (grant K24AG021507) from the National Institute on Aging.
Acknowledgment: We thank Denise Shepard, BSN, MBA, Martha Oravetz, RN, Alice Kossack, Barbara Foster, BS, and Alice Van Wie, BA, BS, for assistance with data collection; Wanda Carr and Geraldine Hawthorne for assistance with data entry and management; Peter Charpentier, MPH, for development of the participant tracking system; Joanne McGloin, MDiv, MBA, for leadership and advice as the Project Director; and Susan E. Hardy, MD, PhD, for her review of an earlier draft of this manuscript.
REFERENCES


