Four-Year Lower Extremity Disability Trajectories Among African American Men and Women

Fredric D. Wolinsky,1,2 Thomas R. Miller,2 Theodore K. Malmstrom,3 J. Philip Miller,4 Mario Schootman,4 Elena M. Andresen,5,6 and Douglas K. Miller7,8

1Iowa City Department of Veterans Affairs Medical Center (VAMC), Iowa. 2The University of Iowa, Iowa City. 3Saint Louis University, St. Louis, Missouri. 4Washington University in St. Louis, Missouri. 5Gainesville VAMC, Florida. 6University of Florida, Gainesville. 7Indiana University, Indianapolis. 8Regenstrief Institute, Inc., Indianapolis, Indiana.

Background. We examined 4-year lower extremity disability trajectories.

Methods. Nine hundred ninety-eight African American men and women 49–65 years old were evaluated at baseline and at four annual follow-ups. Lower extremity disability was the number of difficulties with nine standard activities of daily living (ADL), instrumental ADLs (IADL), and lower body function items. Mixed-effect models were used.

Results. The 9-item lower extremity disability measure had factorial validity and high reliability ($\alpha = 0.88$). The mean baseline lower extremity disability score was 2.43, and at the subsequent follow-ups it was 2.23, 2.35, 2.60, and 2.70. The mixed-effect model included significant random intercept and aging effects. Fixed factors with the largest effect sizes (all $p \leq 0.001$) were physical performance ($-0.238$ lower extremity disabilities per point on the Short Physical Performance Battery [SPPB]), fear of falling (1.094), poor or fair self-rated health (0.735), self-reported arthritis (0.659), clinically relevant levels of depression symptoms (0.641), body mass index (0.047 per kg/m² unit), aging (0.082 per year), and asthma (0.558).

Conclusions. To improve lower extremity disability trajectories among African Americans, interventions should focus on improving SPPB scores. In addition, fear of falling, poor or fair self-rated health, and clinically relevant levels of depression symptoms should be considered potential intervention candidates warranting further evaluation.

DISABILITY trajectories are central to gerontology and geriatrics (1,2). The increasing availability of longitudinal studies of older adults [e.g., Established Populations for the Epidemiologic Study of the Elderly (3), the Women’s Health and Aging Study (4), the Health and Retirement Study (5), the Survey on Assets and Health Dynamics Among the Oldest Old (6), and the National Long-Term Care Study (7)] brings into play more analytic options (8,9).

For the most part, previous studies of disability trajectories may be characterized by three groups. The first involves event history methods and whether disability onset occurs during a fixed time period (i.e., logistic regression) (10,11). This approach focuses on a single change segment and assumes that the underlying process is linear. Furthermore, attrition (i.e., loss to follow-up) is very difficult to address in these models, because lost participants are fully excluded.

The second group of disability trajectory studies considers multiple change segments within each participant, typically using Markov process models (12,13). This approach represents a notable advance because the longitudinal nature of the data is more fully exploited. Moreover, the reliance on unidirectional and linearity assumptions is eliminated, and attrition is more readily addressed (i.e., all available data on all participants is used prior to censoring). Limitations, however, remain. Principal among these limitations are the absence of “memory” from the analysis of one segment to another within participants and the failure to link multiple analytic segments within each participant into a meaningful growth curve or trajectory.

These limitations are fundamentally overcome in the last group of disability trajectory studies, which make use of mixed-effect models (14,15). Like Markov models, mixed-effect analyses use multiple segments within participants and address attrition (i.e., all available data on all participants is used prior to censoring). In addition, multiple segments within participants are linked, resulting in meaningful trajectory analyses. Moreover, random effects reflecting heterogeneity due to unmeasured factors can be included, as can interactions between main (i.e., fixed) effects and time (i.e., data collection wave), which represent slope differentials or discontinuities.

Regardless of which analytic method is used to examine disability trajectories, the contribution of such analyses depends entirely on the quality of the disability measure being studied. To be psychometrically sound, disability measures should have demonstrated validity, reliability, responsiveness to change, and clearly identifiable clinically important change thresholds (2). Despite considerable
work on the development and evaluation of disability measures, the psychometric properties of many indices that have been routinely used have not been fully documented, especially with regard to reliability and validity (1,2,16–18). Failure to document the psychometric properties of those measures has limited confidence in disability trajectory studies.

In this article, we use mixed-effects models to examine lower extremity disability trajectories using a psychometrically sound measure, as well as the associations of those trajectories with baseline factors. Data are from a large, population-based probability sample of older African Americans who were evaluated at five time points over a 48-month follow-up period. The set of baseline factors considered is large, diverse, and theory-driven, including sociodemographic, psychosocial, socioeconomic, physical performance, and morbidity measures.

Methods

Sample

We used data from the African American Health (AAH) project, for which the sampling design has been described elsewhere (19,20). The AAH includes 998 African Americans who were born from 1936 through 1950, and who lived in a poor, inner-city area of St. Louis, Missouri, or the near northwest suburbs. Equal numbers of participants were recruited from both areas, but given the smaller population of the inner-city area, disproportionate stratified random sampling was used. Therefore, all analyses involve probability-weighted data and are fully representative of the two sampling areas.

Inclusion criteria involved self-reported black or African American race, Mini-Mental Status Examination (MMSE) scores ≥ 16 (only 15 participants scored < 20) (21), and the ability and willingness to sign informed consent. At baseline (September 2000 through July 2001), in-home evaluations averaged 2.5 hours, and the response rate was 76%. Telephone follow-ups were conducted at 12, 24, and 48 months after baseline, and averaged about 15 minutes. In-home evaluations were performed 36 months after baseline, and averaged about 2 hours. Of the 998 original participants, 929, 893, 862, and 830 (unweighted Ns) were successfully re-interviewed at the 12-, 24-, 36-, and 48-month follow-ups, respectively. Because 61 deaths (i.e., a 1.5% annual mortality rate) were verified by the 48-month follow-up, the 4-year re-interview rate among the 936 surviving participants was 89%.

Lower Extremity Disability Trajectory

At each wave of data collection, a core set of standard questions was asked of all AAH respondents to monitor their ability to perform daily activities or their lower extremity physical functions. These questions were preceded by the following statement:

“The next group of questions are about how well you are able to do some activities by yourself and without using special equipment. Exclude any difficulties that you expect to last less than three months.”

The introduction for each question was: “Because of a health or physical problem, do you have any difficulty . . .?” For those participants who reported that they did not perform the activity or function for a nonhealth reason, we asked an additional question: “If you tried . . . , would you have any difficulty because of health or physical problems?”

The lower extremity disability measure used here is the simple sum of responses to nine of these core activities of daily living (ADL), instrumental ADL (IADL), and lower body function items or to the follow-on questions for respondents who reported not performing an activity or function due to nonhealth reasons. Bathing, transferring to bed, getting to places outside of walking distance, walking one-quarter mile, walking one-half mile, climbing up and down a flight of steps, standing for 2 hours or more, stooping–crouching–kneeling, and lifting and carrying 10 pounds were the activities and lower body functions addressed. Affirmative responses were coded as 1, and negative responses were coded as 0.

Baseline Correlates

Numerous risk factors have been identified for the onset of and changes in ADLs, IADLs, and lower extremity disabilities over time (22). Disability trajectory studies frequently classify these as sociodemographic, psychosocial, socioeconomic, physical performance, and morbidity factors (1,2). We included several measures from each category. Sociodemographic characteristics included age, sex, marital status, and falls history. Psychosocial factors included fear of falling and depression symptoms. Socioeconomic status indicators included education, income, and self-rated neighborhood quality. Physical performance measures included peak expiratory flow (L/s), grip strength (kg), and the Short Physical Performance Battery (SPPB) summary score (based on gait, chair stands, and hierarchical balance tests; 0 = worst, 12 = best) (23). Morbidity indicators included self-reports of vision, hearing, and overall health; cognitive ability (MMSE); body mass index (BMI, kg/m²); and self-reported history of arthritis, asthma, congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), kidney disease, diabetes, cancer, and stroke. More detailed information on the specific coding algorithms, reliability, and validity of the measures is available elsewhere (19,20).

Analytic Methods

The validity of the lower extremity disability measure was evaluated using exploratory factor analysis and standard simple structure criteria, including unidimensionality (only one eigenvalue ≥ 1.0), factor loadings > 0.50, communalities > 0.30, and scale-explained variation > 50%) (24–26). Reliability was assessed using internal consistency methods (α > 0.80) (24). Because the responsiveness to change and clinical importance of lower body function, ADL, and IADL items based on yes/no difficulty responses are well established (1,2); these are not further evaluated.

Mixed-effect model development followed standard guidelines (8,9), beginning with estimation of the most comprehensive model hypothesized to explain the mean
response using restricted maximum likelihood estimation. That model was then re-estimated using multiple covariance constraints to identify the proper structure. After selecting the appropriate structural constraint, all factors associated with the disability trajectory at the $p \leq .10$ level were retained, along with age and the intercept. Finally, each interaction term involving a fixed effect and aging (wave) was serially evaluated for inclusion in the model.

Because not all participants took part in each wave of data collection, the potential for attrition bias exists. Sensitivity analyses were conducted to evaluate this possibility. First, we used the baseline covariates included in the final mixed-effects model in a multiple logistic regression analysis to model loss to any follow-up. Because our mixed-effects models incorporate all follow-up data until the participant is censored from the analyses (i.e., lost to any further follow-up interviews for whatever reason, including death), the dependent variable in this model identified the 44 participants with no peak expiratory flow test is fully dissipated by the 48-month follow-up. It indicates that the initially protective effect for the binary risk factors, the effect estimates reflect the difference contrast between the two categories at baseline, or an intercept difference. For example, the partial effect of sex is that the average trajectory intercept for men is 0.41 lower extremity disability points lower for each additional SPPB point. Because the interval between each wave of data collection was 12 months, the fixed effect for wave is an aging (longitudinal) effect, whereas the fixed effect for age reflects cross-sectional differences in age. The only interaction effect (i.e., slope differential or discontinuity effect) identified was for the absence of a peak expiratory flow test at baseline. It indicates that the initially protective effect for participants with no peak expiratory flow test is fully dissipated by the 48-month follow-up.

The largest relative effect sizes (all $p \leq .001$), in descending order, were for physical performance ($-0.238$

### Results

#### Sample Characteristics

Of the 998 participants, 932, 888, 853, and 820 (weighted $N$s) were successfully re-interviewed at the 12-, 24-, 36-, and 48-month follow-ups, respectively. Among participants re-interviewed at least once, the mean age at baseline was 56.3 years (median age = 56 years), 37% were men, 37% reported annual income < $20K, 37% had fallen in the previous year, and 31% were afraid of falling. Forty percent rated their health as fair or poor, 48% reported arthritis, 12% reported asthma, 5% reported CHF, 6% reported COPD, 5% reported kidney disease, 8% reported stroke, and 23% reported clinically relevant levels of depression symptoms. The mean BMI was 30.4 kg/m², the mean grip strength was 34 kg, and the mean summary score on the SPPB was 8.0. The mean baseline lower extremity disability score was 2.43, and at successive follow-ups it was 2.23, 2.35, 2.60, and 2.70.

#### Psychometric Analyses

Table 1 contains the results of the exploratory factor analyses. As shown, the lower extremity disability measure satisfied all standard simple structure criteria, including unidimensionality (i.e., only one factor had an eigenvalue $\geq 1.00$), principal factor loadings (i.e., the smallest factor loading at any wave of data collection was .561), and item- and scale-explained variation at each wave of data collection (i.e., all communalities were $\geq .30$, and at each wave of data collection $\geq 50\%$ of the variance in the 9-item scale was explained by the single factor). Also shown in Table 1 are the internal consistency coefficients for the lower extremity disability measure, which were $\geq .80$ at all waves of data collection.

#### Mixed Model Analyses

Table 2 contains the crude and partial fixed-effect estimates from the mixed model analyses for all variables that were retained throughout the model building process, for which the covariance constraint was unstructured. Crude effect estimates were obtained from a series of models in which each independent variable (i.e., row) was individually included, along with wave as both a fixed and random effect. Adjusted effect estimates were obtained from a model including all of the independent variables as fixed effects, along with the inclusion of wave and the intercept as random effects.

The interpretation of the fixed-effect estimates is analogous to that in multiple linear regression analysis. Thus, for the binary risk factors, the effect estimates reflect the contrast between the two categories at baseline, or an intercept difference. For example, the partial effect of sex is that the average trajectory intercept for men is 0.41 lower extremity disability points higher than that for women. For the four nonbinary risk factors (i.e., age, BMI, SPPB score, and grip strength), the effect estimates reflect the difference in the intercept of the lower extremity disability trajectory at baseline associated with each 1-point increase in that independent variable. For example, the partial effect of the SPPB is that the average trajectory intercept is 0.24 lower extremity disability points lower for each additional SPPB point. Because the interval between each wave of data collection was 12 months, the fixed effect for wave is an aging (longitudinal) effect, whereas the fixed effect for age reflects cross-sectional differences in age. The only interaction effect (i.e., slope differential or discontinuity effect) identified was for the absence of a peak expiratory flow test at baseline. It indicates that the initially protective effect for participants with no peak expiratory flow test is fully dissipated by the 48-month follow-up.

The largest relative effect sizes (all $p \leq .001$), in descending order, were for physical performance ($-0.238$...
lower extremity disabilities per SPPB point), fear of falling (1.094), poor or fair self-rated health (0.735), arthritis (0.659), clinically relevant levels of depression symptoms (0.641), BMI (0.047 per BMI unit), aging (0.082 per year), and asthma (0.558). Statistically significant, albeit much more modest effects (i.e., < 0.400 or p > .001) were also found for each of the remaining covariates except age, neighborhood quality, COPD, and kidney disease, which were not independently associated with lower extremity disability trajectories. Overall, the model provided a good fit to the data, with a residual estimate of 1.52 (standard error [SE] = .046, p < .001). The estimate of the random effect for wave (aging) was 0.13 (SE = .016, p < .001), reflecting natural heterogeneity due to unmeasured factors.

**Sensitivity Analyses**

Given the limited number of participants for whom no follow-up interviews were conducted (unweighted N = 44; weighted N = 36), the multiple logistic regression analysis to evaluate the potential for attrition bias used forward stepwise inclusion techniques (results not shown). Included in those analyses were all of the baseline characteristics shown in Table 2. Of these, only CHF (adjusted odds ratio [AOR] = 6.521; p = .001), grip strength (AOR = 0.955; p = .015), and not performing a peak flow test at baseline (AOR = 2.302; p = .033) were independently associated with attrition. Although that model fit the data reasonably well (Hosmer–Lemeshow statistic = 9.757 at 8 df, p = .283; C statistic = .695), the inclusion of these factors in the model depicted in Table 2 and their significant effects suggest that the minimal potential attrition bias that might accrue from those few participants never having been re-interviewed at any follow-up wave is not problematic.

The second sensitivity approach repeated the psychometric analyses shown in Table 1 among the subset of participants who were successfully re-interviewed at all four annual follow-ups (unweighted N = 780, weighted N = 779). Although remarkably similar, the factor loadings for Waves 1 and 2 in those results (data not shown) were closer to the factor loadings for Waves 3 and 4. This finding indicates that the apparent improvements in factor loadings across waves in Table 1 results from the loss of the early attriters. The final sensitivity analyses involved replication of the mixed-effects model shown in Table 2 among the subset of participants who were successfully re-interviewed at all four annual follow-ups (unweighted N = 780, weighted N = 779). Those results (data not shown) were also remarkably consistent with those in Table 2. Taken together, these three sensitivity analyses suggest that it is unlikely that our findings result from potential attrition bias.

**Discussion**

This article makes five contributions to the literature. First, the analysis is based on a 9-item lower extremity disability measure for which substantial factorial validity and high reliability has been established at all waves of data collection. Second, the use of mixed models to estimate fixed and random effects for a large and diverse number of baseline risk factors notably reduces the risk that any observed associations are artifacts of potential confounding. Taken together, the robustness of this lower extremity...
disability measure and the rigour of the mixed-model analyses enhance confidence in the findings reported here.

The third contribution of this article is the identification of the SPPB summary score as the main driver predicting lower extremity disability trajectories. Although the identification of the import of the SPPB summary score is not new, it does validate prior studies by Guralnik and colleagues (23). What is new is that this study extends those results to a 48-month lower extremity disability trajectory to late middle-aged African Americans and to a multivariable analysis that includes a large and diverse number of potential confounders. These extensions are particularly salient in light of current federal efforts to reduce racial and ethnic disparities in health and health care (27), and suggest that targeting lower extremity physical abilities may be successful.

The fourth contribution of this study is the identification of three psychosocial factors—fair or poor self-rated health, fear of falling, and clinically relevant levels of depression symptoms—as also having important roles in predicting worse lower extremity disability trajectories. Because of the ease of asking participants about these factors (one simple screening question could suffice for each), we recommend that one simple screening question for each be considered for inclusion in the routine outpatient setting. Patients indicating elevated risk should be considered potential intervention candidates warranting further evaluation.

Perhaps the most interesting contribution of this study is the last. Lower extremity disability trajectories were not modified much over time. That is, there were no meaningful Wave × Fixed-effect interactions, suggesting that lower extremity disability trajectories are not dynamic. Further evidence of this was found in a model that included only age and wave (data not shown), in which both were significant and of about equal magnitude. This finding suggests a simple linear effect accumulation associated with age and aging.

Our study is not without limitations, and four of these warrant further mention. First, despite its demonstrated reliability and validity across all waves of data collection, our lower extremity disability measure ignores the upper body, which is also quite important to older adults. Time constraints in our telephone follow-up interviews at Waves 2, 3, and 5, however, precluded fielding a multi-item scale targeting upper body disabilities. Accordingly, we are unable to conduct similar analyses on upper extremity disability trajectories. Second, our study is limited by having only four follow-up interviews that occurred at annual waves of data collection. It is possible that the lower extremity trajectories may exhibit dynamic properties (i.e., baseline Characteristic × Time interactions) after our cohort has been followed for an entire decade. Thus, the lack of evidence of dynamic trajectories in these data should not be considered definitive. Third, although our study relied on probability-based sampling of all housing units and had excellent participation and retention rates, it is nonetheless restricted to participants in just two geographic areas of St. Louis, MO. Therefore, our results may not reflect what would be observed in other major metropolitan areas, or in a national study.

Finally, and most importantly, our study is restricted to African American men and women who were 49–65 years old at baseline. As a result, we cannot directly compare and contrast these results with those for whites or Hispanics. Moreover, we cannot determine which of the plausible explanations of our results are operative. On the one hand, we have previously shown that this sample has more health disadvantages than do African Americans, whites, or Hispanics observed at the same time in nationally representative samples (28), and that the same was true for an older sample that we previously studied from this inner-city area (29). Thus, the absence of evidence of dynamic lower extremity disability trajectories in the present AAH sample may reflect the fact that our participants reached high levels of lower extremity disability (and its concomitant ceiling effects) earlier than their counterparts, for whom dynamic trajectories may still be observed (until they catch up and plateau). On the other hand, and as suggested above, it may be that our AAH sample will exhibit dynamic lower extremity trajectory properties after our cohort has been followed for an entire decade, and aged into its sixth (for those 49–59 years old at baseline) and seventh (for those 60–65 years old at baseline) decades. Only time will tell.

ACKNOWLEDGMENTS

This work was supported, in part, by National Institutes of Health Grant R01 AG-10436 to Douglas K. Miller, Dr. Wolinsky is Associate Director of the Center for Research in the Implementation of Innovative Strategies in Practice (CRIISP) at the Iowa City VA Medical Center, and Dr. Andresen is a Core Investigator with the Rehabilitation Outcomes Research Center (RORC) at the North Florida/South Georgia Veterans Health System, both of which are funded through the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service.

CORRESPONDENCE

Address correspondence to Fredric D. Wolinsky, PhD, the John W. Colloton Chair in Health Management and Policy, College of Public Health, the University of Iowa, 200 Hawkins Drive, E-205 General Hospital, Iowa City, Iowa 52242. E-mail: fredric-wolinsky@uiowa.edu

REFERENCES


Received March 31, 2006
Accepted June 7, 2006
Decision Editor: Darryl Wieland, PhD, MPH