Validation and Comparison of Two Frailty Indexes: The MOBILIZE Boston Study

Dan K. Kiely, MPH, MA, L. Adrienne Cupples, PhD,† and Lewis A. Lipsitz, MD*‡§*

OBJECTIVES: To validate two established frailty indexes and compare their ability to predict adverse outcomes in a diverse, elderly, community-dwelling sample of men and women.

DESIGN: Prospective observational study.

SETTING: A diverse defined geographic area of Boston.

PARTICIPANTS: Seven hundred sixty-five community-dwelling participants in the Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly Boston Study.

MEASUREMENTS: Two published frailty indexes, recurrent falls, disability, overnight hospitalization, emergency department (ED) visits, chronic medical conditions, self-reported health, physical function, cognitive ability (including executive function), and depression. One index was developed from the Study of Osteoporotic Fractures (SOF) and the other from the Cardiovascular Health Study (CHS).

RESULTS: The SOF frailty index classified 77.1% as robust, 18.7% as prefrail, and 4.2% as frail. The CHS frailty index classified 51.2% as robust, 38.8% as prefrail, and 10.0% as frail. Both frailty indexes (SOF; CHS) were similar in their ability to predict key geriatric outcomes such as recurrent falls (hazard ratio $(HR)_{frail} = 2.2$, 95% confidence interval (CI) = 1.2–4.0; $HR_{frail} = 1.9$, 95% CI = 1.2–3.1), overnight hospitalization (odds ratio $(OR)_{frail} = 3.5$, 95% CI = 1.5–8.0; $OR_{frail} = 4.4$, 95% CI = 2.4–8.2), ED visits ($OR_{frail} = 3.5$, 95% CI = 1.4,8.8; $OR_{frail} = 3.1$, 95% CI = 1.6–5.9), and disability ($OR_{frail} = 5.4$, 95% CI = 2.3–12.3; $OR_{frail} = 7.7$, 95% CI = 4.0–14.7), as well as chronic medical conditions, physical function, cognitive ability, and depression.

CONCLUSION: Two established frailty indexes were validated using an independent elderly sample of diverse men and women; both indexes were good at distinguishing geriatric conditions and predicting recurrent falls, overnight hospitalization, and ED visits according to level of frailty. Although both indexes are good measures of frailty, the simpler SOF index may be easier and more practical in a clinical setting. **J Am Geriatr Soc 57:1532–1539, 2009.**

Key words: frailty; community-dwelling; MOBILIZE Boston study; Study of Osteoporotic Fractures; Cardiovascular Health Study

Frailty has become a particularly important geriatric topic since a 1990 American Medical Association report emphasized the growing population of vulnerable older adults.[1] This vulnerable population, referred to as frail, has been identified as older adults with an intrinsic vulnerability to stressors and high risk for decline and adverse health-related characteristics such as disability and comorbidity. For reasons related to difficulties distinguishing these entities, and because many factors have been reported to be associated with frailty in older adults, there is no single consensus definition of frailty despite numerous definitions proposed by researchers. There is general agreement in the literature that frailty is a biological syndrome of low reserve and resistance to stressors resulting from cumulative declines across multiple physiological systems that cause vulnerability to adverse outcomes.[2]

In an attempt to solidify the concept of frailty and operationalize its definition, Fried et al.[2] proposed a phenotype of frailty involving at least three of the following five components: unintentional weight loss, self-reported low energy level, weak grip strength, slow walking speed, and a low level of physical energy. Using a frailty index based on this phenotype, Fried et al. and other researchers[2–8] have reported its association with falls, hospitalization, disability, and death.

This index has been useful in identifying frail older adults, although its use is impractical in the clinical setting. Assessing strength, walking speed, and physical activity (3 components of this index) not only depends on sex and

From the *Institute for Aging Research, Hebrew SeniorLife, Boston, Massachusetts; †Department of Biostatistics, Boston University, Boston, Massachusetts; ‡Harvard Medical School, Boston, Massachusetts; and §Gerontology Division, Beth Israel Deaconess Medical Center, Boston, Massachusetts.

Address correspondence to Dan K. Kiely, Institute for Aging Research, Hebrew SeniorLife, 1200 Centre St, Boston, MA 02131.
E-mail: kiely@hrca.harvard.edu

DOI: 10.1111/j.1532-5415.2009.02394.x

0002-8614/09/$15.00

body mass, but also requires knowledge of the underlying distribution of the measure in a given population. Moreover, assessing some of these components may not be feasible. For example, physical activity assessments and timed walks are often impractical to evaluate in a clinic because of scheduling and space constraints.[6]

A simpler frailty index has recently been proposed, requiring at least two of three components (weight loss, the inability to rise from a chair 5 times without the use of arms, and self-reported reduced energy level).[6] This index might be more suitable for assessing frailty in a busy clinical practice setting. Based on a prospective cohort study (N = 6,701) designed to examine osteoporosis and fractures in older women,[6] the index was compared with the Fried et al. frailty index on outcomes such as falls, recurrent falls, disability, fractures, and death. Both indexes were strongly associated with these outcomes, and their effect measures (hazard ratios (HRs) and odds ratios (ORs)) were similar. The conclusion was that the simpler index had similar predictive properties to the Fried et al. frailty index, provided a useful definition of frailty, and could be used to identify older women at risk of adverse health outcomes in clinical practice setting. However, one important shortcoming of this study was its limited generalizability, because it did not include men or African-American women.

The objective of this study was to validate and compare these two indexes using an independent diverse sample of men and women (including African Americans). The ability of the indexes to predict recurrent falls, overnight hospitalizations, and emergency department (ED) visits and their association with disability, chronic medical conditions, self-reported health, physical function, cognitive function (including executive function), and depression were examined. To the authors' knowledge, this is the first validation and comparison of these two indexes using an independent data source that includes men and African-American women. It is also the first to examine executive function measures across frailty levels.

METHODS

Study Sample

Subjects were participants in the Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly (MOBILIZE) Boston Study (MBS). The MBS is a prospective observational study designed to examine novel risk factors for falls in a large, diverse population of older individuals in the greater Boston area. The recruitment strategy targeted older persons living within a 5-mile radius of the Institute for Aging Research by using probability sampling from town lists and census information.

Eligibility criteria included aged 70 and older, ability to speak and understand English, ability to walk across a room, visual ability to read written material, and the expectation that the participant would be living in the area for at least 3 years. Companions or spouses who were aged 65 and older living with a participant also were allowed to join the study because it was recognized early on that recruitment of one spouse or companion without the other would limit participation. Study participation was limited to English speakers, because it was not feasible to translate the study instruments and conduct the interviews in the many languages that are spoken within Boston's minority communities.

Once recruited through door-to-door visits, research staff contacted potential participants over the telephone to confirm eligibility and schedule the two-part baseline data collection, which included an extensive 3-hour in-home interview, followed within 4 weeks by a 3-hour in-clinic examination. During the home visit, participants were given a set of monthly falls calendar postcards designed to record the number of falls and instructed as to how to complete and mail them to the Institute for Aging Research at the end of each month during the 18-month follow-up.

The baseline MBS assessments of 765 participants were included in these analyses. The sex and racial distribution of these subjects matches that of the greater Boston metropolitan area population. Details of the study design have previously been published.[9] The results reported in this study used baseline measures and recurrent falls, overnight hospitalization, and ED visit measures during follow-up (average follow-up 10.4 ± 8.2 months; maximum 32.2 months). The institutional review boards of Hebrew SeniorLife approved the MBS, as well as this specific study.

Frailty Index Definitions (Cardiovascular Health Study and Study of Osteoporotic Fractures)

Cardiovascular Health Study Frailty Index

The Fried et al.[2] frailty index was originally based on data from the Cardiovascular Health Study (CHS). An adaptation of this frailty index was constructed for the current study using the MBS data. The five components of the CHS frailty index were unintentional weight loss, weakness, poor endurance or exhaustion (low energy level), slowness (slow gait), and low physical activity. Using the MBS data, self-reported unintentional weight loss was defined using the MBS question "In the last year, have you lost more than 10 pounds unintentionally, that is, not due to dieting or exercise?"). Weakness was defined according to the sit-stand test time, part of the Short Physical Performance Battery (SPPB).[10] Time required to perform five repetitions of sit to stand was measured and used as a proxy for leg strength. The cohort was stratified according to sex and then according to body mass index (BMI) (in quartiles, four strata for each sex) to adjust for the effects of sex and BMI on leg strength. From each stratum, the highest quintile (20%) of sit-to-stand times (including participants who could not perform the task) was chosen to represent weakness. Low energy level was determined according to the Center for Epidemiologic Studies Depression Scale, Hopkins Revision (CESD-R) question,[11] "Over the past week or so, did you feel like you could not get going?" Those who reported symptoms occurring on 3 days or more in the previous week were considered as demonstrating low energy level. Slow gait was defined from the timed 4-m walk. Two trials were performed, and the fastest time was used. The time scores were stratified according to sex and then according to height (2 strata per sex). Participants who used ambulatory assistive devices were included. In each stratum, those in the slowest quintile were considered to have slow gait. Daily activity was determined using the Physical Activity Scale for the Elderly (PASE).[12] The PASE score is a weighted sum of hours spent doing activities of various vigor. The PASE score was stratified according to sex, with participants scoring in the bottom quintile considered to exhibit low daily activity. Consistent with the original CHS frailty index, frailty status was defined as robust (previously referred to as "not frail"; 0 components), prefrail (previously referred to as "intermediate"; 1–2 components), and frail (3–5 components).

Study of Osteoporotic Fractures Frailty Index

This frailty index[6] was originally derived using data from the Study of Osteoporotic Fractures (SOF) and was constructed from the MBS data using three components: unintentional weight loss (MBS question "In the last year, have you lost more than 10 pounds unintentionally, that is, not due to dieting or exercise?"), inability to rise from a chair five times without the use of arms, and low energy level. Low energy level was determined by interviewing participants using a question on the CESD-R:[11] "Over the past week or so, did you feel like you could not get going?" Those who reported that this feeling had occurred 3 days or more in the previous week were considered as demonstrating low energy level. Consistent with the original SOF frailty index, frailty status was defined as robust (0 components), prefrail (previously referred to as "intermediate"; 1 component), and frail ($\geq$2 components).

Outcomes

Recurrent Fallers

Fall status was determined from the falls calendars, on which participants recorded falls each day that they occurred during a given month throughout the follow-up. On any given month, approximately one-third of the participants had to be contacted over the telephone to return the completed calendars. This included reminding participants to mail the calendars by the 15th of each month and asking questions related to filling in missing information on the previously received calendar. Fewer than 1% of calendars were missing each month. Participants were considered to be recurrent fallers if they recorded two or more falls during follow-up. The date of each fall was ascertained from the falls calendar. One-time fallers were treated as nonfallers.

Hospitalization and ED Visits

Overnight hospitalization was determined from a question included in the falls calendar that asked whether the participant had been hospitalized overnight during a given month throughout the follow-up. ED visits were determined from a question included in the falls calendar that asked whether the participant had visited an ED during a given month throughout the follow-up.

Short Physical Performance Battery

The SPPB, which includes measures of standing balance, 4-m usual-paced walking speed, and ability and time to rise from a chair five times, was used to measure lower extremity mobility performance.[10] The validity of this scale has been demonstrated by showing a gradient of risk for admission to a nursing home and mortality along the full range of the scale, from 0 to 12.[13,14]

Activities of Daily Living, Instrumental Activities of Daily Living, and Disability

The activity of daily living (ADL) scale included bathing, dressing, transferring, using the toilet, and eating.[15] The instrumental activity of daily living (IADL) scale included shopping, preparing meals, and housework.[16] Response options for the ADL and IADL items included asking individuals to identify their inability or level of difficulty (none, a little, some, or a lot) in performing each ADL and IADL activity. Each scale was classified into three levels: no difficulty, little or some difficulty, and a lot of difficulty or inability to do one or more activities. Because MBS participants are community dwelling, IADL measures were used to define disability. IADL disability was defined as a lot of difficulty or inability to do one or more IADLs.

Chronic Medical Conditions

A number of chronic medical conditions (yes = 1 or no = 0) were summed into a scale that included heart disease, heart attack, myocardial infarction, angina pectoris, or chest pain; congestive or chronic heart failure; high blood pressure; diabetes mellitus cancer other than skin cancer; osteoarthritis; asthma, emphysema, or chronic bronchitis; stroke; Parkinson's disease; and Alzheimer's disease or dementia. This variable was categorized as 0, 1, 2, 3, 4, and 5 or more, because having more than five comorbidities was rare.

Self-Rated Health

The participants were asked, "In general, would you say your health is excellent (1), very good (2), good (3), fair (4), or poor (5). Lower scores indicate better self-rated health.

Cognitive Measures

Verbal memory functioning was assessed using the Hopkins Verbal Learning Test—Revised (HVLT-R). The HVLT-R is a 12-item wordlist learning test that has been identified as an ideal memory measure for elderly patients and those suspected of having dementia.[17] Higher scores are better. Reliability and validity of the HVLT-R have been shown in older adults and persons with frontal lesions.[18,19] The Mini-Mental State Examination (MMSE),[20] a valid and reliable brief examination of general cognitive function, assesses memory, concentration, attention, and language, yielding a maximum (best) score of 30. Verbal fluency was assessed using phonemic (word-list generation) and semantic (animal) fluency tasks.[21,22] The Trail Making Test Parts A and B (Trails A and B) requires the individual to connect encircled items in sequential order in a timed test. This test is a measure of executive function, is frequently used in the clinical setting, and has been shown to be sensitive to the presence of frontal lobe pathology and cerebrovascular risk.[23] Higher values (seconds) indicate that it took longer to complete the test. The Clock-in-a-Box Test,[24] a modification of the commonly used Clock Drawing test,[25,26] was designed as a cognitive screening measure for use in the medical setting and has increasingly been used as a measure of executive function.[27] Higher scores represent better performance.

Depression

Depression symptomatology was measured using a modification of the 20-item CESD scale.[11] The instrument has been shown to be valid, reliable, and sensitive to change in older populations.[28,29] The CESD-R was used in the MBS. Depressive syndrome burden scores were calculated using item response theory,[30,31] and the metric was set relative to the mean and variance of the MBS sample aged 70 to 74 at baseline interview using a mean of 50 and a standard deviation of 10. The items that constitute the CESD-R had high internal consistency (coefficient $\alpha = 0.86$).

Covariates

A number of variables were included in adjusted analyses: age, sex, race, education, income, diabetes mellitus, stroke, hypertension, and hyperlipidemia. Race was defined as Caucasian versus non-Caucasian; 80% were Caucasian, 16% were African American, and the remaining categories presented 2% or less. Education was defined as less than high school, high school graduate, and college graduate. Income was measured in ordinal categories ranging from less than $5,000 to $45,000 or more by $5,000 increments. Education and income were used as proxies for socioeconomic status. Diabetes mellitus and stroke were self-reported. Hypertension was defined as systolic blood pressure of 140 mmHg or higher, diastolic blood pressure of 90 mmHg or higher, told by participant's physician that he or she had high blood pressure or hypertension, or receiving any hypertensive medication. Hyperlipidemia was considered positive if cholesterol was 200 mg/dL or higher, low-density lipoprotein cholesterol was 130 mg/dL or higher, or the participant was taking a cholesterol-lowering drug. Except for stroke and hyperlipidemia, covariates were chosen because they were reported to be associated with frailty in the CHS or SOF study.

Statistical Analysis

Means ± standard deviations and frequencies (percentages) were calculated to characterize the study sample. A weighted kappa was calculated to estimate the association between the frailty categories of the two indexes. Analysis of variance was used to compare mean values across frailty status groups, and multiple comparison tests were performed when appropriate. Cox proportional hazards analyses (unadjusted and adjusted) were performed, and HRs and corresponding 95% confidence intervals (CIs) were calculated to estimate the association between frailty status and recurrent falls. Recurrent fallers were coded positive if they fell two or more times during follow-up and 0 otherwise. One-time fallers were coded as nonfallers. The analysis modeled time to first fall for recurrent fallers. Indicator variables (dummy variables) were created using robust as the referent group.

Logistic regression analyses (unadjusted and adjusted) were performed, and ORs and corresponding 95% CIs were calculated to estimate the association between frailty status and overnight hospitalization, ED visits, and disability. Indicator variables (dummy variables) were created using robust as the referent group. The adjusted models included age, sex, diabetes mellitus, stroke, hypertension and hyperlipidemia. An alpha level of 0.05 was used in all analyses to determine statistical significance and guide inference. SAS, Version 9.1 for Windows (SAS Institute, Inc., Cary, NC) was used for statistical analyses.

RESULTS

The study sample of 765 participants was characterized in Table 1. The average age was 78, 63.9% of participants were women, 78.0% were white, and 21.2% were disabled. Thirty percent of participants fell at least twice (recurrent faller), 34.8% were hospitalized overnight, and 45.0% visited an ED during follow-up. Table 1 also lists information on chronic medical conditions, self-reported health, physical and cognitive ability, depression, and the distribution of frailty categories for each index. The weighted kappa for the agreement between the two indexes was 0.51, which represents a moderate association.

Table 2 presents the means of participant characteristics, clinical conditions, functional and cognitive ability, and depression across frailty categories for each index. The frailty group characteristics for each index were similar. Older age, greater number of chronic medical conditions, and worse self-reported health were associated with greater frailty. Greater impairments in physical (SPPB, ADL, IADL) and cognitive function (MMSE, HVLT-R, phonemic (word-list generation) and semantic (animal) fluency tasks, Trails A, Trails B, Clock-in-a-Box) were associated with greater frailty. Finally, frailty was associated with higher depression scores. There was a statistically significant ($P < .05$) difference between the mean values of each variable between the frailty groups for both indexes.

Table 3 presents the unadjusted and adjusted HRs and corresponding 95% CIs for the association between frailty groups and recurrent falls for each index. In the adjusted analysis, frail participants were 2.19 (95% CI = 1.19–4.03; SOF) and 1.90 (95% CI = 1.17–3.10; CHS) times as likely to experience a recurrent fall as robust participants. Prefrail participants were 1.62 (95% CI = 1.14–2.32; SOF) and 1.10 (95% CI = 0.80–1.50; CHS) times as likely to experience a recurrent fall as robust participants.

Table 4 presents the unadjusted and adjusted ORs and corresponding 95% CIs for the association between frailty indexes and ED visits and overnight hospitalization. In adjusted analyses, frail participants were 3.49 (95% CI = 1.53–7.98; SOF) and 4.45 (95% CI = 2.42–8.18; CHS) times as likely to experience an overnight hospitalization as robust participants. Prefrail participants were 2.64 (95% CI = 1.74–4.01; SOF) and 1.97 (95% CI = 1.37–2.84; CHS) times as likely to experience an overnight hospitalization as robust participants. In adjusted analyses, frail participants were 3.54 (95% CI = 1.43–8.79; SOF) and 3.10 (95% CI = 1.64–5.86; CHS) times as likely to experience an ED visit as robust participants. Prefrail participants were 2.19 (95% CI = 1.43–3.33; SOF) and 1.34 (95% CI = 0.95–1.89; CHS) times as likely to experience an ED visit as robust participants.

Table 5 presents the unadjusted and adjusted ORs and corresponding 95% CIs for the association between frailty groups and prevalent IADL disability for each index. In adjusted analyses, frail participants were 5.38 (95% CI = 2.34–12.35; SOF) and 7.68 (95% CI = 4.01–14.74; CHS) times as likely to be disabled as robust participants.

Table 1. Descriptive Information on Maintenance of Balance, Independent Living, Intellect, and Zest in the Elderly Boston Study Participants (N = 765)

Characteristic	Value
Age, mean ± SD	78.1 ± 5.4
Female, n (%)	489 (63.9)
Race, n (%)	
Caucasian	596 (78.0)
African American	121 (15.8)
Asian	10 (1.3)
American Indian	4 (0.5)
Multiracial	17 (2.2)
Other	16 (2.1)
Education, n (%)	
<High school	85 (11.1)
High school graduate	178 (23.3)
College graduate	501 (65.6)
Annual income, $, n (%)	
<5,000	18 (2.6)
5,000–9,999	70 (10.1)
10,000–14,999	82 (11.8)
15,000–24,999	114 (16.4)
25,000–34,999	75 (10.8)
35,000–44,999	77 (11.1)
≥45,000	258 (37.2)
Height, m, mean ± SD	1.6 ± 0.1
Weight, kg, mean ± SD	73.1 ± 15.4
Body mass index, kg/m^2, mean ± SD	27.3 ± 5.2
Diabetes mellitus, n (%)	141 (18.7)
Stroke, n (%)	76 (10.0)
Hypertension, n (%)	598 (79.1)
Hyperlipidemia, n (%)	359 (46.9)
Number of chronic medical conditions, n (%)	
0	125 (16.3)
1	278 (36.3)
2	212 (27.7)
3	109 (14.2)
4	28 (3.7)
≥5	13 (1.7)
Self-reported health, n (%)	
Excellent	126 (16.5)
Very good	251 (32.8)
Good	274 (35.8)
Fair	100 (13.1)
Poor	14 (1.8)
Recurrent fallers, n (%)	233 (30.5)
Disability, n (%)	162 (21.2)
Overnight hospitalization, n (%)	266 (34.8)
Emergency department visit, n (%)	344 (45.0)
Physical function, mean ± SD	
Short Physical Performance Battery*	9.3 ± 2.5
Activity of daily living scale†	0.3 ± 0.6
Instrumental activity of daily living scale†	0.6 ± 0.8
Cognitive function, mean ± SD	
Mini-Mental State Examination*	27.1 ± 2.7

(*Continued*)

Table 1. (Contd.)

Characteristic	Value
Hopkins Verbal Learning Test—Revised*	0.02 ± 0.76
Phonemic fluency task (word-list generation), number of words*	36.6 ± 14.5
Semantic fluency task (animal), number of words*	15.8 ± 5.2
Trail Making Test Part A, seconds†	57.3 ± 35.5
Trail Making Test Part B, seconds†	143.6 ± 78.6
Clock-in-a-Box*	6.3 ± 1.5
Center for Epidemiologic Studies Depression Scale—Hopkins Revision, mean ± SD†	50.6 ± 10.0
Items used in frailty indexes, n (%)	
Weight loss (unintentional)	56 (7.4)
Inability to rise from a chair 5 times	87 (11.4)
Low energy level‡	67 (8.8)
Sit-to-stand time, seconds, mean ± SD	12.9 ± 3.8
Gait speed, seconds, mean ± SD	4.6 ± 1.6
Physical Activity Scale for the Elderly, mean ± SD†	107.4 ± 70.9
Study of Osteoporotic Fractures frailty index, n (%)	
Robust	590 (77.1)
Prefrail	143 (18.7)
Frail	32 (4.2)
Cardiovascular Health Study frailty index, n (%)	
Robust	389 (51.2)
Prefrail	295 (38.8)
Frail	76 (10.0)

* Lower score indicates greater impairment.
† Higher score indicates greater impairment.
‡ Low energy level was assessed using the Center for Epidemiologic Studies Depression Scale (CESD) question, "Over the past week or so, did you feel like you could not get going?"
SD = standard deviation.

Prefrail participants were 2.88 (95% CI = 1.81–4.58; SOF) and 2.73 (95% CI = 1.64–4.40; CHS) times as likely to be disabled as robust participants.

Because of concern that the five missing values in the CHS frailty index might have differentially affected the comparison with the SOF frailty index, all SOF frailty index analyses were rerun excluding the five participants who had missing CHS frailty index values. The results of these analyses were nearly identical to the original analyses.

DISCUSSION

The results of this study demonstrate that the CHS and SOF frailty indexes provide useful definitions of frailty, are associated with cognitive and functional deficits, and are good predictors of adverse outcomes. Both indexes were similar in their ability to distinguish the incidence of recurrent falls, overnight hospitalizations, ED visits, and the prevalence of IADL disability, chronic medical conditions, self-reported health, physical function, cognitive ability, and depression based on different frailty levels. The simpler SOF frailty index may be more useful in clinical practice, because its components are easier to define and do not require knowledge of population distributions. The results of this study validate the findings of the SOF and CHS studies in a diverse elderly community-dwelling population that, unlike the

Table 2. Select Variables Stratified According to Frailty Status (Robust, Prefrail, Frail) for Two Frailty Indexes

Variable	Study of Osteoporotic Fractures Frailty Index (n = 765): Robust n = 590 (77.1%)	Prefrail n = 143 (18.7%)	Frail n = 32 (4.2%)	*P*-Value†	CHS Frailty Index (n = 760)*: Robust n = 389 (51.2%)	Prefrail n = 295 (38.8%)	Frail n = 76 (10.0%)	*P*-Value†
Age, mean ± SD	77.6 ± 5.3	79.7 ± 5.4	81.4 ± 5.9	<.001(1,2)	76.8 ± 5.0	79.1 ± 5.6	80.8 ± 5.5	<.001(1,2,3)
Body mass index, kg/m^2, mean ± SD	27.4 ± 5.1	27.1 ± 5.3	27.5 ± 5.6	.82	26.8 ± 4.7	27.5 ± 5.1	28.9 ± 6.9	.005(2,3)
Number of chronic medical conditions, mean ± SD	1.5 ± 1.1	1.8 ± 1.2	2.2 ± 1.6	<.001(1,2,3)	1.3 ± 1.0	1.8 ± 1.1	2.2 ± 1.5	<.001(1,2,3)
Self-reported health, mean ± SD	2.3 ± 0.9	3.0 ± 1.0	3.4 ± 1.0	<.001(1,2,3)	2.2 ± 0.9	2.7 ± 0.9	3.2 ± 1.0	<.001(1,2,3)
Short Physical Performance Battery, mean ± SD‡	10.0 ± 1.8	7.3 ± 2.9	5.0 ± 2.3	<.001(1,2,3)	10.7 ± 1.2	8.5 ± 2.3	5.3 ± 2.3	<.001(1,2,3)
Activity of daily living scale, mean ± SD§	0.18 ± 0.47	0.66 ± 0.75	1.03 ± 0.86	<.001(1,2,3)	0.10 ± 0.36	0.41 ± 0.67	0.95 ± 0.78	<.001(1,2,3)
Instrumental activity of daily living scale, mean ± SD§	0.49 ± 0.74	0.99 ± 0.90	1.28 ± 0.89	<.001(1,2)	0.35 ± 0.65	0.80 ± 0.85	1.26 ± 0.88	<.001(1,2,3)
Mini-Mental State Examination, mean ± SD‡	27.3 ± 2.5	26.3 ± 3.0	26.0 ± 2.9	<.001(1,2)	27.6 ± 2.3	26.8 ± 2.7	25.4 ± 3.2	<.001(1,2,3)
Hopkins Verbal Learning Test, mean ± SD‡	0.07 ± 0.75	−0.13 ± 0.79	−0.16 ± 0.70	.007(1)	0.16 ± 0.72	−0.08 ± 0.75	−0.30 ± 0.83	<.001(1,2,3)
Phonemic fluency task (word-list generation), number of words, mean ± SD‡	37.3 ± 14.5	34.77 ± 14.4	30.7 ± 13.0	.01(2)	38.7 ± 14.2	35.6 ± 14.2	29.3 ± 13.4	<.001(1,2,3)
Semantic fluency task (animal), number of words, mean ± SD‡	16.2 ± 5.1	14.7 ± 5.0	12.8 ± 4.6	<.001(1,2)	16.8 ± 5.1	15.3 ± 5.0	12.6 ± 4.6	<.001(1,2,3)
Trail Making Test Part A, mean ± SD§	53.4 ± 31.1	69.8 ± 44.1	77.1 ± 50.0	<.001(1,2)	48.0 ± 23.8	60.3 ± 31.4	94.4 ± 65.4	<.001(1,2,3)
Trail Making Test Part B, mean ± SD§	135.7 ± 75.8	168.4 ± 81.3	190.8 ± 84.8	<.001(1,2)	128.5 ± 71.2	153.0 ± 79.2	205.5 ± 81.1	<.001(1,2,3)
Clock-in-a-Box, mean ± SD‡	6.4 ± 1.5	6.0 ± 1.6	5.5 ± 1.7	<.001(1,2)	6.5 ± 1.4	6.2 ± 1.4	5.3 ± 2.1	<.001(1,2,3)
Center for Epidemiologic Studies Depression Scale—Hopkins Revision, mean ± SD§	49.0 ± 9.1	55.1 ± 10.9	61.2 ± 10.8	<.001(1,2,3)	48.2 ± 9.1	52.5 ± 9.8	59.9 ± 11.8	<.001(1,2,3)
Female, %	63.7	65.7	59.4	.78	63.2	64.1	64.5	.96
Caucasian, %	78.8	75.5	75.0	.64	80.7	76.9	67.1	.03
Disability, %	14.9	39.2	56.2	<.001	9.5	27.5	55.3	<.001
Overnight hospitalization, %	29.1	51.7	62.5	<.001	24.7	40.3	64.5	<.001
Emergency department visit, %	39.7	60.1	75.0	<.001	36.8	48.1	73.7	<.001

* The Cardiovascular Health Study (CHS) frailty index could not be created for five participants due to missing values.
† *P*-values from overall test of group differences in analyses of variance. Specific group differences ($P<.05$): 1 = robust and prefrail, 2 = robust and frail, 3 = prefrail and frail.
‡ Lower score indicates greater impairment.
§ Higher score indicates greater impairment.
SD = standard deviation.

Table 3. Association Between Frailty (Two Indexes) and Recurrent Falls

	Hazard Ratio (95% Confidence Interval) *P*-Value	
Risk of Recurrent Falls	Study of Osteoporotic Fractures Frailty Index (N = 765)	Cardiovascular Health Study Frailty Index (N = 760)
Unadjusted		
Robust	Referent	
Prefrail	1.43 (1.04–1.95) .03	1.06 (0.80–1.40) .71
Frail	1.55 (0.88–2.72) .13	1.61 (1.09–2.39) .02
Adjusted*		
Robust	Referent	
Prefrail	1.62 (1.14–2.32) .008	1.10 (0.80–1.50) .57
Frail	2.19 (1.19–4.03) .01	1.90 (1.17–3.10) .01

Indicator (dummy) variables represent prefrail and frail groups relative to robust.
* Adjusted for age, sex, race, diabetes mellitus, stroke, hypertension, hyperlipidemia, education, and income.

Table 5. Association Between Frailty (Two Indexes) and Instrumental Activity of Daily Living (IADL) Disability

	Odds Ratio (95% Confidence Interval) *P*-Value	
Frailty Category	Study of Osteoporotic Fractures Frailty Index (n = 765)	Cardiovascular Health Study Frailty Index (n = 760)
IADL disability (a lot of difficulty or inability to perform $\geq$1 IADLs)		
Unadjusted		
Robust	Referent	
Prefrail	3.66 (2.44–5.49) <.001	3.62 (2.37–5.53) <.001
Frail	7.32 (3.51–15.25) <.001	11.75 (6.68–20.67) <.001
Adjusted*		
Robust	Referent	
Prefrail	2.88 (1.81–4.58) <.001	2.73 (1.69–4.40) <.001
Frail	5.38 (2.34–12.35) <.001	7.68 (4.01–14.74) <.001

Indicator (dummy) variables represent prefrail and frail groups relative to robust.
* Adjusted for age, sex, race, diabetes mellitus, stroke, hypertension, hyperlipidemia, education, and income.

previous comparison of SOF and CHS, included men and African-American women.

It is not surprising that the SOF and CHS frailty indexes are similar, because the three items used in the SOF index are similar to three of the five items included in the CHS index. This is particularly true in the modified CHS index, because chair stand time was substituted for a measure of weakness, although the weighted kappa for the agreement between the two indexes was 0.51, which represents a moderate rather than a high association. Furthermore, although the associations between the indexes and various outcomes were similar, in the same direction, and statistically significant, the magnitude of the association varied slightly across indexes. For example, using the robust frailty category as the referent, individuals who were considered to be frail according to the SOF were 2.19 times as likely to be recurrent fallers, compared with 1.90 times as likely for individuals who were considered to be frail according to the CHS. Similar small differences were observed with the outcomes overnight hospitalization (CHS has a higher risk), ED visits (SOF has a higher risk), and IADL-defined disability (CHS has a higher risk).

Table 4. Association Between Frailty (Two Indexes) and Overnight Hospitalization and Emergency Department Visits

	Odds Ratio (95% Confidence Interval) *P*-Value	
Frailty Category	Study of Osteoporotic Fractures Frailty Index (n = 765)	Cardiovascular Health Study Frailty Index (n = 760)
Overnight hospitalizations		
Unadjusted		
Robust	Referent	
Prefrail	2.61 (1.79–3.78) <.001	2.06 (1.49–2.86) <.001
Frail	4.05 (1.94–8.47) <.001	5.54 (3.28–9.35) <.001
Adjusted*		
Robust	Referent	
Prefrail	2.64 (1.74–4.01) <.001	1.97 (1.37–2.84) <.001
Frail	3.49 (1.53–7.98) .003	4.45 (2.42–8.18) <.001
Emergency department visits		
Unadjusted		
Robust	Referent	
Prefrail	2.29 (1.58–3.33) <.001	1.60 (1.17–2.17) .003
Frail	4.56 (2.02–10.33) <.001	4.82 (2.78–8.35) <.001
Adjusted*		
Robust	Referent	
Prefrail	2.19 (1.43–3.33) <.001	1.34 (0.95–1.89) .10
Frail	3.54 (1.43–8.79) .006	3.10 (1.64–5.86) <.001

Indicator (dummy) variables represent prefrail and frail groups relative to robust.
* Adjusted for age, sex, race, diabetes mellitus, stroke, hypertension, hyperlipidemia, education, and income.

Results from Table 2 show that the percentage of Caucasians was not significantly different between categories of the SOF frailty index but was significantly different between categories of the CHS frailty index. This may lend support to a recent study[32] suggesting that standardization of frailty items without consideration of ethnic variations is problematic and may lead to misclassification of frailty categories for non-Caucasians. Moreover, the fact that body mass was not significantly different between categories of the SOF frailty index but was significantly different between categories of the CHS frailty index may support the perspective that higher BMI values in African Americans than Caucasians may lead to overclassification of weakness in African Americans, which could partially account for the higher prevalence of frailty found in the CHS classification of frailty.[32]

The strengths of this study include its prospective ascertainment of recurrent falls, overnight hospitalizations, and ED visits and its inclusion of men and African-American subjects. Furthermore, this is the first study to compare these two established frailty indexes (SOF and CHS) using a third source of data.

One limitation of this study was that it did not have the exact same measures used in the originally reported indexes. Another limitation was that outcomes such as falls, overnight hospitalizations, and ED visits were ascertained according to self-report. Also, as in the SOF and CHS studies, the MBS data were collected for other purposes that are not directly related to the study of frailty. Like the SOF and CHS results, the generalizability of the findings may be limited to community-dwelling elderly people. Finally, the MBS data did not contain an adequate number of deaths to include mortality as an outcome.

In conclusion, two established frailty indexes were validated and compared using an independent data source that included men and African-American women, and it was shown that both indexes were good at distinguishing relevant geriatric conditions, functional and cognitive impairments, and predicting adverse outcomes and acute care service use according to level of frailty. Both indexes are good measures of frailty and are able to distinguish differences according to level of frailty. The simpler SOF index may be easier and more practical to use in a clinical setting than the CHS index.

ACKNOWLEDGMENTS

We thank the participants, staff, and investigators of the MBS who made this study possible.

Conflict of Interest: The editor in chief has reviewed the conflict of interest checklist provided by the authors and has determined that the authors have no financial or any other kind of personal conflicts with this paper.

This work was supported by a National Institute on Aging Program Project P01AG004390 and Research Grant R37AG025037.

Author Contributions: DKK and LAL: study concept and design, analysis and interpretation of data, and preparation of manuscript. LAC: analysis and interpretation of data, and preparation of manuscript.

Sponsor's Role: None.

REFERENCES

1. Fried LP, Ferrucci L, Darer J et al. Untangling the concepts of disability, frailty, and comorbidity: Implications for improved targeting and care. J Gerontol A Biol Sci Med Sci 2004;59A:255–263.
2. Fried LP, Tangen CM, Walston J et al. Frailty in older adults: Evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56A:M146–M156.
3. Bandeen-Roche K, Xue QL, Ferrucci L et al. Phenotype of frailty: Characterization in the Women's Health and Aging Studies. J Gerontol A Biol Sci Med Sci 2006;61A:262–266.
4. Boyd CM, Xue QL, Simpson CF et al. Frailty, hospitalization, and progression of disability in a cohort of disabled older women. Am J Med 2005;118:1225–1231.
5. Cawthon PM, Marshall LM, Michael Y et al. Frailty in older men: Prevalence, progression, and relationship with mortality. J Am Geriatr Soc 2007;55:1216–1223.
6. Ensrud KE, Ewing SK, Taylor BC et al. Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women. Arch Intern Med 2008;168:382–389.
7. Ensrud KE, Ewing SK, Taylor BC et al. Frailty and risk of falls, fracture, and mortality in older women: The study of osteoporotic fractures. J Gerontol A Biol Sci Med Sci 2007;62A:744–751.
8. Woods NF, LaCroix AZ, Gray SL et al. Frailty: Emergence and consequences in women aged 65 and older in the Women's Health Initiative Observational Study. J Am Geriatr Soc 2005;53:1321–1330.
9. Leveille SG, Kiel DP, Jones RN et al. The MOBILIZE Boston study: Design and methods of a prospective cohort study of novel risk factors for falls in an older population. BMC Geriatr 2008;8:1–42.
10. 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 1994;49:M85–M94.
11. Radloff L. The CES-D scale: A self report depression scale for research in the general population. Appl Psychol Meas 1977;1:385–401.
12. Washburn RA, Smith KW, Jette AM et al. The Physical Activity Scale for the Elderly (PASE): Development and evaluation. J Clin Epidemiol 1993;46:153–162.
13. Guralnik JM, Ferrucci L, Pieper CF et al. Lower extremity function and subsequent disability: Consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000;55A:M221–M231.
14. Guralnik JM, Ferrucci L, Simonsick EM et al. 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.
15. Katz S, Ford AB, Moskowitz RW et al. Studies of illness in the aged. The Index of ADL: A standardized measure of biological and psychosocial function. JAMA 1963;185:914–919.
16. Lawton MP, Brody EM. Assessment of older people: Self-maintaining and instrumental activities of daily living. Gerontologist 1969;9:179–186.
17. Benedict RHB, Schretlen D, Groninger L et al. The Hopkins Verbal Learning Test Revised: Normative data and analysis of inter-form and test-retest reliability. Clin Neuropsychol 1998;12:43–45.
18. Miceli G, Caltagirone C, Gainotti G et al. Neuropsychological correlates of localized cerebral lesions in non-aphasic brain-damaged patients. J Clin Neuropsychol 1981;3:53–63.
19. Perret E. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia 1974;12:323–330.
20. 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.
21. Benton AL. Development of a multilingual aphasia battery. Progress and problems. J Neurol Sci 1969;9:39–48.
22. Benton AL, Hamsher K. Multilingual Aphasia Examination. Iowa City: University of Iowa, 1976.
23. Pugh KG, Kiely DK, Milberg WP et al. Selective impairment of frontal-executive cognitive function in African Americans with cardiovascular risk factors. J Am Geriatr Soc 2003;51:1439–1444.
24. Milberg WP, Hebben NA, Kaplan E. The Boston Process Approach to Neuropsychological Assessment. New York: Oxford University Press, 1986.
25. Royall DR, Cordes JA, Polk M. CLOX: An executive clock drawing task. J Neurol Neurosurg Psychiatry 1998;64:588–594.
26. Royall DR, Mulroy AR, Chiodo LK et al. Clock drawing is sensitive to executive control: A comparison of six methods. J Gerontol B Psychol Sci Soc Sci 1999;54B:P328–P333.
27. Shulman KI. Clock-drawing: Is it the ideal cognitive screening test? Int J Geriatr Psychiatry 2000;15:548–561.
28. Berkman LF, Berkman CS, Kasl S et al. Depressive symptoms in relation to physical health and functioning in the elderly. Am J Epidemiol 1986;124:372–388.
29. Yesavage JA, Brink TL, Rose TL et al. Development and validation of a geriatric depression screening scale: A preliminary report. J Psychiatr Res 1982;17:37–49.
30. Jones RN, Fonda SJ. Use of an IRT-based latent variable model to link different forms of the CES-D from the Health and Retirement Study. Soc Psychiatry Psychiatr Epidemiol 2004;39:828–835.
31. Lord F, Novick M. Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley, 1968.
32. Espinoza SE, Hazuda HP. Frailty in older Mexican-American and European-American adults: Is there an ethnic disparity? J Am Geriatr Soc 2008;56:1744–1749.