GeroScience
https://doi.org/10.1007/s11357-026-02183-1

ORIGINAL ARTICLE

Revisiting the social frailty index in aging adults: relevance and implications for an emerging construct

Charles W. Collinge · **Alessandro Bartolomucci** · **Monica Luciana**

Received: 22 January 2026 / Accepted: 20 February 2026
© The Author(s) 2026

Abstract Frailty is an established benchmark of aging-related decline, yet most measures of frailty focus on physical decline. Increasing evidence that social environments influence trajectories of aging has led to growing interest in social frailty. Recently, a 10-item Social Frailty Index (SFI-10) was developed using data from the Health and Retirement Study (HRS) that used 8 social items plus chronological age and sex to predict mortality. This study evaluates the validity of the SFI-10, providing considerations for social frailty measurement. Utilizing HRS cohorts ($n=8264$, ages 66 to 101) and cohorts from the Midlife in the United States (MIDUS) study ($n=5619$, ages 40 to 75), the SFI-10 was evaluated alongside an index including only its social-environmental factors (SFI-8). Logistic regressions were used to predict mortality from the SFI-10, SFI-8, and only age and sex. Within MIDUS, the SFI-10, SFI-8, and age were further evaluated against external criteria relevant for social frailty. Findings indicate that chronological age is the most important predictor of mortality risk across both samples. Social-environmental factors predict mortality only marginally in older adults but more strongly in midlife. When associated with external criteria relevant for social frailty, the SFI-8 exhibited stronger criterion validity than the SFI-10. These findings suggest that social frailty may be best reflected by measures focusing exclusively on social behaviors and resources, excluding demographic characteristics. Our work underscores the need for well-designed measures of social frailty.

Keywords Frailty · Geroscience · Well-being · Loneliness · Social determinants of health

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11357-026-02183-1.

C. W. Collinge (✉) · M. Luciana
Department of Psychology, University of Minnesota, Minneapolis, MN, USA
e-mail: collinge@umn.edu

A. Bartolomucci
Department of Integrative Biology & Physiology, University of Minnesota, Minneapolis, MN, USA

Introduction

Frailty is a useful benchmark for scientists, clinicians, and policymakers alike. Defined as a state of increased vulnerability to adverse outcomes resulting from the cumulative biological decline that occurs with aging, frailty reflects variations in the pace of aging between individuals [1–5]. Moreover, frailty is a well-established predictor of disability, falls, hospitalizations, and mortality [6–10]. Among the many extant methods to quantify frailty, two distinct approaches have been most extensively validated: the *phenotype* and *deficit accumulation* approaches [6, 11–13]. Common to these two approaches is an emphasis on items that are physical in nature,

Published online: 24 March 2026

yielding measures of "physical frailty" [6, 14, 15]. However, hallmarks of biological aging do not occur in a vacuum; they interact with social contexts [16–18]. Robust associations have been established between social factors and longevity across time, place, and species, highlighting the importance of integrating geroscience with social science to produce a holistic account of the aging process and its determinants [19–23]. This necessitates innovative methods to define social dimensions of frailty.

Recognition that frailty is multidimensional has existed for over three decades, yet "social frailty" remains relatively understudied and lacks consensus regarding its measurement [24, 25]. Nonetheless, an influential framework for social frailty in older adults was advanced by Bunt and colleagues [26]. Hypothesizing that social frailty represents an inability to fulfill one's basic social needs, thereby compromising well-being, Bunt et al. [26] used the Social Production Functions (SPF [27]) theory as a heuristic to define social frailty as "a continuum of being at risk of losing, or having lost, social and general resources, activities, or abilities that are important for fulfilling one or more basic social needs during the life span," and detail a theory-based hierarchical process through which social frailty occurs. This framework has not yet been comprehensively empirically evaluated, however, and is not explicitly concordant with geroscience and frameworks of frailty. Specifically, how social frailty relates to an increased risk for adverse age-related outcomes such as mortality, or relates to the aging process itself, is not defined. To provide support for the scientific and clinical utility of social frailty as a construct, an empirically derived measure of social frailty must be linked to (1) outcomes proximal to this framework, such as social adjustment and well-being, (2) decline resultant of the aging process, and (3) distal adverse age-related outcomes, such as mortality.

Recently, Shah and colleagues [28] addressed some of these limitations, developing a 10-item Social Frailty Index (SFI-10) based on Bunt's framework and keyed on items that predicted mortality outcomes. Using data from the Health and Retirement Study (HRS), Shah et al. identified items within the HRS's Psychosocial and Lifestyle Questionnaire (PLQ) that fit within Bunt's framework. Least absolute shrinkage and selection operator (LASSO) regression on 4-year mortality outcomes was then used in a cohort of older adults to identify a subset of the candidate items that best predicted mortality. This process produced the SFI-10, which consists of items spanning general resources and life history, social resources, social activities, fulfillment of basic social needs, and two demographic features, chronological age and sex. In a separate evaluation cohort of older adults, the SFI-10 exhibited fair discrimination for mortality, activities of daily living, and nursing home stay outcomes, all assessed over 4 years.

Nevertheless, SFI-10 items were ultimately selected based on their utility in predicting mortality outcomes, and the index is functionally a logistic regression equation that directly quantifies mortality risk. This raises the question of what construct their index truly measures: is the SFI-10 a measure of social frailty as it is labeled, or is it best construed as a mortality risk prediction index? Additionally, the SFI-10 includes chronological age and sex. Robust relationships have been established between age, sex, and longevity [29–32]. Thus, it is not clear whether the social components of the SFI-10 predict mortality, beyond risks associated with age and sex. Lastly, like other existing social frailty measurements [33–38], the SFI-10 was developed and evaluated in adults aged > 65 years. Although high levels of social frailty may not manifest until later in the life course, the translational and clinical utility of a social frailty measurement would be enhanced if it could meaningfully detect age-related variations in social functioning at midlife to inform intervention strategies during periods of relative health [39–41].

The present study revisits the SFI-10, considering how it relates to the broader social frailty construct, as well as *how* it predicts mortality outcomes from midlife to older age. Accordingly, this study had three aims. First, we re-calculated the SFI-10 in the HRS cohorts used by Shah et al. [28] and derived an 8-item index consisting of only its social components (SFI-8) to investigate the extent to which they alone contribute to mortality prediction. Second, using cohorts and data from the Midlife in the United States (MIDUS) study, we derived the SFI-10 and SFI-8 using analogous items to examine the generalizability of our findings to mortality prediction in adults at midlife through the transition to older age. Third, we identified external criterion outcomes in MIDUS that fit within Bunt's framework of social frailty and examined the extent to which each index (SFI-10 and

SFI-8) relates to these criteria to evaluate the validity of the SFI-10 and SFI-8 as measures of social frailty, rather than as measures of mortality risk.

Methods

Sample

Four cohorts from two large longitudinal studies on aging adults in the USA comprise the analytic samples for the present study. Two cohorts were drawn from the HRS, mirroring the approach used to initially develop and validate the SFI-10 [28], and two were drawn from the MIDUS study to assess the generalizability and validity of the SFI-10.

The HRS is a longitudinal panel study that has surveyed a nationally representative sample of more than 46,000 individuals over age 50 since its inception in 1992 [42–44]. HRS interviews are conducted biennially in-person, via phone, or internet surveys and consist of multiple measures, one of which is the Psychosocial and Lifestyle Questionnaire [45]. The PLQ is administered to a random half of the overall HRS sample during every biennial interview wave, assuring no overlap between individuals in subsequent waves. Mirroring the methodology employed for developing the SFI-10 [28], we selected HRS cohorts consisting of individuals aged > 65 years who completed the PLQ in 2010 or 2012, yielding: (1) a 2010 cohort, used for model development, and (2) a 2012 cohort, used for model evaluation. A total of 4308 individuals were included in our HRS 2010 cohort, and 3956 individuals were included in our HRS 2012 cohort. These values differ from those previously reported by $n=6$ ($n=4302$ [28]) and $n=8$ individuals ($n=3948$ [28]), respectively despite using their published inclusion criteria. Both HRS cohorts were similar to one another with respect to descriptive statistics and health/functional status (Table S1). By the mortality outcome assessment date, 22% (935/4308) of the 2010 cohort and 23% (906/3956) of the 2012 cohort were deceased. Despite using the same criteria for case selection, these values differ slightly from those reported by Shah et al., undershooting their 2010 mortality rate by 25 deaths (960 deaths [28]) and overshooting their 2012 mortality rate by 24 deaths (882 deaths [28]). We could not reconcile these differences after corresponding with the authors. The present analyses were carried out using our calculated mortality rates.

MIDUS is a longitudinal study initiated in 1995–1996 (MIDUS 1), surveying a nationally representative sample of 7108 individuals ages 25–74 [40, 46]. A longitudinal follow-up assessment of the MIDUS 1 sample was conducted in 2004–2006 (MIDUS 2; ages 28–84). Between 2011 and 2014, a separate sample of 4085 individuals ages 25–75 was enrolled to “refresh” the original sample (MIDUS Refresher 1). Survey project data via phone interview and self-administered questionnaire from the MIDUS 2 and Refresher 1 samples were used for the present analyses. Considering our aim to evaluate the criterion validity of the SFI-10 and its generalizability to a younger sample consisting of adults at midlife and the transition to old age, our MIDUS 2 and Refresher 1 cohorts consisted of participants aged 40–70 years with complete data across all outcome criterion variables described below, yielding: (1) a MIDUS 2 (M2) cohort, used for model development, and (2) a MIDUS Refresher 1 (MR1) cohort, used for model evaluation. 3542 individuals were included in the MIDUS M2 cohort, and 2077 individuals were included in the MR1 cohort (Table S1). By the mortality outcome assessment date, 4% (144/3542) of the M2 cohort and 4% (88/2077) of the MR1 cohort were deceased. The MIDUS cohorts were similar to one another in composition based on descriptive statistics and health/functional status (Table S1). The MIDUS cohorts reported higher rates of fair or poor health but lower rates of age-related non-communicable diseases compared to the HRS cohorts. These differences between HRS and MIDUS are expected given each study’s age distributions.

Mortality outcomes

Shah et al. included all-cause mortality as the primary outcome for SFI-10 model development and evaluation [28]. Correspondence with the authors (unpublished) revealed that all outcomes were assessed using fixed timepoints for each cohort at the end of the fourth year after the last interview in each wave, which resulted in variable outcome assessment timeframes spanning from 4.17 to 5.75 years. We replicated this approach for model development and evaluation. Mortality was determined using December 31, 2015, as the outcome assessment date for the

HRS 2010 cohort and December 31, 2017, for the HRS 2012 cohort. Individuals whose deaths occurred before the outcome assessment date were assigned a positive mortality status. Of note, Shah et al. used both the National Death Index (NDI) and HRS data to determine mortality status at the time of their study. However, since the HRS references the NDI to complete its mortality ascertainment, we did not redundantly access the NDI and only used HRS data to determine mortality status [47].

This approach for defining mortality outcome assessment dates was applied to the MIDUS M2 and MR1 cohorts to determine all-cause mortality outcomes. December 31, 2009, was used as the outcome assessment date for the M2 cohort. Two outcome assessment dates were applied for the MR1 cohort: December 31, 2016, for those between ages 40 and 54, and December 31, 2018, for those between ages 55 and 70. Individuals whose deaths occurred before the outcome assessment date were assigned a positive mortality status.

Model development and application

Within the HRS cohorts, the SFI-10 was constructed as described in [28]. See the Supplemental Appendix for an expanded description of model development and application. The ten items that comprise the SFI-10 include age, sex, neighborhood cleanliness, perceived control over financial situation, frequency of meeting with children, activities with grandchildren, nieces/nephews, or neighborhood children (youth), work status, volunteering, feeling isolated from others, and being treated with courtesy/respect (Table 1). All items were used to fit a logistic regression model that predicted mortality outcomes in the HRS 2010 cohort: the SFI-10. The final model was evaluated in the HRS 2012 cohort.

To replicate the SFI-10 within the MIDUS cohorts, minor adaptations were required based on the content of the MIDUS survey (Table 1). Three of ten items—age, sex, and perceived financial control—were measured identically between HRS and MIDUS and did not require adaptation. Two of ten items from HRS—neighborhood cleanliness and work status—had analogous survey items in MIDUS that differed only slightly in the item wording or measurement scale. Five of ten items from the HRS required minor item adaptation due to the absence of analogous survey items in MIDUS. For these items, substitutions were made to develop the MIDUS SFI-10 to ensure that the constructs measured by the SFI-10 across the two studies were congruent (Supplemental Appendix). The model development approach described for the HRS was identically applied in the MIDUS M2 cohort to produce an analog to the SFI-10, which was subsequently evaluated in the MIDUS MR1 cohort.

To evaluate item-level contributions to the SFI-10's performance, three logistic regression models predicting mortality outcomes composed of components of the SFI-10 were fit in the HRS and MIDUS development cohorts—the SFI-8, an age/sex-based model, and an age-based model—and applied to respective evaluation cohorts. Each of these models incorporates the methodology used for the SFI-10 (Supplemental Appendix).

External criteria measures

Shah et al. used mortality as a primary outcome when developing and evaluating the predictive utility of the SFI-10 [28]. This approach was replicated within the HRS and extended within MIDUS to include additional external criterion outcomes that would theoretically constitute the construct of social frailty under the predominant current framework and its theoretical basis [26, 27]. Criteria included self-evaluated physical and mental/emotional health, basic and instrumental activities of daily living, scales of psychological well-being [49], scales of social well-being [50], positive and negative affect [51, 52], scales of positive and negative trait emotionality from the Multidimensional Personality Questionnaire [53], and both familial and friendship affectual solidarity [54]. Each criterion was measured via self-report questionnaire concurrently with SFI-10 items. A detailed description of all criteria can be found in the Supplemental Appendix.

Statistical analysis

Statistical analyses were completed using R Statistical Software (v4.5.1 [55]). Significance testing was conducted at the $\alpha = 0.05$ level of significance, with directional alternative hypotheses specified where appropriate. The ggplot2 R package (v3.5.1 [56]) was used for data visualization.

Table 1 Social Frailty Index items in HRS and MIDUS

Domain	SFI-10 item	**HRS Psychosocial and Lifestyle Questionnaire** *Survey item*	**MIDUS Project 1: Survey** *Adapted survey item*
Demographics	Age	*Calculated* via*: date of interview - date of birth* *Rounded to the nearest year	*Calculated* via*: date of interview - date of birth* *Rounded to the nearest year
Demographics	Sex[a]	*Respondent's sex* 1, female 0, male	*Respondent's sex* 1, female 0, male
General resources	Neighborhood cleanliness	*How do you feel about your local area?* 1, this area is kept very clean; 7, this area is always full of rubbish and litter	*Please indicate how much "my neighborhood is kept clean" describes your situation* 1, a lot; 4, not at all
General resources	Perceived control over financial situation	*How would you rate the amount of control you have over your financial situation these days?* 0, no control; 10, very much control	*How would you rate the amount of control you have over your financial situation these days?* 0, no control; 10, very much control
Social resources/social activities	Has children/frequency of meeting up with children	*On average, how often do you meet up (arranged/chance) with your children, not counting any that live with you?* 1, 3 or more times a week; 6, less than once a year or never	*How often are you in contact with any members of your family—that is any of your brothers, sisters, parents, or children who do not live with you—including visits, phone calls, letters, or email?* 1, several times a day; 8, never or hardly ever
Social activities	Activities with grandchildren, nieces/nephews, or neighborhood children	*How often do you do activities with grandchildren, nieces/nephews, or neighborhood children?* 1, daily; 7, never/not relevant	*How often are you in contact with any of your friends—including visits, phone calls, letters, or email?* 1, several times a day; 8, never or hardly ever
Social activities	Work status	*Are you currently working?* 1, yes 0, no	*(What is) your current employment situation?* 1, working/self-employed/temp. laid off/leave 0, else
Social activities	Volunteering	*How often do you do any other [not with children or young people] volunteer or charity work?* 1, daily; 7, never/not relevant	*On average, how many hours per month do you spend doing formal volunteer work of the following types? [Hospital/nursing home; political organization/causes; other organization/cause]*

Table 1 (continued)

Domain	SFI-10 item	**HRS Psychosocial and Lifestyle Questionnaire** *Survey item*	**MIDUS Project 1: Survey** *Adapted survey item*
Fulfillment of basic social needs	Feeling isolated from others	*How much of the time do you feel isolated from others?* 1, often; 3, hardly ever or never	*During the past 30 days, how much of the time did you feel ____? [lonely; close to others; like you belong]* 1, all of the time; 5, none of the time *Adapted from [48]. 1st item reverse-coded, item scores summed to create loneliness composite. Rescaled to 1–13.
Fulfillment of basic social needs	Treated with less courtesy/respect	*In your day-to-day life, how often [does it happen that] you are treated with less courtesy or respect than other people?* 1, almost every day; 6, never	*How often on a day-to-day basis do you experience each of the following types of discrimination? You are treated with less (courtesy; respect) than other people* 1, often; 4, never *Both items summed. Rescaled to 1–7.

Items in the SFI-10 and their corresponding survey items in the HRS and MIDUS. Domain specifiers in relation to Bunt et al. (2017)'s social frailty framework are adapted from [28]. Only minimum and maximum scale values provided for each item. Each item scale consists of response options within respective ranges separated by 1-point intervals (denoted by ";"). [a]This predictor was labeled "gender" by Shah et al., which may be due to the former labeling of the variable in HRS data products as "gender." However, the verbiage used in the HRS core questionnaire refers to biological sex rather than gender. As such, we refer to this item as "sex"

Model development and application were conducted as described. To maintain consistency with Shah et al.'s methodology, BIC was used to determine which model was the best fit to development cohort data [28]. Assumptions of this approach are described elsewhere [57].

Identical methods were employed in aims one and two to investigate the extent to which the social components of the SFI-10 contribute to its performance in mortality prediction in the HRS and MIDUS cohorts, respectively. Model performance in the external evaluation cohorts was primarily measured using area under the receiver operating characteristic curve (AU_{ROC} [58]). AU_{ROC} was used as the primary model evaluation metric to maintain consistency with Shah et al.'s approach [28]. However, the HRS cohorts and MIDUS cohorts both exhibit class imbalance (a high number of individuals in the negative—e.g., survival—class, and a low number of individuals in the positive—e.g., mortality—class) with respect to mortality outcomes, with the latter imbalance being more extreme (HRS 2010: 22%, 2012: 23% deceased by assessment date; MIDUS M2: 4%, MR1: 4% deceased by assessment date). In such cases, there has been a proliferation of recommendations to use area under the precision-recall curve (AU_{PRC}) over AU_{ROC} for model performance evaluation [59, 60]. Thus, we replicated our performance evaluation analyses using AU_{PRC} to complement our AU_{ROC}-based analyses. These are presented in the supplementary material. It should be noted, however, that recent insights have challenged this recommendation and support the use of AU_{ROC} as an interpretable and robust metric under class imbalance [58, 61].

AU_{ROC} analyses were implemented via the pROC R package (v1.18.5 [62]). Bias-corrected and accelerated (BCa) confidence intervals generated via $n = 9999$ stratified bootstrap replications are reported for all AU_{ROC} estimates. To compare the performance of models within each cohort, BCa confidence intervals for differences in AU_{ROC} values between ROC curves were generated via $n = 9999$ stratified bootstrap replications. All bootstrap resampling was implemented in the boot R package (v1.3–31 [63, 64]).

To assess the contribution of each predictor to model performance of the SFI-10 and SFI-8 in the external evaluation cohorts, permutation-based feature importance analyses [65, 66] were conducted using a model-agnostic approach implemented within the iml R package (v0.11.4 [67]) using $n = 1000$ permutations per feature. Feature importance to model performance was measured via AU_{ROC} loss after feature permutation. This loss function was chosen to maintain consistency with the primary performance metric used to evaluate models.

To address the third study aim and evaluate the criterion validity of the SFI-10 and SFI-8 as measures of social frailty, Spearman rank correlation coefficients were utilized to determine the strength of associations between criteria relevant for social frailty and the SFI-10, SFI-8, and age in the MIDUS MR1 cohort. Predicted mortality risk on the logit scale from the SFI-10 and SFI-8 was used in correlational analysis. Confidence intervals for correlations were produced by Fisher's z-transformation and are reported at the 95% level. Steiger's z test implemented in the cocor R package (v1.1.4 [68]) was used to evaluate differences in effect size between SFI-10 and SFI-8 correlations with social frailty criteria, and Benjamini-Hochberg corrections [69] were applied to control the false discovery rate at 5% for multiple comparisons. Where relevant, Pearson product-moment correlation coefficients are reported for bivariate correlations between two continuous variables and between continuous and dichotomous variables.

Results

Revisiting the Social Frailty Index in the HRS

Cohort and SFI-10 replication

The 2010 ($n = 4308$) and 2012 ($n = 3956$) HRS cohorts were curated with minor differences from Shah et al. (see "Methods"; Table S1). The SFI-10 was reconstructed (Table 1; Supplemental Appendix) with very minor differences from Shah et al., given differences in cohort composition and functional forms analysis relating each item in the SFI-10 to mortality outcomes (Table S2). Of relevance, in the original Shah et al. paper, sex was labeled as "gender" [28]. However, the verbiage used in the HRS core questionnaire refers to biological sex rather than gender identity; thus, "sex" is a more accurate term and is used herein.

Incremental validity of social frailty items for mortality prediction

Next, the predictive utility of the SFI-10's social items was examined beyond the predictive utility of age and sex. We hypothesized that age and sex would emerge as the most important items for the prediction of mortality, but that items reflective of social-environmental factors would be incrementally predictive. To test this hypothesis, the predictive ability of the SFI-10 for mortality outcomes was compared to that of (i) an index comprised of the eight social items from the SFI-10 (termed the SFI-8), (ii) a model including age and sex only, and (iii) a model that included only age (Supplemental Appendix).

Within the HRS 2010 cohort, the SFI-10 (BIC = 3876.54) was the best fit model to the data that was used to develop the models, favored over the model consisting of only age (BIC = 3989.72), age and sex (BIC = 4019.24), and social items alone (SFI-8; BIC = 4247.93).

The performance of models in the external HRS 2012 cohort was primarily evaluated using the area under the receiver operating characteristic curve (AU_{ROC}), a widely used and interpretable metric for evaluating the performance of binary classification models [58]. AU_{ROC} values were assessed based on previously published criteria [70]. When applied to the external HRS 2012 cohort for evaluation, the SFI-10 had the best discriminatory capacity for mortality outcomes compared to the other models, with an AU_{ROC} value in the fair range (Fig. 1A; AU_{ROC} = 0.740 [95% CI 0.722–0.758]), comparable to what Shah et al. reported [28]. The model based on age and sex had the second-best discriminatory capacity for mortality outcomes (Fig. 1A; AU_{ROC} = 0.724 [95% CI 0.704–0.743]). Notably, the full SFI-10 exhibited a significantly larger AU_{ROC} than the model based on age and sex alone (ΔAU_{ROC} = 0.016 [one-sided, 90% CI 0.007–0.026]). Although the difference in AU_{ROC} observed between the SFI-10 and the age and sex model was statistically significant, the magnitude of this difference was small, with a value of 0.016. In contrast, a model based on age alone exhibited lower discriminatory capacity while still in the fair range (Fig. 1A; AU_{ROC} = 0.720 [95% CI 0.700–0.739]), and the SFI-8 exhibited the worst discriminatory capacity, in the "not useful" range (Fig. 1A; AU_{ROC} = 0.639 [95% CI 0.619–0.660]). Model performance was also evaluated using the area under the precision-recall curve (AU_{PRC}) to complement the AU_{ROC}-based analysis. The rank ordering of model performance was confirmed by AU_{PRC}-based comparisons (Fig. S1A). Notably, the SFI-10 *did not* exhibit a significantly larger AU_{PRC} value than the model based on age and sex alone (ΔAU_{PRC} = 0.012 [one-sided, 90% CI −0.004–0.029]), constituting a lack of support for the incremental utility of the social items in the SFI-10 for mortality prediction.

Item-level contributions to mortality prediction

To understand which individual items were most important to the SFI-10's predictive ability, beyond their contributions in aggregate, a model-agnostic permutation-based feature importance approach [65, 66] was applied using a loss function based on reduction in AU_{ROC} (Fig. 1B). In the HRS 2012 cohort, age emerged as the most important item in the SFI-10 for mortality prediction by a substantial margin. Sex was the second most important item in the SFI-10, but marginally so. *Volunteering* was the most important social item of the SFI-10. The remaining social items had low importance estimates. *Meeting with children*, being *treated with courtesy and respect*, and *neighborhood cleanliness* were the least important items in the SFI-10 and all had negative values as the lower bound of the 5th–95th percentile range for estimates of their feature importance (Table S3), indicating relative unimportance to the SFI-10's practical predictive ability for mortality. An identical approach using a loss function based on reduction in AU_{PRC} largely confirmed these findings, with slight differences in rank order between metrics (Fig. S1B; Table S4).

The same feature importance approach was applied to the SFI-8 in the HRS 2012 cohort (Fig. 1C; Table S3). The feature importance estimates were unanimously marginal, given the SFI-8's low AU_{ROC} value. *Activities with grandchildren*, *nieces/nephews*, *or neighborhood children (youth)* emerged as the most important item in the SFI-8 for mortality prediction. *Volunteering* was the second most important predictor based on AU_{ROC} loss. Again, *meeting with children*, *neighborhood cleanliness*, and *being treated with courtesy and respect* were the least important items in the SFI-8 and had negative values as the lower bound of the 5th–95th percentile range for estimates of their feature importance. Results based on

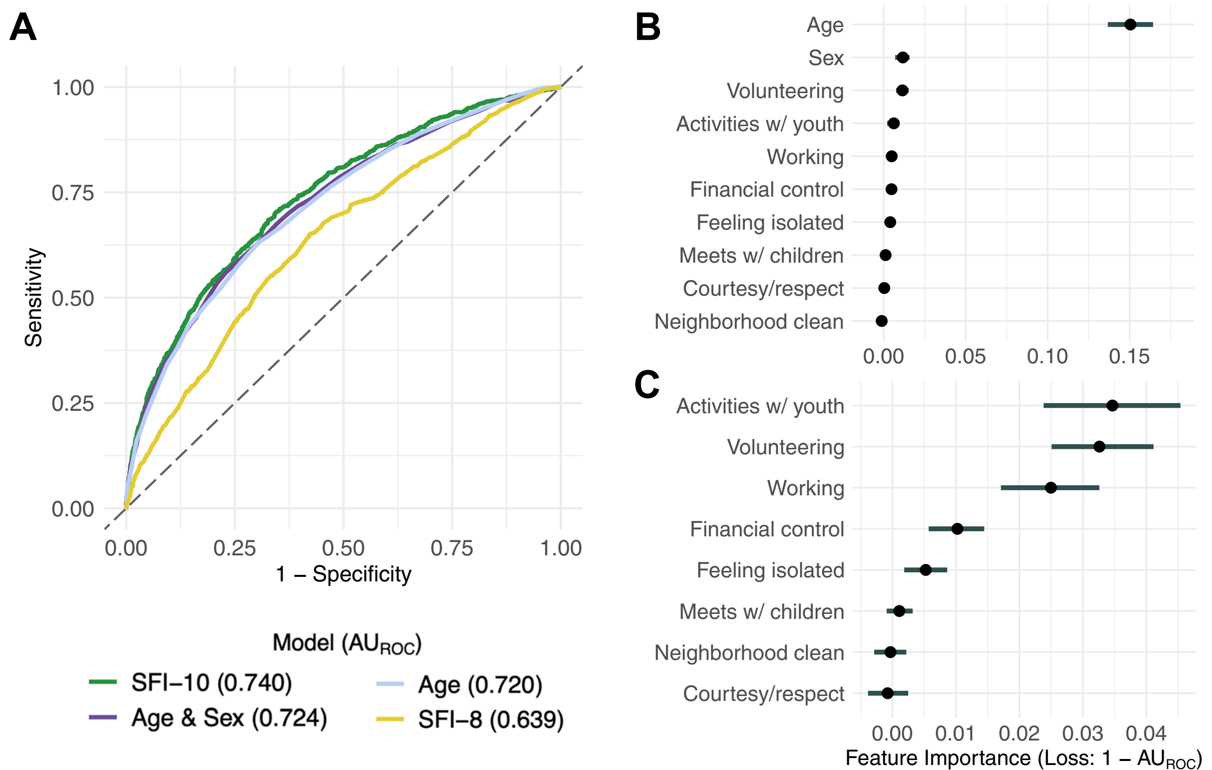

Fig. 1 Model comparisons and evaluation in HRS. Comparison of model performance in the HRS 2012 evaluation cohort. **A** ROC curves for all models evaluated in the 2012 cohort. Sensitivity = recall; true positive rate. 1 – Specificity = false positive rate. AU$_{ROC}$ is bound by 0 to 1, with higher values indicating better model performance. AU$_{ROC}$ values are comparable between samples and can be interpreted as the probability that a randomly selected case (death) will be ranked higher by a model than a randomly selected control (survival). AU$_{ROC}$ values have a universal random baseline of 0.50 (dashed line), indicating a model with no predictive value beyond chance. Permutation-based feature importance estimates for items in the **B** SFI-10 and **C** SFI-8 using loss functions based on reduction in AU$_{ROC}$. Estimates for median importance are plotted along with 5th–95th percentile bands for each item based on $n = 1000$ permutations per item. Greater loss values reflect greater feature importance. Activities w/youth: Activities with grandchildren, nieces/nephews, or neighborhood children

AU$_{PRC}$ largely confirmed these findings (Fig S1C; Table S4).

Application of the Social Frailty Index to MIDUS

Applying the SFI-10 to MIDUS

Functional forms analysis relating each predictor in the SFI-10 to mortality outcomes in the M2 cohort ($n = 3542$) yielded different relationships than those found in the HRS 2010 cohort for two items: *financial control* and *volunteering* (Table S5). Due to slight differences in item content between the HRS and MIDUS, the SFI-10 applied in MIDUS (Table 1; Supplemental Appendix) differs moderately from the HRS SFI-10, although differences in coefficients must be interpreted in the context of effect coding conventions used by Shah et al. [28] and differences across the two studies in SFI-10 item content.

Incremental validity of social frailty items for mortality prediction in MIDUS

Informed by the strong dependency of the SFI-10 on age and sex exhibited in the HRS 2012 cohort and the marginal contributions from the social items to the predictive ability of the SFI-10 for mortality outcomes, we assessed whether similar relationships would emerge in the MIDUS sample. Within the MIDUS cohorts, we compared the predictive ability

of the SFI-10 for mortality outcomes with that of (i) the SFI-8, (ii) age and sex alone, and (iii) age alone (Supplemental Appendix).

Within the M2 cohort, the SFI-10 was not the best fit to the data used to develop the models (BIC = 1141.61), as the models based on age (BIC = 1124.60) and both age and sex (BIC = 1125.96) had lower BIC values. The SFI-8 was the worst fit to the M2 cohort data (BIC = 1194.99).

When applied to the external MR1 evaluation cohort ($n = 2077$), the SFI-10 had the best discriminatory capacity for mortality outcomes of all models (Fig. 2A). The model based on age and sex had the second-best discriminatory capacity for mortality outcomes in the MR1 cohort, with a value in the fair range (Fig. 2A; $AU_{ROC} = 0.708$ [95% CI 0.655–0.759]). The SFI-10 again exhibited a significantly larger AU_{ROC} value than the model based on age and sex alone ($\Delta AU_{ROC} = 0.050$ [one-sided, 90% CI 0.017–0.087]), and unlike in the HRS, the magnitude of this difference was of a larger practical utility at 0.050. The SFI-8 exhibited the third-best discriminatory capacity, in the fair range (Fig. 2A; $AU_{ROC} = 0.703$ [95% CI 0.646–0.757]). Age alone exhibited the worst performance of all models, in the not useful AU_{ROC} range (Fig. 2A; $AU_{ROC} = 0.693$ [95% CI 0.640–0.743]). The rank ordering of model performance differed in AU_{PRC}-based comparisons (Fig. S2A). The SFI-10 was confirmed as the best performing model; however, the SFI-8 was second, the

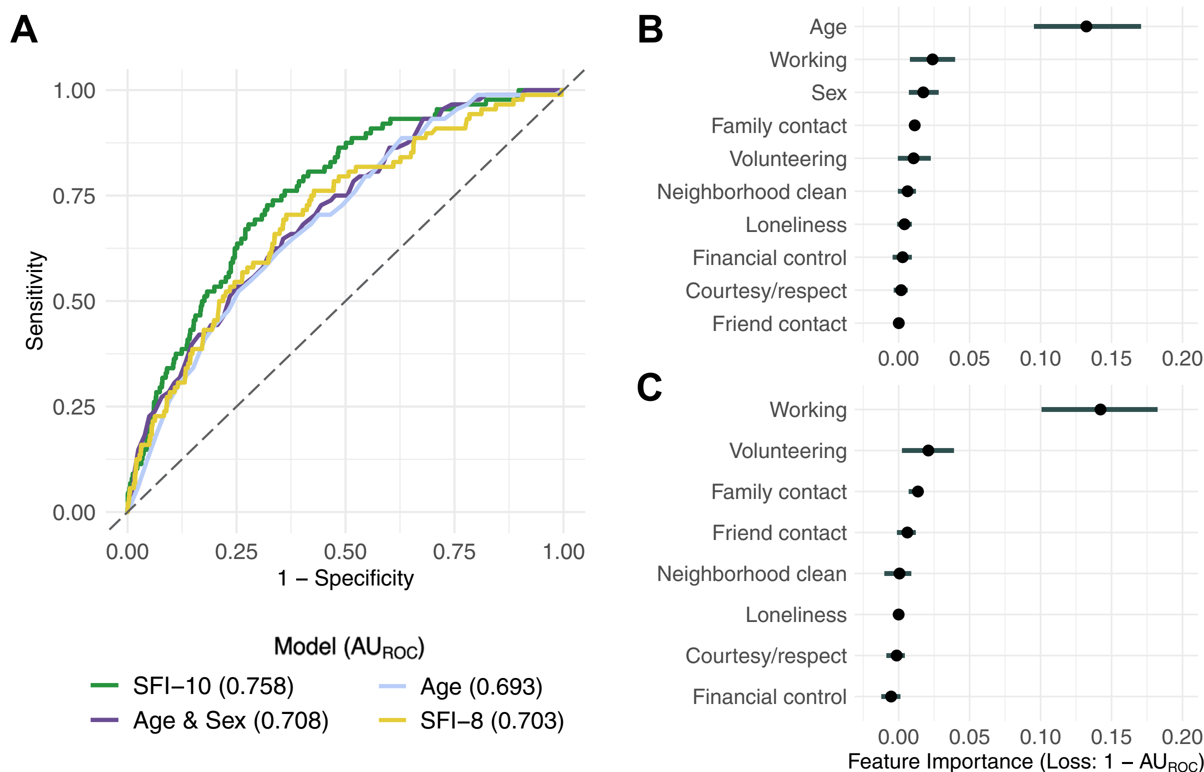

Fig. 2 Model comparisons and evaluation in MIDUS. Comparison of model performance in the MIDUS Refresher 1 (MR1) evaluation cohort. **A** ROC curves for all models evaluated in the MR1 cohort. Sensitivity = recall; true positive rate. 1 – Specificity = false positive rate. AU_{ROC} is bound by 0 to 1, with higher values indicating better model performance. AU_{ROC} values are comparable between samples and can be interpreted as the probability that a randomly selected case (death) will be ranked higher by a model than a randomly selected control (survival). AU_{ROC} values have a universal random baseline of 0.50 (dashed line), indicating a model with no predictive value beyond chance. Permutation-based feature importance estimates for items in the **B** SFI-10 and **C** SFI-8 using loss functions based on reduction in AU_{ROC}. Estimates for median importance are plotted along with 5th–95th percentile bands for each item based on $n = 1000$ permutations per item. Greater loss values reflect greater feature importance

model based on age and sex alone was third, and the model based on age alone was worst. The SFI-10 also exhibited a significantly larger AU_{PRC} value than the model based on age and sex alone ($\Delta AU_{PRC} = 0.040$ [one-sided, 90% CI 0.00002–0.088]), confirming the heightened incremental importance of the social items within the SFI-10 for mortality prediction in this sample.

Item-level contributions to mortality prediction

The model-agnostic permutation-based feature importance approach used in the HRS was applied to understand which items were most important to mortality prediction in the younger MR1 cohort (Fig. 2 B and C; Table S6). Age distinguished itself as the pre-eminent item from the SFI-10 for mortality prediction (Fig. 2B). *Working status* emerged as the second most important item in the SFI-10, while sex was the third most important item. Contextualizing its relative importance for mortality prediction, *working status* had the highest correlation with age of all the social items in the SFI-10 (H_0: $\rho \leq 0$, H_1: $\rho > 0$, $r_p(2075) = 0.49$, $t = 25.48$, $p < 0.001$). The remaining social items had marginal feature importance values. *Volunteering*, *neighborhood cleanliness*, *loneliness*, *perceived financial control*, *being treated with courtesy and respect*, and *friend contact* all had negative values as the lower bound of the 5th–95th percentile range for estimates of their feature importance. Results based on AU_{PRC} were largely confirmatory (Fig S2B; Table S7). Of note, they indicated that sex was the second most important item in the SFI-10, while *working status* was third, and additionally suggested lesser importance of the social items beyond that indicated by AU_{ROC}-loss.

From the SFI-8 (Fig. 2C; Table S6), *working* emerged as the most important item for mortality prediction. The remaining SFI-8 items had marginal feature importance estimates. *Volunteering* was the second most important item, followed by family contact. All other items of the SFI-8 had negative values as the lower bound of the 5th–95th percentile range for estimates of their feature importance, with *being treated with courtesy and respect* and *perceived financial control* having negative values for their feature importance estimates themselves, indicating unimportance to the SFI-8's practical predictive ability for mortality outcomes. AU_{PRC} loss–based analyses largely confirmed these findings, as all items besides *working* had negative values as the lower bound of the 5th–95th percentile range for estimates of their feature importance (Fig. S2C; Table S7).

Associations between social frailty, age, and external criteria

To validate the SFI-10 as a measure of social frailty in MIDUS, measures concordant with Bunt et al.'s [26] definition of social frailty—spanning health and functional status, well-being, and social adjustment—were selected as external criteria. Hypotheses regarding associations between a measure of social frailty and each criterion were formulated prior to analysis (Supplemental Appendix). Monotonic correlations between SFI-10 total index score and each criterion measure were calculated (Table 2). To examine how the social-environmental factors alone related to external criteria, correlations were additionally examined between the SFI-8 total index score and the criterion measures as well as between age and the criterion measures.

The hypothesized directional associations between the SFI-10 and each criterion were partially supported (Table 2). Reasonable effect sizes ($r_s > |0.10|$ [71]) were observed for 13 of the 26 criteria evaluated. When the SFI-8 was evaluated against these same criteria, hypothesized directional correlations of reasonable effect sizes were observed for 19 of the 26 criteria evaluated (Table 2). Comparing the effect sizes observed for the SFI-10 and SFI-8 for each criterion, the SFI-8 yielded correlations of significantly greater effect size in the hypothesized direction compared to the SFI-10 (H_0: SFI-8 ≤ SFI-10, H_1: SFI-8 > SFI-10) for 16 of the 26 criteria evaluated. Further, correlations with effect sizes that opposed directional hypotheses were observed between age and 15 of the 26 criteria evaluated (Table 2).

Discussion

Social frailty is a novel and potentially useful concept in geroscience, yet it is an underexplored construct with little empirical evaluation and a lack of consensus on measurement. Accordingly, the present study examined the construct of social frailty and how it can be best measured. Utilizing

Table 2 Correlations between the SFI-10, SFI-8, and age with external criteria in MIDUS

		r_s [95% CI]			
	Mean (SD)	SFI-10	Age	SFI-8	*p*
Health					
Physical	3.48 (1.11)	**−0.17 [−0.21, −0.13]**	−0.02 [−0.06, 0.03]	**−0.22 [−0.26, −0.17]**	.003
Mental/emotional	3.71 (1.02)	**−0.12 [−0.17, −0.08]**	0.05 [0.01, 0.09]	**−0.17 [−0.22, −0.13]**	.003
Activities of daily living					
Basic	1.41 (0.74)	**0.29 [0.25, 0.32]**	**0.18 [0.13, 0.22]**	**0.31 [0.27, 0.35]**	.141
Instrumental	1.89 (0.93)	**0.33 [0.29, 0.36]**	**0.24 [0.20, 0.28]**	**0.30 [0.26, 0.34]**	.985
Psychological well-being					
Autonomy	37.23 (6.97)	0.00 [−0.04, 0.05]	0.07 [0.03, 0.11]	−0.03 [−0.08, 0.01]	.019
Environmental mastery	37.46 (7.77)	−0.05 [−0.09, 0.00]	0.16 [0.11, 0.20]	**−0.12 [−0.16, −0.07]**	< 0.001
Personal growth	38.58 (6.73)	**−0.21 [−0.25, −0.17]**	−0.03 [−0.07, 0.01]	**−0.22 [−0.26, −0.18]**	.351
Positive relations	39.69 (7.28)	**−0.13 [−0.17, −0.08]**	0.10 [0.06, 0.15]	**−0.14 [−0.18, −0.10]**	.208
Purpose in life	38.26 (7.15)	**−0.21 [−0.25, −0.17]**	−0.01 [−0.05, 0.03]	**−0.23 [−0.27, −0.19]**	.177
Self-acceptance	37.25 (8.53)	−0.09 [−0.13, −0.05]	0.12 [0.08, 0.16]	**−0.15 [−0.19, −0.11]**	< 0.001
Social well-being					
Social coherence	9.32 (3.14)	**−0.19 [−0.23, −0.15]**	**−0.10 [−0.14, −0.05]**	**−0.21 [−0.25, −0.17]**	.170
Social integration	14.22 (4.00)	**−0.14 [−0.18, −0.10]**	0.09 [0.05, 0.13]	**−0.23 [−0.27, −0.18]**	< 0.001
Social acceptance	13.41 (3.45)	−0.03 [−0.07, 0.01]	0.14 [0.10, 0.18]	**−0.11 [−0.15, −0.07]**	< 0.001
Social contribution	15.70 (3.66)	**−0.26 [−0.30, −0.22]**	−0.06 [−0.11, −0.02]	**−0.30 [−0.34, −0.26]**	.014
Social actualization	11.63 (4.01)	**−0.17 [−0.21, −0.12]**	−0.05 [−0.10, −0.01]	**−0.18 [−0.22, −0.14]**	.227
Affect					
Positive	3.38 (0.77)	−0.08 [−0.12, −0.03]	0.13 [0.09, 0.17]	**−0.11 [−0.15, −0.07]**	.039
Negative	1.54 (0.68)	0.02 [−0.03, 0.06]	−0.17 [−0.21, −0.13]	**0.13 [0.09, 0.17]**	< 0.001
Multidimensional Personality Questionnaire					
Positive emotionality					
Well-being	9.00 (1.85)	−0.06 [−0.10, −0.02]	0.08 [0.04, 0.12]	**−0.11 [−0.15, −0.07]**	.004
Social potency	10.27 (2.39)	**−0.11 [−0.15, −0.07]**	−0.05 [−0.10, −0.01]	**−0.16 [−0.20, −0.11]**	.006
Achievement	12.33 (2.29)	−0.08 [−0.13, −0.04]	−0.02 [−0.06, 0.02]	**−0.13 [−0.18, −0.09]**	.003
Social closeness	11.47 (2.45)	**−0.12 [−0.16, −0.07]**	−0.01 [−0.05, 0.03]	−0.06 [−0.10, −0.01]	.999
Negative emotionality					
Stress reactivity	6.26 (2.25)	0.05 [0.00, 0.09]	−0.12 [−0.16, −0.08]	0.09 [0.05, 0.14]	.004
Aggression	5.52 (1.85)	0.07 [0.02, 0.11]	−0.07 [−0.12, −0.03]	0.06 [0.02, 0.11]	.650
Alienation	5.31 (1.99)	0.03 [−0.01, 0.07]	−0.13 [−0.17, −0.09]	0.08 [0.03, 0.12]	.006
Affectual solidarity					
Family	3.18 (0.56)	0.03 [−0.01, 0.08]	0.18 [0.14, 0.22]	−0.04 [−0.08, 0.01]	< 0.001
Friends	3.25 (0.47)	−0.08 [−0.12, −0.03]	0.07 [0.03, 0.12]	−0.07 [−0.11, −0.02]	.800

r_s, Spearman's rank correlation coefficient; degrees of freedom across all bivariate correlations (*df*) = 2075. *SD*, standard deviation. *SFI-10*, log odds of mortality modeled by the SFI-10. *SFI-8*, log odds of mortality modeled by the SFI-8. *p*, Benjamini-Hochberg adjusted *p*-values testing H_0: r_s SFI-8 $\leq r_s$ SFI-10 vs. H_1: r_s SFI-8 $> r_s$ SFI-10 in hypothesized directions specified for each criterion. Bold: $r_s > |.10|$ in hypothesized direction

cohorts from the HRS that originally produced the SFI-10 [28], as well as cohorts from the MIDUS study, we demonstrate that the SFI-10 consistently predicts mortality in adults from midlife to old age. However, our findings raise concerns about whether the SFI-10 should be conceptualized as a measure of social frailty despite being labeled as such. Specifically, (1) we found that the SFI-10 exhibited a striking reliance on chronological age to predict mortality outcomes within older adults

from the HRS, while its social items contributed little to its predictive ability; (2) this marked reliance on age for mortality prediction persisted when the scale was applied to adults at midlife through the transition to old age from MIDUS, although a heightened contribution of social items, in aggregate, was observed in this younger cohort; lastly, (3) in examining associations between the SFI-10 and external criterion measures relevant for social frailty, we demonstrated that an index derived from the eight social-environmental factors in the SFI-10 shows stronger associations with many criteria than the full index, which also includes age and sex. The patterning of negative associations between age and the same criteria provides evidence that the direct inclusion of age within the 10-item index may obscure associations between social-environmental factors and relevant external criteria that reflect other aspects of social behavior, subjective well-being, and psychological health. Overall, although the SFI-10 predicts mortality risk, our results caution against construing it as a measure of social frailty. We propose that a measure that exclusively indexes social behaviors and resources is needed to best capture the social frailty construct.

The SFI-10 predicts mortality from middle to older age

The naming and development of the SFI-10 tether it to social frailty as a construct, and the present study evaluated it as such. However, it must be acknowledged that the SFI-10 was developed for mortality prediction and risk stratification. Notwithstanding the items it utilizes to accomplish this purpose, we have successfully replicated its fair ability to predict mortality outcomes in two independent cohorts. Although the absolute acceptability of $AU_{ROC/PRC}$ values depends on the problem to which a model is applied, the SFI-10 performed modestly at predicting mortality outcomes for older adults in the HRS, and slightly better in adults from midlife to the transition to old age from MIDUS. Given its parsimonious design, the SFI-10 may be suited to inform decision-making and intervention planning in clinical settings where mortality risk is a consideration.

Marginal contributions of social-environmental factors to mortality prediction in older adults

We hypothesized that age and sex would emerge as the most important items for predicting mortality outcomes in older adults from the HRS, but that the SFI-10's social items would be incrementally useful. The former hypothesis was strongly supported and the directional effects of these items on mortality risk in the HRS SFI-10—increasing chronological age and being biologically male corresponding to an increased mortality risk—as well as their importance for mortality prediction are consistent with previous reports [29–32].

However, the hypothesis that the social items would be predictive of mortality in and of themselves received marginal support. Incremental validity-wise, the SFI-10 showed a significant increase in mortality prediction performance compared to the model based on age and sex alone only with respect to AU_{ROC}. Although statistically significant, the social items in aggregate only added 0.016 to the AU_{ROC}, a value that may have little practical significance. While the SFI-10 emerged as the best fit model to the development cohort data, this marginal practical predictive improvement of the SFI-10 over the model without the social items calls into question the incremental validity of the SFI-10 for mortality prediction over age and sex alone. Moreover, the lack of predictive ability exhibited by the SFI-8 in the HRS evaluation cohort provides evidence against the relevance of the social items it includes for predicting mortality in older adults. In aggregate, this analysis offers evidence that the index developed by Shah et al. [28] in older adults from the HRS is largely dependent on non-social items to accomplish its purpose, which is to predict mortality. This pattern of findings reinforces the notion that risk prediction should not be conflated with construct measurement.

In terms of risk prediction, prima facie, these findings imply that the social-environmental factors reflected within the SFI-10 do not predict mortality in older adults. Given that countless studies link various social determinants to health across the lifespan [19, 21–23], this patterning may suggest that different social contributions to mortality risk should be indexed or that they might vary depending on when in the lifespan they are measured. The reliance on both chronological age and sex as predictors for

performance is not detrimental to the SFI-10's construct validity as a measure of mortality risk, as a mortality risk prediction model should contain items most predictive of mortality outcomes.

Yet we are still left with the question of how to best measure social frailty. Simply put, a construct-valid measure of social frailty should ostensibly contain items that reflect social behaviors and resources (e.g., age and sex are not in Bunt et al.'s framework [26]). The presence of these attributes would constitute social robustness, while the absence would constitute frailty. Theoretically, a social frailty index constituted from this item content should be primarily predictive of proximal outcomes such as subjective well-being, although distally predictive of aging-related outcomes such as mortality. Thus, although the inclusion of age and sex is conducive to the SFI-10's ability to predict mortality, it limits the interpretation of the index as a measure of social frailty.

Differential importance of social-environmental factors for mortality prediction from midlife to old age

Results in the MIDUS MR1 cohort varied meaningfully from results obtained in the HRS 2012 cohort, indicating that the importance of social-environmental factors for mortality prediction may vary across age groups. Compared to lackluster support for the added utility of the social items in the HRS, significant and practically useful increases in model performance metrics for the SFI-10 over the model based on age and sex alone observed in MIDUS indicate that the social items had heightened contributions to mortality prediction. Further, the SFI-8's AU_{ROC} and AU_{PRC} values indicated its usefulness for mortality prediction in MIDUS, compared to its relatively low values in the HRS.

Acknowledging that age was again unequivocally the most important SFI-10 item for mortality prediction in the MIDUS MR1 cohort, noteworthy differences in which social items ranked highest in importance emerged between evaluation cohorts. In the HRS, volunteering was the most important social item for mortality prediction using the 10-item index and activities with youth was the most important when using the 8-item index. Across both indices in the MIDUS MR1 cohort, working status was the most important social item. Although it showed a lack of importance AU_{PRC} loss-wise using the 10-item index, working status exhibited an inflated importance for mortality prediction across metrics using the 8-item index, potentially due to its high correlation with age. These results suggest that different facets of one's social experience—volunteering and activities with youth in old age vs. working status in midlife—may be more predictive of mortality across age cohorts, perhaps revealing malleability in social-environmental risk factors with increasing age. Interestingly, each of these items that contribute most to mortality prediction is within the social behavior/activities domain of Bunt's social frailty framework [26]. The importance of items indexing social behavior/activity for predicting mortality is congruent with theories of successful aging and empirical support linking a lack of social activity to increased mortality risk [72–75], while our results provide evidence that the relationship between the social behavior/activities domain of Bunt's framework and mortality is strongest relative to other domains. Patterns in importance of specific items indexing social behavior/activities may reflect changes in time use with increasing age and opportunities available for social activities once individuals reach retirement age [76–78].

Overall, the heightened contribution of social-environmental factors for mortality prediction closer to midlife compared to old age illustrates the plausibility of leveraging modifiable social behaviors and factors of one's social environment within interventions at midlife seeking to promote healthy aging trajectories through old age. This underscores the potential utility of social frailty assessment in prevention and early intervention efforts, rather than purely in risk stratification, and highlights the translational and clinical relevance of social frailty as a construct.

The SFI-10 and the nomological network of social frailty: implications for an emerging construct

The SFI-10 was originally developed as an index of social risk factors that predict mortality. Nevertheless, it was labeled a "social frailty index" by its creators, implying that it is a measure of social frailty. To be considered a valid measure of a construct, a measurement must relate to observables within said construct (or rather, its nomological network [79]). For example, a "depression index" should show a positive relationship with known depressive symptoms

to be considered a true measurement of depression. Thus, to determine if a social frailty index is indeed a measure of social frailty, it must be subjected to a set of criterion validity tests to evaluate its coherence with the social frailty construct. As measured in different ways by various instruments, social frailty has been associated with functional decline and disability [35, 37, 38, 80], depressive symptoms [36, 81], malnutrition risk [36], physical frailty [82, 83], dementia [84], and all-cause mortality [28, 33, 38, 80]. The current study is one of the first attempts (see [85]) to take a step back from such distal criteria and comprehensively explore the psychosocial aspects of the nomological network of social frailty to empirically evaluate whether a proposed social frailty index intersects with observables that cohere with contemporary formulations of the construct such as psychological health, subjective well-being, and social adjustment.

While the full SFI-10 does indeed show an intersection with observable outcomes within Bunt's social frailty framework, its social items appear to largely drive this intersection, as indicated by the presence of stronger associations between those items (i.e., the SFI-8) and external criteria, as well as age correlations that often directionally oppose these associations.

In this light, associations between the SFI-8 and social frailty criteria provide insight into what aspects of Bunt's theoretical framework are empirically supported. Correlations of reasonable effect size in hypothesized directions were observed between the SFI-8 and indicators of subjective well-being, as well as physical and social well-being, which underlie subjective well-being in SPF theory [27]. These relationships strongly suggest that subjective, social, and physical well-being are critical to the construct of social frailty. Interestingly, correlations of reasonable effect size were not observed between the SFI-8 and measures of affective solidarity. Affectual solidarity captures the type and degree of positive and negative sentiments held in social relationships and the degree of reciprocity of such sentiments and was included in our analyses to capture the degree to which basic social needs are met [54]. The lack of strong support for this relationship raises questions about the lack of basic social needs fulfillment being an etiologic mechanism for the manifestation of social frailty.

This notion leads to two overarching questions that will determine the nature of social frailty's application and usefulness in research, clinical, and policy-making settings. These concern (1) the nature of social frailty and (2) how we can best quantify it.

Before Bunt et al.'s work [26], most early studies on social frailty defined it as the sum of component items in a measurement but lacked any explicit theoretical basis motivating why items were originally selected. Bunt et al. produced their definition of social frailty by superimposing SPF theory on studies addressing social frailty through a scoping review, but research supporting the structure of their framework is scarce (see [36]). Alternatively, Bessa et al. conducted a systematic review of frailty instruments with social components, taking an atheoretical approach to summarize their items into eleven domains that potentially comprise social frailty [24]. Additional efforts to define social frailty are undoubtedly required and must be concordant with established frailty frameworks, geroscience principles, integrate existing social frailty research, and employ data-driven approaches to identify social factors that are relevant to the construct. These, in turn, should be validated against concurrent aspects of health and well-being as we attempted with the SFIs here and then tested in terms of their utility in predicting adverse age-related outcomes. In defining social frailty, it may be important to delineate between social, environmental, and psychological phenomena to ensure that the construct stays true to its social nature and does not become amorphous. For example, although negative psychosocial outcomes (e.g., low positive affect, low psychological well-being) may result from social frailty, they should not be considered part of the construct. In this regard, it is notable that the SFI-8 shows stronger associations with psychological well-being in our analyses than with narrowly defined social relationship indicators, e.g., affectual solidarity, which are more "true" to the social frailty construct [26]. Because the items that comprise the SFI-8 were keyed for mortality prediction, they may not best represent social frailty as conceptualized by dominant theoretical frameworks. Accordingly, a renewed effort to devise a psychometrically valid index of social frailty that conceptualizes the construct in relation to social behaviors and resources is needed. Focusing on social behaviors and resources would anchor social frailty in the social domain and avoid conflating construct measurement with risk prediction.

Informed by the results of the present study, we recommend using the SFI-10 as a measure of mortality risk, not as a measure of social frailty. The SFI-8, on the other hand, reflecting only social-environmental factors, shows promise as a measure of social frailty but requires further scrutiny in terms of its item content and implementation. A primary consideration in its use is how the index is constructed quantitatively. The SFI-8, like the SFI-10, was derived from a logistic regression model quantifying the probability of mortality, which we construed as a continuous measure of social frailty. Independent of our study, Matison et al. [84] recently used the items in the SFI-8, sans courtesy/respect, and a sum score approach to define social frailty under the *phenotype* model, finding that it predicted dementia risk better than other social frailty measures. We suggest that careful scrutiny of item content and evaluation against external criteria related to social behavior is needed to ensure that what is being measured is, indeed, social frailty.

Related to social frailty, the deficit accumulation model of frailty has been adapted to quantify "social vulnerability" [86, 87]. Social vulnerability focuses on the social environment of an individual and is used as a predictor of outcomes not necessarily always related to health or the aging process [87]; however, applied to items reflective of social behavior, the *deficit accumulation* model may be a useful framework for quantifying social frailty. Given the evolutionary-conserved relevance of social factors and determinants for health and aging [22, 88], endeavors to do so may use tools developed in preclinical literature to reverse translate quantifiable behavioral items into novel human indices of social frailty. For example, a recent attempt to quantify the social dimensions of aging in animal models led to the development of the Mouse Social Frailty Index, which includes seven assays that measure social behavioral functioning in mice across social communication, social interaction, and social functional ability [89, 90]. A translational framework for social frailty quantification across species would support geroscience, translational findings, and clinical understanding with regard to the social dimensions of the aging process [88].

It is within the outlined constraints that we advance a slightly revised definition of social frailty: *a continuum defined by a lack of adaptive social behaviors and resources that are necessary for healthy aging.* Social frailty, defined in this manner, is predicted to show general increases with age and confer risk for various adverse mental and physical health outcomes. Rates of increase in social frailty and associated risks vary between individuals, with higher levels reflecting accelerated biological aging. Future research should focus on empirically supporting this definition, advancing the construct of social frailty across species, and producing a construct-valid dimensional measure of social frailty that captures aging-related social declines that increase vulnerability to adverse health outcomes.

Strengths and limitations

The present study has limitations that must be considered. Differences between the HRS and MIDUS cohorts with respect to both demographics and response frequencies on items comprising the SFIs limit our comparison across studies and age groups, as differences interpreted as effects of life stage may be artifacts of these cohort effects. There were also differences in scale content between the HRS and MIDUS surveys. Although item substitutions were made to ensure that the same constructs were being measured across scales, exact equivalence may not have been achieved. Additionally, our methodological approach was constrained by our attempt to directly replicate Shah and colleagues' approach. For example, a survival analysis framework may have been more appropriate for mortality risk prediction.

Nonetheless, our study does have marked strengths supporting the robustness of our conclusions. We successfully replicated Shah et al.'s SFI using the original HRS dataset in which it was developed, confirming its predictive utility in an independent analysis. Furthermore, we extended its application to the MIDUS sample, demonstrating the index's generalizability across age cohorts and survey instruments. Importantly, our use of two large nationally representative samples of aging adults in the USA provides inferential power to support our conclusions and their generalizability.

Conclusion

Here we replicated, extended, and reevaluated the Social Frailty Index introduced by Shah and

colleagues [28], highlighting its utility as a tool to predict mortality in aging individuals, but questioning its validity as a measure of social frailty. In aggregate, our findings suggest that while the SFI-10 has potential for mortality risk prediction and stratification applications, it may not be a valid index of social frailty. Further, social-environmental factors included in the SFI-10 do not appear to predict mortality in and of themselves in older adults although they are more strongly predictive closer to midlife. Future efforts seeking to quantify social frailty should unite theory and empiricism in measurement creation, focus on social behaviors and resources, and consider translational approaches that connect the social dimensions of aging across species. Clarifying what constitutes social frailty and how best to measure it remains a critical step for advancing theory, informing interventions, and guiding policy aimed at the social dimensions of aging and their implications.

Acknowledgements The authors would like to express sincere gratitude to Scott Vrieze and the members of the Brain and Behavioral Processes Laboratory for invaluable discussions and feedback through the development and drafting of this manuscript.

Author contribution C.W.C., A.B., and M.L. conceptualized the study; C.W.C. and M.L. designed the study; C.W.C. performed research and analyzed the data; C.W.C., A.B., and M.L. wrote the paper.

Funding This work was supported by the Graduate Student Research Fellowship from the Department of Psychology at the University of Minnesota (C.W.C.) and NIH/NIA R61/R33 AG078520 (A.B.). The HRS is sponsored by the National Institute on Aging (U01 AG009740) and is conducted by the University of Michigan. Since 1995, the MIDUS study has been funded by the John D. and Catherine T. MacArthur Foundation Research Network and the National Institute on Aging (P01 AG020166, U19 AG051426).

Data availability Data from the HRS used in this study is available through registration at: https://hrs.isr.umich.edu/. RAND's 2020 HRS longitudinal file (v2 [91]) was used to curate the analytic cohort and for descriptive analysis. RAND's 2010 HRS fat file (v6A [92]) and 2012 HRS fat file (v3A [93]) were used for analysis. Data from MIDUS used in this study is available through registration at: https://midus.colectica.org/. Custom variable lists were used to curate the analytic cohort and for analysis (downloaded 6/30/2025). Our analyses were not preregistered. Code for data preparation and analysis to reproduce the results herein will be made publicly available at: https://github.com/ccollinge

Declarations

Ethics approval and and consent to participate All individuals included from the HRS and MIDUS provided informed consent. Because this study was a secondary data analysis of deidentified data, it was determined exempt from oversight by the University of Minnesota-Twin Cities Institutional Review Board.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet. 2013;381(9868):752–62. https://doi.org/10.1016/S0140-6736(12)62167-9.
2. Diebel LWM, Rockwood K. Determination of biological age: geriatric assessment vs biological biomarkers. Curr Oncol Rep. 2021;23(9):104. https://doi.org/10.1007/s11912-021-01097-9.
3. Elliott ML, Caspi A, Houts RM, Ambler A, Broadbent JM, Hancox RJ, et al. Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy. Nat Aging. 2021;1(3):295–308. https://doi.org/10.1038/s43587-021-00044-4.
4. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet. 2019;394(10206):1365–75. https://doi.org/10.1016/S0140-6736(19)31786-6.
5. Kane AE, Sinclair DA. Frailty biomarkers in humans and rodents: current approaches and future advances. Mech Ageing Dev. 2019;180:117–28. https://doi.org/10.1016/j.mad.2019.03.007.
6. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci. 2001;56(3):M146–57. https://doi.org/10.1093/gerona/56.3.M146.
7. Kojima G. Frailty as a predictor of future falls among community-dwelling older people: a systematic review and meta-analysis. J Am Med Dir Assoc.

2015;16(12):1027–33. https://doi.org/10.1016/j.jamda.2015.06.018.
8. Kojima G. Frailty as a predictor of hospitalisation among community-dwelling older people: a systematic review and meta-analysis. J Epidemiol Community Health. 2016;70(7):722–9. https://doi.org/10.1136/jech-2015-206978.
9. Kojima G, Iliffe S, Walters K. Frailty index as a predictor of mortality: a systematic review and meta-analysis. Age Ageing. 2018;47(2):193–200. https://doi.org/10.1093/ageing/afx162.
10. Rockwood K, Blodgett JM, Theou O, Sun MH, Feridooni HA, Mitnitski A, et al. A frailty index based on deficit accumulation quantifies mortality risk in humans and in mice. Sci Rep. 2017;7(1):1. https://doi.org/10.1038/srep43068.
11. Dent E, Kowal P, Hoogendijk EO. Frailty measurement in research and clinical practice: a review. Eur J Intern Med. 2016;31:3–10. https://doi.org/10.1016/j.ejim.2016.03.007.
12. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. Sci World J. 2001;1:323–36. https://doi.org/10.1100/tsw.2001.58.
13. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci. 2007;62(7):722–7. https://doi.org/10.1093/gerona/62.7.722.
14. Morley JE, Vellas B, van Abellan Kan G, Anker SD, Bauer JM, Bernabei R, et al. Frailty consensus: a call to action. J Am Med Dir Assoc. 2013;14(6):392–7. https://doi.org/10.1016/j.jamda.2013.03.022.
15. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr. 2008;8(1):24. https://doi.org/10.1186/1471-2318-8-24.
16. Crimmins EM. Social hallmarks of aging: suggestions for geroscience research. Ageing Res Rev. 2020;63:101136. https://doi.org/10.1016/j.arr.2020.101136.
17. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: an expanding universe. Cell. 2023;186(2):243–78. https://doi.org/10.1016/j.cell.2022.11.001.
18. Lyons CE, Razzoli M, Bartolomucci A. The impact of life stress on hallmarks of aging and accelerated senescence: connections in sickness and in health. Neurosci Biobehav Rev. 2023;153:105359. https://doi.org/10.1016/j.neubiorev.2023.105359.
19. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med. 2010;7(7):e1000316. https://doi.org/10.1371/journal.pmed.1000316.
20. Moffitt TE. Behavioral and social research to accelerate the geroscience translation agenda. Ageing Res Rev. 2020;63:101146. https://doi.org/10.1016/j.arr.2020.101146.
21. Schrempft S, Belsky DW, Draganski B, Kliegel M, Vollenweider P, Marques-Vidal P, et al. Associations between life-course socioeconomic conditions and the pace of aging. J Gerontol A Biol Sci Med Sci. 2022;77(11):2257–64. https://doi.org/10.1093/gerona/glab383.
22. Snyder-Mackler N, Burger JR, Gaydosh L, Belsky DW, Noppert GA, Campos FA, et al. Social determinants of health and survival in humans and other animals. Science. 2020;368(6493):eaax9553. https://doi.org/10.1126/science.aax9553.
23. Stringhini S, Carmeli C, Jokela M, Avendaño M, Muennig P, Guida F, et al. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. Lancet. 2017;389(10075):1229–37. https://doi.org/10.1016/S0140-6736(16)32380-7.
24. Bessa B, Ribeiro O, Coelho T. Assessing the social dimension of frailty in old age: a systematic review. Arch Gerontol Geriatr. 2018;78:101–13. https://doi.org/10.1016/j.archger.2018.06.005.
25. Rockwood K, Fox RA, Stolee P, Robertson D, Beattie BL. Frailty in elderly people: an evolving concept. Can Med Assoc J. 1994;150(4):489–95.
26. Bunt S, Steverink N, Olthof J, van der Schans CP, Hobbelen JSM. Social frailty in older adults: a scoping review. Eur J Ageing. 2017;14(3):323–34. https://doi.org/10.1007/s10433-017-0414-7.
27. Ormel J, Lindenberg S, Steverink N, Verbrugge LM. Subjective well-being and social production functions. Soc Indic Res. 1999;46(1):61–90. https://doi.org/10.1023/A:1006907811502.
28. Shah SJ, Oreper S, Jeon SY, Boscardin WJ, Fang MC, Covinsky KE. Social frailty index: development and validation of an index of social attributes predictive of mortality in older adults. Proc Natl Acad Sci U S A. 2023;120(7):e2209414120. https://doi.org/10.1073/pnas.2209414120.
29. Austad SN. Chapter 23—sex differences in longevity and aging. In E. J. Masoro & S. N. Austad (Eds.), *Handbook of the biology of aging (seventh edition)*. Academic Press; 2011, pp. 479-495. https://doi.org/10.1016/B978-0-12-378638-8.00023-3
30. Austad SN, Fischer KE. Sex differences in lifespan. Cell Metab. 2016;23(6):1022–33. https://doi.org/10.1016/j.cmet.2016.05.019.
31. Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos Trans R Soc Lond. 1825;115:513–83. https://doi.org/10.1098/rstl.1825.0026.
32. Kirkwood TBL. Deciphering death: a commentary on Gompertz (1825) 'on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies.' Philos Trans R Soc Lond B Biol Sci. 2015;370(1666):20140379. https://doi.org/10.1098/rstb.2014.0379.
33. Hao M, Hu Z, Zhang X, Li X, Wang S, Li Y, et al. Social frailty and mortality risk in middle-aged and older adults: a prospective cohort study. Geroscience. 2025. https://doi.org/10.1007/s11357-025-01879-0.
34. Ma L, Sun F, Tang Z. Social frailty is associated with physical functioning, cognition, and depression, and predicts mortality. J Nutr Health Aging. 2018;22(8):989–95. https://doi.org/10.1007/s12603-018-1054-0.
35. Makizako H, Shimada H, Tsutsumimoto K, Lee S, Doi T, Nakakubo S, et al. Social frailty in community-dwelling older adults as a risk factor for disability. J Am Med Dir Assoc. 2015;16(11):1003.e7-1003.e11. https://doi.org/10.1016/j.jamda.2015.08.023.

36. Pek K, Chew J, Lim JP, Yew S, Tan CN, Yeo A, et al. Social frailty Is independently associated with mood, nutrition, physical performance, and physical activity: insights from a theory-guided approach. Int J Environ Res Public Health. 2020. https://doi.org/10.3390/ijerph17124239.
37. Teo N, Gao Q, Nyunt MSZ, Wee SL, Ng T-P. Social frailty and functional disability: findings from the Singapore Longitudinal Ageing Studies. J Am Med Dir Assoc. 2017;18(7):637.e13-637.e19. https://doi.org/10.1016/j.jamda.2017.04.015.
38. Yamada M, Arai H. Social frailty predicts incident disability and mortality among community-dwelling Japanese older adults. J Am Med Dir Assoc. 2018;19(12):1099–103. https://doi.org/10.1016/j.jamda.2018.09.013.
39. Belsky DW, Caspi A, Houts R, Cohen HJ, Corcoran DL, Danese A, et al. Quantification of biological aging in young adults. Proc Natl Acad Sci U S A. 2015;112(30):E4104–10. https://doi.org/10.1073/pnas.1506264112.
40. Brim OG, Ryff CD, Kessler RC. Chapter one. The MIDUS national survey: an overview. In O. G. Brim, C. D. Ryff, & R. C. Kessler (Eds.), *How healthy are we?: A national study of well-being at midlife*. University of Chicago Press; 2004. pp. 1-34 https://doi.org/10.7208/9780226074764-002
41. Moffitt TE, Belsky DW, Danese A, Poulton R, Caspi A. The Longitudinal Study of Aging in Human Young Adults: knowledge gaps and research agenda. J Gerontol A Biol Sci Med Sci. 2017;72(2):210–5. https://doi.org/10.1093/gerona/glw191.
42. Heeringa SG, Connor JH. Technical description of the health and retirement study sample design. *University of Michigan*; 1995. Retrieved February 27, 2025 from: https://hrs.isr.umich.edu/publications/biblio/5310
43. Juster FT, Suzman R. An overview of the Health and Retirement Study. J Hum Resour. 1995;30:S7. https://doi.org/10.2307/146277.
44. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the Health and Retirement Study (HRS). Int J Epidemiol. 2014;43(2):576–85. https://doi.org/10.1093/ije/dyu067.
45. Smith J, Ryan L, Larkina M, Sonnega A, Weir D. Psychosocial and Lifestyle Questionaire 2006—2022. *University of Michigan*; 2023. Retrieved March 6, 2025 from: https://hrs.isr.umich.edu/publications/biblio/12903
46. Ryff CD, Krueger RF. Chapter one. Approaching human health as an integrative challenge: introduction and overview. In *The Oxford handbook of integrative health science, Carol D. Ryff, Robert F. Krueger (eds.)*. Oxford University Press; 2018. pp. 3-22.
47. Weir D. Validating mortality ascertainment in the Health and Retirement Study. *University of Michigan*; 2016. Retrieved March 7, 2025 from: https://hrs.isr.umich.edu/sites/default/files/biblio/Weir_mortality_ascertainment.pdf
48. Freilich CD, Mann FD, Krueger RF. Comparing associations between personality and loneliness at midlife across three cultural groups. J Pers. 2023;91(3):653–66. https://doi.org/10.1111/jopy.12765.
49. Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. J Pers Soc Psychol. 1989;57(6):1069–81. https://doi.org/10.1037/0022-3514.57.6.1069.
50. Keyes CLM. Social well-being. Soc Psychol Q. 1998;61(2):121–40. https://doi.org/10.2307/2787065.
51. Mroczek DK. Chapter seven. Positive and negative affect at midlife. In O. G. Brim, C. D. Ryff, & R. C. Kessler (Eds.), *How healthy are we?: a national study of well-being at midlife*. University of Chicago Press; 2004. pp. 250-226. https://doi.org/10.7208/9780226074764-002
52. Mroczek DK, Kolarz CM. The effect of age on positive and negative affect: a developmental perspective on happiness. J Pers Soc Psychol. 1998;75(5):1333–49. https://doi.org/10.1037/0022-3514.75.5.1333.
53. Patrick CJ, Curtin JJ, Tellegen A. Development and validation of a brief form of the Multidimensional Personality Questionnaire. Psychol Assess. 2002;14(2):150–63. https://doi.org/10.1037/1040-3590.14.2.150.
54. Grzywacz JG, Marks NF. Family solidarity and health behaviors: evidence from the National Survey of Midlife Development in the United States. J Fam Issues. 1999;20(2):243–68. https://doi.org/10.1177/019251399020002004.
55. R Core Team. (2024). *R: a langauge and environment for statistical computing* (version 4.4.1) [computer software]. R Foundation for Statistical Computing. https://www.R-project.org/
56. Wickham H. Ggplot2: elegant graphics for data analysis. Springer-Verlag New York; 2016.
57. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods. 2012;17(2):228–43. https://doi.org/10.1037/a0027127.
58. Richardson E, Trevizani R, Greenbaum JA, Carter H, Nielsen M, Peters B. The receiver operating characteristic curve accurately assesses imbalanced datasets. Patterns. 2024. https://doi.org/10.1016/j.patter.2024.100994.
59. Davis J, Goadrich M (2006) The relationship between precision-recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning, ICML '06, 233–240. https://doi.org/10.1145/1143844.1143874
60. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One. 2015;10(3):e0118432. https://doi.org/10.1371/journal.pone.0118432.
61. McDermott MBA, Zhang H, Hansen LH, Angelotti G, Gallifant J. A closer look at AUROC and AUPRC under class imbalance (arXiv:2401.06091). arXiv. 2025. https://doi.org/10.48550/arXiv.2401.06091.
62. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12(1):77. https://doi.org/10.1186/1471-2105-12-77.
63. Canty A, Ripley B. *boot: Bootstrap R (S-Plus) functions* (Version 1.3–31) [Computer software]. 2024. https://cran.r-project.org/web/packages/boot/index.html
64. Davison AC, Hinkley DV. Bootstrap methods and their application. Cambridge University Press; 1997. https://doi.org/10.1017/CBO9780511802843.

65. Breiman L. Random forests. Mach Learn. 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324.
66. Fisher A, Rudin C, Dominici F. All models are wrong, but many are useful: learning a variable's importance by studying an entire class of prediction models simultaneously. J Mach Learn Res. 2019;20:177.
67. Molnar C, Casalicchio G, Bischl B. Iml: an R package for interpretable machine learning. J Open Source Softw. 2018;3(26):786. https://doi.org/10.21105/joss.00786.
68. Diedenhofen B, Musch J. Cocor: a comprehensive solution for the statistical comparison of correlations. PLoS One. 2015;10(4):e0121945. https://doi.org/10.1371/journal.pone.0121945.
69. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57(1):289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
70. Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: evaluation and interpretation of receiver operating characteristic curves. Surgery. 2016;159(6):1638–45. https://doi.org/10.1016/j.surg.2015.12.029.
71. Gignac GE, Szodorai ET. Effect size guidelines for individual differences researchers. Pers Individ Differ. 2016;102:74–8. https://doi.org/10.1016/j.paid.2016.06.069.
72. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med. 2000;51(6):843–57. https://doi.org/10.1016/S0277-9536(00)00065-4.
73. Herzog AR, Ofstedal MB, Wheeler LM. Social engagement and its relationship to health. Clin Geriatr Med. 2002;18(3):593–609. https://doi.org/10.1016/S0749-0690(02)00025-3.
74. Lemon BW, Bengtson VL, Peterson JA. An exploration of the activity theory of aging: activity types and life satisfaction among in-movers to a retirement community1. J Gerontol. 1972;27(4):511–23. https://doi.org/10.1093/geronj/27.4.511.
75. Rowe JW, Kahn RL. Successful aging. Gerontologist. 1997;37(4):433–40. https://doi.org/10.1093/geront/37.4.433.
76. Herzog AR, Kahn RL, Morgan JN, Jackson JS, Antonucci TC. Age differences in productive activities. J Gerontol. 1989;44(4):S129–38. https://doi.org/10.1093/geronj/44.4.S129.
77. Marcum CS. Age differences in daily social activities. Res Aging. 2013;35(5):612–40. https://doi.org/10.1177/0164027512453468.
78. Verbrugge LM, Gruber-Baldini AL, Fozard JL. Age differences and age changes in activities: Baltimore Longitudinal Study of Aging. J Gerontol B Psychol Sci Soc Sci. 1996;51B(1):S30–41. https://doi.org/10.1093/geronb/51B.1.S30.
79. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychol Bull. 1955;52(4):281–302. https://doi.org/10.1037/h0040957.
80. Goto T, Kishimoto T, Fujiwara S, Shirayama Y, Ichikawa T. Social frailty as a predictor of all-cause mortality and functional disability: a systematic review and meta-analysis. Sci Rep. 2024;14(1):3410. https://doi.org/10.1038/s41598-024-53984-3.
81. Tsutsumimoto K, Doi T, Makizako H, Hotta R, Nakakubo S, Kim M, et al. Social frailty has a stronger impact on the onset of depressive symptoms than physical frailty or cognitive impairment: a 4-year follow-up longitudinal cohort study. J Am Med Dir Assoc. 2018;19(6):504–10. https://doi.org/10.1016/j.jamda.2018.02.008.
82. Makizako H, Shimada H, Doi T, Tsutsumimoto K, Hotta R, Nakakubo S, et al. Social frailty leads to the development of physical frailty among physically non-frail adults: a four-year follow-up longitudinal cohort study. Int J Environ Res Public Health. 2018;15(3):3. https://doi.org/10.3390/ijerph15030490.
83. Nagai K, Tamaki K, Kusunoki H, Wada Y, Tsuji S, Itoh M, et al. Physical frailty predicts the development of social frailty: a prospective cohort study. BMC Geriatr. 2020;20(1):1. https://doi.org/10.1186/s12877-020-01814-2.
84. Matison AP, Samtani S, Brodaty H, Sachdev PS, Reppermund S. How well do social frailty indices predict incident dementia in older adults? J Gerontol B Psychol Sci Soc Sci. 2025;80(10):gbaf148. https://doi.org/10.1093/geronb/gbaf148.
85. Coundouris SP, Grainger SA, Schweitzer D, Hubbard RE, Laakso E-L, Henry JD. How social frailty is operationalized matters: relationships with health and wellbeing in late adulthood. Int Psychogeriatr. 2025. https://doi.org/10.1016/j.inpsyc.2024.100032.
86. Andrew MK, Mitnitski AB, Rockwood K. Social vulnerability, frailty and mortality in elderly people. PLoS One. 2008;3(5):e2232. https://doi.org/10.1371/journal.pone.0002232.
87. Mah JC, Penwarden JL, Pott H, Theou O, Andrew MK. Social vulnerability indices: a scoping review. BMC Public Health. 2023;23(1):1. https://doi.org/10.1186/s12889-023-16097-6.
88. Bartolomucci A, Kane AE, Gaydosh L, Razzoli M, McCoy BM, Ehninger D, et al. Animal models relevant for geroscience: current trends and future perspectives in biomarkers, and measures of biological aging. J Gerontol A Biol Sci Med Sci. 2024;79(9):glae135. https://doi.org/10.1093/gerona/glae135.
89. Collinge CW, Razzoli M, Mansk R, McGonigle S, Lamming DW, Pacak CA, et al. The mouse Social Frailty Index (mSFI): a novel behavioral assessment for impaired social functioning in aging mice. Geroscience. 2024. https://doi.org/10.1007/s11357-024-01263-4.
90. Collinge CW, Razzoli M, Bartolomucci A. The Mouse Social Frailty Index (mSFI): a standardized protocol. Bio-Protoc. 2025;15(8):e5272. https://doi.org/10.21769/BioProtoc.5272.
91. Health and Retirement Study. (2024). *RAND HRS 2020 longitudinal file* (Version 2). Retrieved March 17, 2025.
92. Health and Retirement Study. (2022). *RAND HRS 2010 fat file* (Version 6A). Retrieved March 17, 2025.
93. Health and Retirement Study. (2021). *RAND HRS 2012 fat file* (Version 3A). Retrieved March 8, 2025.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.