JAMDA 27 (2026) 106161

JAMDA

journal homepage: www.jamda.com

Original Study

Comparison of 4 Frailty Measurements in Relation to Ageing, Physical Performance, and All-Cause Mortality

Lei Peng MSc [a,b], Klaus Hauer PhD [c], Bernd Holleczek PhD [d], Silvia Calderazzo PhD [e], Hermann Brenner PhD [a,f], Ben Schöttker PhD [a,f,*]

[a] Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany
[b] Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany
[c] Geriatric Center, Medical Faculty, Heidelberg University, Heidelberg, Germany
[d] Saarland Cancer Registry, Saarbrücken, Germany
[e] Division of Biostatistics, German Cancer Research Center, Heidelberg, Germany
[f] Network Aging Research, University of Heidelberg, Heidelberg, Germany

Keywords:
Frailty
aging
physical performance
mortality
cohort study

A B S T R A C T

Objectives: Frailty instruments predict mortality and functional decline, but their performance across sex and age groups has not been compared.
Design: This is a prospective population-based cohort study conducted between October 2014 and September 2016, with mortality follow-up through December 31, 2022.
Setting and Participants: This study was conducted in Saarland, Germany, and included 2000 community-dwelling adults aged 63 to 89 years with complete data on 4 frailty instruments.
Methods: We conducted a population-based cohort study of 2000 participants aged 63 to 89 years. The Clinical Frailty Scale (CFS), Fried Frailty Phenotype (FFP), Frailty Index (FI), and FRAIL Scale (FS) were compared with respect to their cross-sectional associations with poor physical performance (Short Physical Performance Battery score ≤6 points) and longitudinal associations with all-cause mortality in models adjusted for age, sex, education, and smoking status.
Results: The frailty prevalence was similar when assessed with the FFP (18.1%), FI (18.7%), FS (18.5%), and CFS (20.2%), and increased with age except for the FS. Frailty assessed by all instruments was significantly associated with poor physical performance and all-cause mortality. The FFP exhibited the strongest association with poor physical performance [odds ratio (OR), 9.19; 95% CI, 4.87-14.30], whereas the ORs of the other instruments had comparable OR point estimates of about 4. The FI yielded the strongest hazard ratio (HR) with mortality (HR, 3.83; 95% CI, 2.47-5.94), closely followed by the CFS and FS, whereas the FFP had the weakest association (HR, 2.62; 95% CI, 1.76-3.90). The FFP showed the best discrimination for poor physical performance across age and sex subgroups. For mortality, the FI showed the best discrimination overall, except among men, in whom the FS performed best.
Conclusions and Implications: This study showed that the performance of frailty instruments varies strongly by outcome of interest, sex, and age. This has implications for the choice of the frailty instrument to use in different clinical and research contexts.
© 2026 The Author(s). Published by Elsevier Inc. on behalf of Post-Acute and Long-Term Care Medical Association. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

The accurate measurement of frailty is vital for assessing vulnerability in older adults. Frailty is a multidimensional state of reduced resilience and increased vulnerability.[1,2] Prior research has firmly established the link between frailty and all-cause mortality. For example, Saum et al[3] demonstrated a dose-response relationship between the Frailty Index (FI) and 10-year mortality, and Li et al[4] reported that frailty accounted for 18.4% of male and 19.2% of female deaths before the age of 80 years. Frailty has also been strongly associated with impaired physical function, including reduced

Funding: The ESTHER study was supported in part by the Baden-Württemberg state Ministry of Science, Research and Arts (Stuttgart, Germany), the German Federal Ministry of Education and Research (Berlin, Germany), and the Saarland state Ministry of Health, Women, Social Affairs and Family.
* Address correspondence to Ben Schöttker, PhD, Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Im Neuenheimer Feld 581, Heidelberg, Germany 69120.
E-mail address: b.schoettker@dkfz.de (B. Schöttker).

https://doi.org/10.1016/j.jamda.2026.106161
1525-8610/© 2026 The Author(s). Published by Elsevier Inc. on behalf of Post-Acute and Long-Term Care Medical Association. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

mobility and muscle strength, which often precede disability and loss of independence.[5]

Over the past decade, studies have developed and refined frailty assessment scales, creating standardized, multidimensional frameworks for evaluation in older adults.[6,7] The most frequently used instruments include the Clinical Frailty Scale (CFS),[8] the Fried Frailty Phenotype (FFP),[9] the FI,[10] and the FRAIL Scale (FS).[11] Despite these advances, debate continues over which instrument is best suited for clinical and research purposes.[12,13] These instruments differ fundamentally in their conceptual focus, ranging from physical performance-based (FFP) and deficit accumulation-based approaches (FI) to clinician-rated (CFS) and self-reported screening instruments (FS). Their relative suitability for different clinical outcomes and research purposes is unknown.

The objective of this study is to conduct a comparative analysis of these 4 frailty measures in a single cohort. An effective frailty tool should demonstrate an increasing prevalence with advancing age, show a strong association with poor physical performance, and reliably predict all-cause mortality. Although numerous studies have compared at least 2 frailty measures using mortality as an outcome,[14-28] none has provided a sex- and age-specific comparison and also examined associations with poor physical function. This study is the first to offer such a comprehensive evaluation of the 4 most widely used frailty tools across different sex and age groups.

Methods

Study Design and Participant Selection

Our analyses use data from the ESTHER study (German name: *Epidemiologische Studie zu Chancen der Verhütung Früherkennung und optimierten Therapie chronischer Erkrankungen in der älteren Bevölkerung*), a comprehensive population-based cohort study conducted in Germany.[3] Detailed descriptions of the study population and design have been shown in Supplementary Text 1.

The only inclusion criterion was participation in the third home visit (conducted from October 2014 to September 2016), to which all ESTHER participants were invited. This time point served as the baseline for the present analysis. Of 2342 home visit participants, 342 were excluded due to age outside 63 to 89 years or completely missing frailty data, leaving 2000 for the final analysis (Supplementary Figure 1).

The ESTHER cohort received approval from the ethics committee of the Heidelberg Medical Faculty of Heidelberg University and the State Medical Board of Saarland, Germany. All participants provided written informed consent.

Frailty Measurements and Short Physical Performance Battery

Supplementary Text 2 describes the assessment methods of the frailty instruments and the Short Physical Performance Battery (SPPB). Supplementary Table 1 details the single items used to calculate their scores/indices. Supplementary Table 2 gives the rationale for the cutoffs chosen for the FI to define prefrailty and frailty. Supplementary Figure 2 shows the cutoffs used to define prefrailty and frailty for all 4 frailty instruments and for the SPPB to define poor, suboptimal, and optimal physical performance.

Mortality Assessment

Deaths occurring between the home visit and December 31, 2022, were ascertained through inquiries at residents' registration offices, which resulted in a mortality follow-up of 5.2 to 7.3 years (median, 6.2 years). Causes of death were classified using the International Classification of Diseases, 10th Revision (ICD-10) codes, with cardiovascular disease defined as I00 to I99 and cancer defined as C00 to C97. The vital status ascertainment was complete for all participants.

Covariates

Covariates considered were age, sex, education, smoking status, body mass index (BMI), and the Cumulative Illness Rating Scale-Geriatric (CIRS-G) score. Their assessment methods are outlined in Supplementary Text 3.

Statistical Analyses

Weighted kappa coefficients were computed to assess the inter-instrument agreement between frailty instruments to define prefrailty and frailty. We used the interpretation of kappa coefficients of Checa-Lopez et al,[14] with <0.01 indicating no agreement, 0.01 to 0.20 indicating poor agreement, 0.21 to 0.40 indicating fair agreement, 0.41 to 0.60 indicating moderate agreement, and 0.61 to 0.80 indicating substantial agreement. To assess the association of age with (pre-)frailty, unadjusted ordinal logistic regression models were used with frailty status (nonfrail, prefrail, and frail) as the dependent variable and age as a continuous independent variable.

Logistic regression models, adjusted for age (continuous), sex, education (categorical: $\leq$9, 10-11, and $\geq$12 years of school attendance), and smoking status (categorical: never, former, and current smoker), were used to estimate odds ratios (ORs) and 95% CIs for the associations of prefrailty and frailty (assessed using the 4 instruments) with poor physical performance. Similarly, Cox proportional hazards regression models, adjusted for the same covariates, were used to estimate hazard ratios (HRs) and 95% CIs for the associations of prefrailty and frailty with all-cause mortality. The proportional hazards assumption was assessed using Schoenfeld residuals.[29] In sensitivity analyses, the models were extended by additional adjustment for BMI (categorical: $<$21.5, 21.5 to $<$25, 25 to $<$30, 30 to $<$35, and $\geq$35 kg/m^2) and CIRS-G score (continuous). These variables were not part of the main model because low BMI and morbidity can be highly correlated with frailty and are also items of some frailty instruments.

The predictive performance of the frailty measures was evaluated using discrimination metrics. For logistic regression models,[30] the area under the receiver operating characteristic curve (AUC) was calculated, and for Cox proportional hazards models, Harrell's C-statistic was used.[31]

Supplementary Text 4 provides details on statistical software used and the handling of missing data.

Results

Baseline Characteristics of the Study Population

The study included 2000 participants with a mean age $\pm$ SD of 74.3 $\pm$ 6.1 years (Table 1). Slightly more than one-half of the participants (53.6%) were women. Almost two-thirds of the study participants had low education of $\leq$9 years (64.4%), whereas 18.6% had attended school for 10 to 11 years, and 17.0% for $\geq$12 years. Most were never smokers (55.3%), whereas the others were either former (38.7%) or current smokers (6.0%). In terms of BMI, nearly one-half of the participants (44.3%) were classified as overweight (25 to $<$30 kg/m^2), whereas 23.9% met the criteria for obesity ($\geq$30 kg/m^2). The burden of comorbidity, as measured by the CIRS-G, showed a mean score $\pm$ SD of 7.49 $\pm$ 5.54, and Supplementary Table 3 shows the detailed rating in all 14 organ systems.

The proportion of study participants classified as prefrail or frail was generally comparable across the 4 frailty measures with

Table 1
Baseline Characteristics of the Study Population (N = 2000)

Variables	Data
Age, y	74.3 ± 6.1
Female	1072 (53.6)
Education, y*	
≤9	1289 (64.4)
10–11	371 (18.6)
≥12	340 (17.0)
Smoking status*	
Never smoker	1107 (55.3)
Former smoker	773 (38.7)
Current smoker	120 (6.0)
BMI, kg/m^2*	
<21.5	132 (6.6)
21.5 to <25	503 (25.2)
25 to <30	887 (44.3)
30 to <35	374 (18.7)
≥35	104 (5.2)
Total CIRS-G morbidity score*	7.49 ± 5.54
FFP*	
Nonfrail	390 (19.5)
Prefrail	1248 (62.4)
Frail	362 (18.1)
CFS	
Nonfrail	438 (21.9)
Prefrail	1159 (57.9)
Frail	403 (20.2)
FI*	
Nonfrail	399 (19.9)
Prefrail	1227 (61.4)
Frail	374 (18.7)
FS*	
Nonfrail	712 (35.6)
Prefrail	919 (45.9)
Frail	369 (18.5)
SPPB*	
Optimal	569 (28.5)
Suboptimal	1033 (51.6)
Poor	398 (19.9)

Data are presented as mean ± SD or n (%).
*Missing values for education (1.3%), smoking status (1.9%), BMI (1.7%), the 14 somatic organ systems of the CIRS-G (0.1-0.5%) (Supplementary Table 2), and items from the FI, FFP, FS, and SPPB (0.1-8.2%) (Supplementary Table 1) were imputed using random forest implemented with the missRanger package in R (version 2.6.0; R Foundation for Statistical Computing, Vienna, Austria).

approximately 60% classified as prefrail and 20% classified as frail. As assessed by the SPPB, 28.5% of participants had optimal, 51.6% had suboptimal, and 19.9% had poor physical performance.

Concordance of Frailty Measures

The agreement between the 4 frailty measures in classifying individuals as prefrail and frail ranged from fair to moderate (Figure 1). Moderate agreement was noted in the total sample between the CFS and FI (kappa: 0.47), whereas the agreement of the comparisons between other frailty measures was fair (kappa values: 0.26-0.40). The agreement of the instruments did not differ substantially by sex. However, the kappa values were generally higher among the older (75-89 years) than the younger (63-74 years) subgroup.

Increasing Frailty Prevalence With Age

As illustrated in Figure 2, the proportion of frail individuals increased progressively with age in the total population for all frailty instruments ($P < .001$ in the ordinal logistic regression model), with exception of the FS ($P > .05$). The age trend was clearest for the FI and FFP. This trend was consistently observed across sex- and age-specific groups.

Baseline Association Between Frailty Scales and Physical Performance

Every fifth study participant had poor physical performance (n = 398, 19.9%). In the total study population, frailty was significantly associated with poor physical performance across all frailty instruments, with frailty defined by the FFP showing the strongest association (OR, 9.19), followed by the other 3 instruments, which were each associated with approximately 4-fold higher odds of poor physical performance (Table 2). Prefrail individuals had >4-fold odds for poor physical performance when assessed by the FFP and approximately 2-fold odds when assessed by the other frailty instruments.

Stratified analyses revealed notable differences in the strength of associations by sex and age (Table 2). However, across all subgroups, the FFP consistently demonstrated the strongest association with poor physical performance; the strength of the association was also consistent with the exception of a slightly weaker association among younger study participants (aged 63-74 years). In women, frailty classified by the FI showed the second strongest association, whereas in men, the FS ranked second. Among younger individuals, frailty classified by the FS had the second strongest association, whereas in older adults, the FI showed the second strongest association.

The observed ranking of model discrimination based on AUC for poor physical performance was FFP > FS > FI > CFS in the total population, FFP > FI > FS > CFS among females, FFP > FS > FI > CFS among males, FFP > FS > FI > CFS among younger individuals, and FFP > FI > FS > CFS among older individuals (Table 2).

Longitudinal Association of Frailty Status and All-Cause Mortality

During a median follow-up of 6.2 years, 368 of the 2000 participants died. Cardiovascular disease and cancer were the leading causes of death, accounting for 30.4% (n = 112) and 24.2% (n = 89) of all deaths, respectively. In adjusted Cox proportional hazard regression models, frailty was significantly associated with increased all-cause mortality across all frailty measures and all populations (Figure 3). The Schoenfeld residual test indicated no violation of the proportional hazards assumption (all $P > .05$). Among all participants, frailty was associated with a 2.6- to 3.8-fold increased mortality across the different frailty measures, with the FI showing the strongest association. Prefrailty was associated with a 1.3- to 1.7-fold increased risk, with again the FI showing the strongest association. Only the association of prefrailty classified by the CFS was not statistically significant.

Frailty classified by the FI showed the highest mortality among women, younger, and older individuals (Figure 3). Among males, the FI ranked second after the FS. The FS was also the only frailty instrument with a stronger association of frailty with mortality among males than females, whereas it was vice versa for all other instruments. The frailty instruments had comparable associations with mortality in the younger and older subgroup with exception of the FFP, which showed a stronger association of frailty with mortality among younger compared with older individuals.

The observed ranking of model discrimination based on C-statistic for all-cause mortality was FI > CFS > FS > FFP in the total population, FI > FFP > CFS > FS among females, FS > FI > FFP > CFS among males, FI > FFP > CFS > FS among younger individuals, and FI > FS > CFS > FFP among older individuals (Figure 3).

Sensitivity Analyses

In sensitivity analyses, additionally adjusting for BMI and CIRS-G, the results were highly consistent with the main analyses for both poor physical performance (Supplementary Table 4) and all-cause

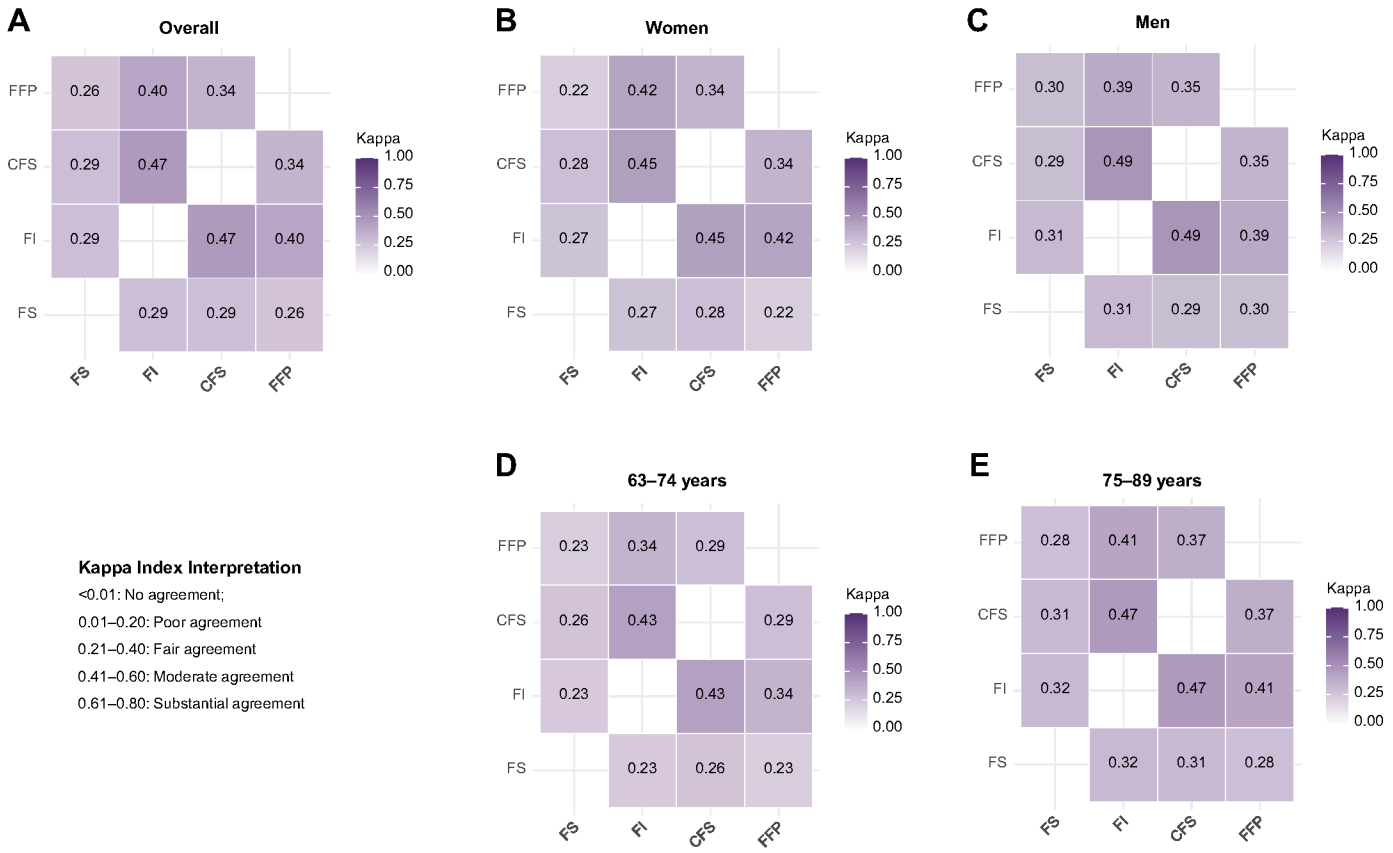

Fig. 1. Interinstrument agreement between the 4 frailty instruments in the classification of nonfrail, prefrail, and frail individuals: (A) overall, (B) women, (C) men, (D) aged 63 to 74 years, and (E) aged 75 to 89 years.

mortality (Supplementary Figure 3). Although all HRs and ORs mostly decreased and AUCs increased in all models, the top ranked frailty instrument always remained the same as in the main analysis.

Discussion

This study provides a comprehensive comparison of 4 widely used frailty measures (FFP, FI, CFS, and FS) in relation to aging, physical performance, and all-cause mortality in a large population-based cohort of older adults. Our findings underscore the differences among these measures regarding their predictive value for the different outcomes, and across sex and age groups.

Although the proportion of patients classified as prefrail and frail was similar across the FFP, FI, FS, and CFS measures, their inter-instrument agreement was fair to moderate. This reflects the conceptual differences in frailty measurements. Each test has a unique methodologic feature. The FS relies only on self-reported information, the CFS relies on clinical judgment of physicians, the FI only includes age-related deficits, and the FFP is the only instrument that includes physical assessments (grip strength and gait speed). The only instruments with a moderate agreement were the CFS and FI. This aligns with prior research suggesting that frailty measures capturing cumulative deficits (eg, FI) tend to correlate more strongly with clinical frailty assessments (eg, CFS) because for both, age-related diseases play an important role.[32] In contrast, the weak agreement of the FS with other measures supports previous concerns regarding its limited overlap with more comprehensive frailty assessment methods.[33] In the nursing home setting, the FI identified more than twice as many frail individuals as the FS (94% vs 41%).[33]

Frailty is increasingly recognized as a clinical manifestation of biological aging, characterized by multisystem dysregulation.[9] Within the geroscience framework, frailty reflects underlying processes such as chronic inflammation, mitochondrial dysfunction, and cellular senescence.[34] The extent to which frailty instruments capture these mechanisms varies.[35] According to our results, the FFP and FI best reflected aging processes, because the prefrail and frailty prevalences estimated by these instruments increased consistently with age. This pattern was less evident for the FS and CFS, which may be explained by their more subjective nature. Self-reported or clinician-rated assessments may correlate less closely with biological age than the FFP, which is partly based on objective measurements of grip strength and gait speed. The high correlation of FI with age is not surprising given that FI components are commonly selected based on their association with aging and health decline.[36]

All 4 frailty measures were strongly associated with poor physical performance, as assessed by the SPPB, with the FFP demonstrating the strongest association. This finding is expected given that the FFP directly incorporates components of physical function, which are also central to the SPPB.

Interestingly, the FI, which is based on the accumulation of deficits across multiple domains, showed a weaker association with poor physical performance compared with the FFP. This may indicate that the FI captures broader aspects of health deterioration beyond musculoskeletal decline, including cognitive and metabolic impairments, which are not directly measured by the SPPB. The FS and CFS also demonstrated weaker associations, which may reflect the fact that they do not include physical performance items in their frailty assessment.

Stratified analyses revealed notable sex and age differences in the strength of associations between frailty and poor physical performance. Among women, the FI showed the strongest association after the FFP, implying that cumulative health deficits play a more

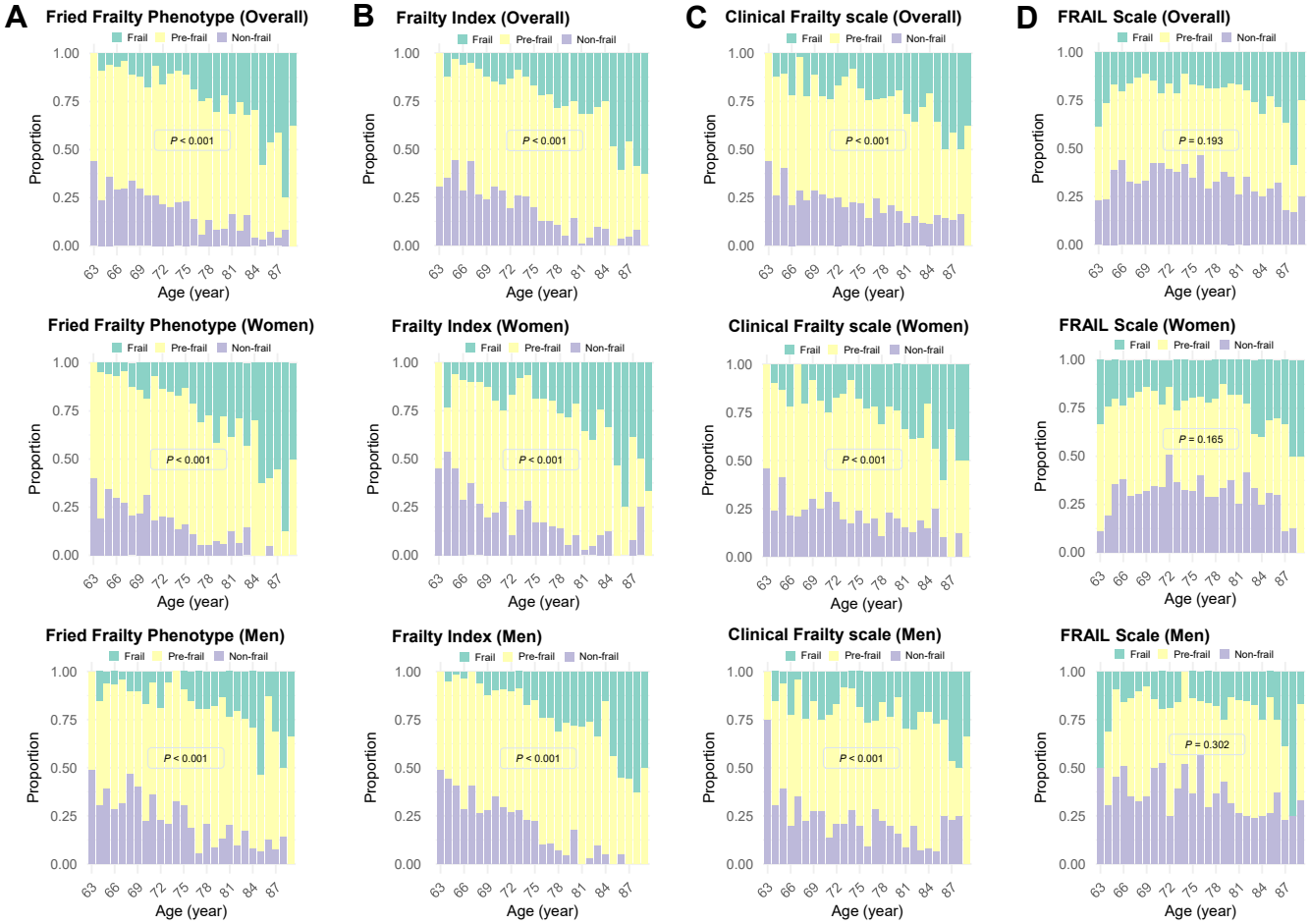

Fig. 2. Age trend of frailty prevalence according to the 4 frailty measures in the total population and specific subgroups: (A) Fried Frailty Phenotype, (B) Frailty Index, (C) Clinical Frailty Scale, and (D) FRAIL Scale. The *P* values were derived from ordinal logistic regression models to test the association between the study participant's age (continuous) and frailty status (nonfrail/prefrail/frail).

prominent role in female frailty-related functional decline. In contrast, the FS exhibited the second strongest association after the FFP, suggesting that self-reported fatigue and weight loss may be particularly relevant predictors of functional decline in men.[37,38] These findings align with prior research showing sex differences, with men more often physically frail and women accumulating more chronic disease–related deficits.[39]

Age-stratified analyses showed that the ranking of frailty measures in modeling poor physical performance differed between relatively younger (63-74 years) and older (75-89 years) individuals in the study population, with the FS ranking second among younger participants and the FI ranking second among older participants. This pattern could be explained by the fact that the FI, which heavily depends on disease burden, may show stronger associations with poor physical performance in older compared with younger individuals, because older individuals tend to have a higher prevalence and greater severity of chronic diseases.

The FI had the strongest association with mortality in the total population and all subgroups except men, which can be explained by its construct as a cumulative deficit model based on age-related items. This close link to aging makes it a robust instrument for all-cause mortality prediction. Surprisingly, the FS outperformed the FI among men. All individual items of the FS are strong predictors of mortality; possibly, some of them predict mortality better among men than women. For all other frailty instruments, the association of frailty with mortality was stronger among women than men, which is consistent with previous literature.[40-43] However, the biological reason for this is unknown.

Interestingly, despite its strong association with physical performance, the FFP had the weakest association with mortality in the total study population. This may be due to its emphasis on short-term functional decline rather than broader systemic health deterioration.[44] Nevertheless, the FFP ranked second in women for mortality prediction, highlighting again sex differences in frailty-health outcome associations and suggesting that, in women, functional decline is more prevalent and occurs earlier in life, with lower baseline physical reserves contributing to reduced resilience against loss of function.

The frailty instruments had comparable associations with mortality in relatively young (63-74 years) and old (75-89 years) individuals with exception of the FFP, which showed a stronger association of frailty with mortality among younger than older individuals. This likely reflects that declines in physical performance are less common earlier in life, making them more potent indicators of mortality risk when present.

Taken together, these findings carry important clinical and public health implications, which are outlined in Supplementary Text 5.

This study has several strengths, including the use of a large, well-characterized, population-based cohort that supports robust statistical analysis. The comprehensive assessment of 4 frailty measures alongside an objective physical performance test and complete mortality assessment allowed a robust analysis of high internal

Table 2
Baseline Associations of Frailty Status Assessed With Different Instruments With Poor Physical Performance

Frailty Status	Total Study Population, n_{total} = 2000	Women, n_{total} = 1072	Men, n_{total} = 928	63–74 y, n_{total} = 978	75–89 y, n_{total} = 1022
	n_{poor} = 398 (19.9%)	n_{poor} = 225 (21.0%)	n_{poor} = 173 (18.6%)	n_{poor} = 118 (12.1%)	n_{poor} = 280 (27.4%)
FFP					
Nonfrail	Ref.	Ref.	Ref.	Ref.	Ref.
Prefrail	**4.80 (2.94–7.82)**	**3.71 (1.74–7.90)**	**5.73 (3.02–10.86)**	**4.39 (2.45–7.85)**	**6.58 (2.51–17.28)**
Frail	**9.19 (4.87–14.30)**	**8.16 (2.12–16.42)**	**7.94 (2.99–14.07)**	**5.86 (2.26–10.06)**	**8.94 (2.91–15.20)**
AUC (95% CI)*	**0.804 (0.780–0.828)**	**0.812 (0.779–0.845)**	**0.791 (0.755–0.827)**	**0.755 (0.709–0.800)**	**0.794 (0.763–0.825)**
CFS					
Nonfrail	Ref.	Ref.	Ref.	Ref.	Ref.
Prefrail	**1.87 (1.18–2.96)**	**2.20 (1.17–4.15)**	1.48 (0.75–2.90)	1.83 (0.86–3.89)	**1.84 (1.01–3.33)**
Frail	**4.06 (2.32–7.11)**	**4.68 (2.14–10.22)**	**3.34 (1.50–7.47)**	**5.74 (2.39–13.76)**	**3.11 (1.49–6.45)**
AUC (95% CI)*	**0.745 (0.717–0.773)**	**0.752 (0.716–0.788)**	**0.737 (0.694–0.780)**	**0.671 (0.620–0.721)**	**0.749 (0.716–0.788)**
FI					
Nonfrail	Ref.	Ref.	Ref.	Ref.	Ref.
Prefrail	**1.78 (1.17–2.71)**	1.35 (0.76–2.41)	**2.30 (1.24–4.28)**	**1.90 (1.12–3.23)**	1.42 (0.70–2.88)
Frail	**3.77 (2.15–6.64)**	**2.83 (1.29–6.18)**	**5.00 (2.18–11.48)**	**4.16 (1.68–10.27)**	**3.00 (1.33–6.79)**
AUC (95% CI)*	**0.762 (0.734–0.789)**	**0.775 (0.740–0.809)**	**0.746 (0.703–0.789)**	**0.684 (0.630–0.737)**	**0.770 (0.737–0.802)**
FS					
Nonfrail	Ref.	Ref.	Ref.	Ref.	Ref.
Prefrail	**2.02 (1.39–2.92)**	**2.33 (1.39–3.91)**	**1.83 (1.06–3.15)**	1.62 (0.85–3.08)	**2.29 (1.44–3.67)**
Frail	**3.97 (1.83–8.64)**	**6.15 (2.33–16.27)**	1.74 (0.41–7.45)	**4.54 (1.37–14.91)**	**4.02 (1.42–11.31)**
AUC (95% CI)*	**0.764 (0.738–0.790)**	**0.760 (0.724–0.796)**	**0.774 (0.736–0.813)**	**0.714 (0.667–0.760)**	**0.758 (0.725–0.791)**

Ref., reference.
Values are OR (95% CI) or as otherwise indicated. Statistically significant ($P < .05$) values are presented in boldface font.
n_{poor} refers to the number of participants with poor physical performance (SPPB score $\leq$ 6 points).
*Poor physical performance was defined by $\leq$6 points for the SPPB. All logistic regression models included the covariates of age (years), sex, education, and smoking status.

validity. Moreover, the comparison of 4 widely used frailty measures provides practical guidance for clinicians and researchers.

However, several limitations warrant consideration. Confounding cannot be excluded. We used a limited number of covariates in the models because it was the aim of this study to compare the predictive value of frailty instruments and not to precisely quantify causal associations of frailty with poor physical performance and mortality. Our study population was healthier than the general population

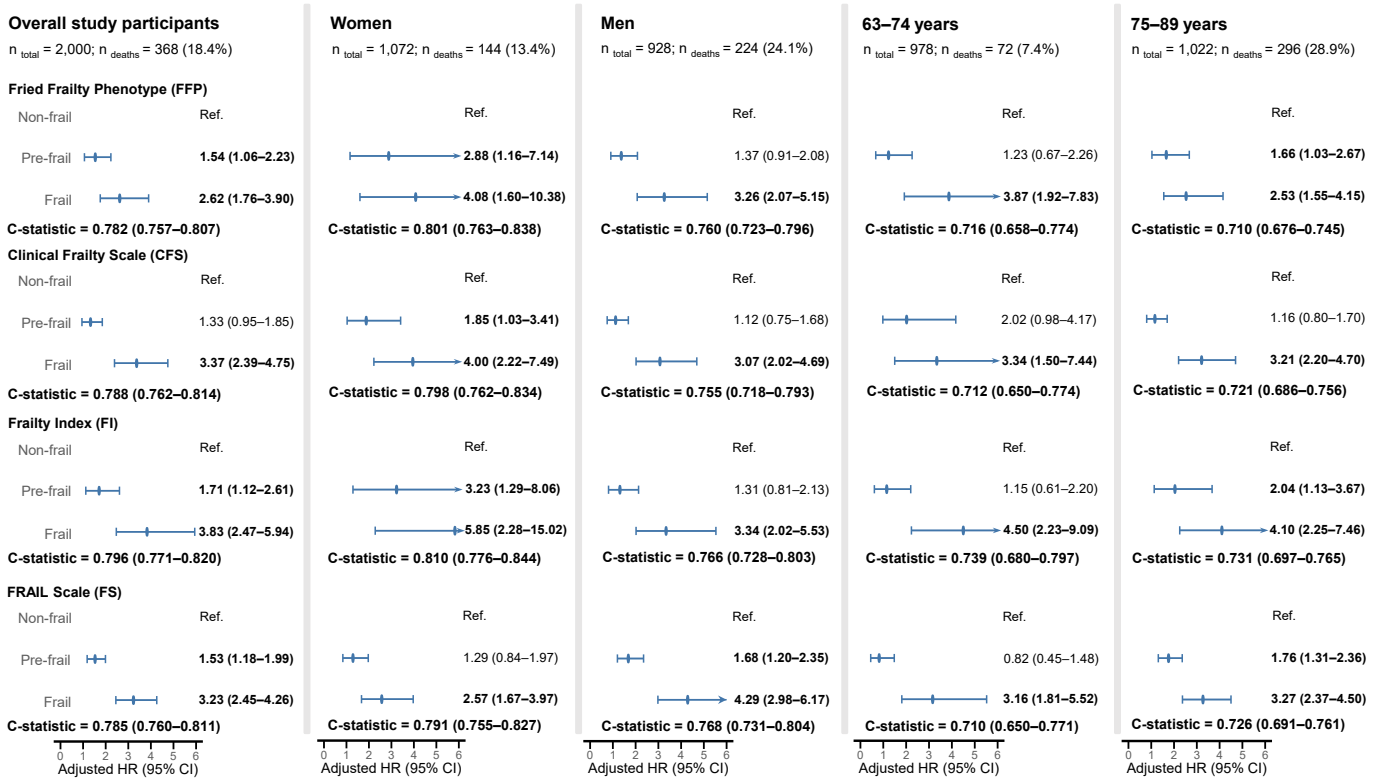

Fig. 3. Longitudinal association of the 4 frailty measures with all-cause mortality in the overall population and specific subgroups. All Cox regression models were adjusted for age (years), sex, education, and smoking status. Statistically significant results ($P < .05$) are highlighted in boldface font. Ref., reference.

because institutionalized adults and less healthy nonparticipants were excluded. This is unlikely to materially affect the relative effect estimates (ORs and HRs) shown in this study. However, it may have led to an underestimation of (pre-)frailty prevalences, and these estimates should be interpreted with caution when extrapolating to broader populations. Therefore, replication of these findings in independent cohorts, particularly those including individuals with higher baseline frailty levels or institutionalized populations, would further strengthen the generalizability of our findings. Finally, the results on physical performance were cross sectional, and future studies should adopt a longitudinal design.

Conclusions and Implications

In this population-based cohort, 4 commonly used frailty instruments were all associated with poor physical performance and all-cause mortality, but their agreement and predictive performance varied substantially by outcome, sex, and age. These findings indicate that frailty instruments are not interchangeable and that their selection should be guided by the intended clinical or research purpose. Instruments incorporating physical performance may be preferable for identifying functional impairment, whereas deficit accumulation approaches appear more suitable for mortality risk stratification. Future studies should explicitly consider outcome- and subgroup-specific performance when selecting or comparing frailty measures.

Conflict of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors thank all study participants and general practitioners of the ESTHER study, and the study staff for their valuable contributions. The authors also thank Dr Utz Benscheid (DKFZ) for his support in preparing and collecting the ESTHER data.

Supplementary Material

Supplementary data related to this article can be found online at https://doi.org/10.1016/j.jamda.2026.106161.

References

1. Hoogendijk EO, Afilalo J, Ensrud KE, et al. Frailty: implications for clinical practice and public health. *Lancet*. 2019;394:1365–1375.
2. Robinson TN, Walston JD, Brummel NE, et al. Frailty for surgeons: review of a National Institute on Aging conference on frailty for specialists. *J Am Coll Surg*. 2015;221:1083–1092.
3. Saum KU, Dieffenbach AK, Müller H, et al. Frailty prevalence and 10-year survival in community-dwelling older adults: results from the ESTHER cohort study. *Eur J Epidemiol*. 2014;29:171–179.
4. Li X, Ploner A, Karlsson IK, et al. The frailty index is a predictor of cause-specific mortality independent of familial effects from midlife onwards: a large cohort study. *BMC Med*. 2019;17:94.
5. 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.
6. Sison SDM, Shi SM, Kim KM, et al. A crosswalk of commonly used frailty scales. *J Am Geriatr Soc*. 2023;71:3189–3198.
7. Dent E, Kowal P, Hoogendijk EO. Frailty measurement in research and clinical practice: a review. *Eur J Intern Med*. 2016;31:3–10.
8. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005;173:489–495.
9. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56:M146–M156.
10. Rockwood K, Mitnitski A. Frailty defined by deficit accumulation and geriatric medicine defined by frailty. *Clin Geriatr Med*. 2011;27:17–26.
11. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. *J Nutr Health Aging*. 2012;16:601–608.
12. Zhang XM, Cao S, Gao M, et al. The prevalence of social frailty among older adults: a systematic review and meta-analysis. *J Am Med Dir Assoc*. 2023;24:29–37.e29.
13. Pavasini R, Guralnik J, Brown JC, et al. Short Physical Performance Battery and all-cause mortality: systematic review and meta-analysis. *BMC Med*. 2016;14:215.
14. Checa-Lopez M, Rodriguez-Laso A, Carnicero JA, et al. Differential utility of various frailty diagnostic tools in non-geriatric hospital departments of several countries: a longitudinal study. *Eur J Clin Invest*. 2023;53:e13979.
15. Shang N, Liu H, Wang N, et al. Comparison of three frailty screening instruments for prediction of adverse outcomes among older adults in the emergency department. *Geriatr Gerontol Int*. 2022;22:851–856.
16. Oviedo-Briones M, Rodríguez-Laso Á, Carnicero JA, et al. The ability of eight frailty instruments to identify adverse outcomes across different settings: the FRAILTOOLS project. *J Cachexia Sarcopenia Muscle*. 2022;13:1487–1501.
17. Zhang D, Tang W, Dou LY, et al. Four different frailty models predict health outcomes in older patients with stable chronic obstructive pulmonary disease. *BMC Geriatr*. 2022;22:57.
18. Arteaga AS, Aguilar LT, González JT, et al. Impact of frailty in surgical emergencies. A comparison of four frailty scales. *Eur J Trauma Emerg Surg*. 2021;47:1613–1619.
19. Li JJ, Jiang S, Zhu ML, et al. Comparison of three frailty scales for prediction of adverse outcomes among older adults: a prospective cohort study. *J Nutr Health Aging*. 2021;25:419–424.
20. Op Het Veld LPM, Beurskens A, de Vet HCW, et al. The ability of four frailty screening instruments to predict mortality, hospitalization and dependency in (instrumental) activities of daily living. *Eur J Ageing*. 2019;16:387–394.
21. Chong E, Ho E, Baldevarona-Llego J, et al. Frailty in hospitalized older adults: comparing different frailty measures in predicting short- and long-term patient outcomes. *J Am Med Dir Assoc*. 2018;19:450–457.e453.
22. Belga S, Majumdar SR, Kahlon S, et al. Comparing three different measures of frailty in medical inpatients: multicenter prospective cohort study examining 30-day risk of readmission or death. *J Hosp Med*. 2016;11:556–562.
23. Widagdo IS, Pratt N, Russell M, et al. Predictive performance of four frailty measures in an older Australian population. *Age Ageing*. 2015;44:967–972.
24. Ritt M, Schwarz C, Kronawitter V, et al. Analysis of Rockwood et Al's Clinical Frailty Scale and Fried et Al's Frailty Phenotype as predictors of mortality and other clinical outcomes in older patients who were admitted to a geriatric ward. *J Nutr Health Aging*. 2015;19:1043–1048.
25. Malmstrom TK, Miller DK, Morley JE. A comparison of four frailty models. *J Am Geriatr Soc*. 2014;62:721–726.
26. Theou O, Brothers TD, Mitnitski A, et al. Operationalization of frailty using eight commonly used scales and comparison of their ability to predict all-cause mortality. *J Am Geriatr Soc*. 2013;61:1537–1551.
27. Ravindrarajah R, Lee DM, Pye SR, et al. The ability of three different models of frailty to predict all-cause mortality: results from the European Male Aging Study (EMAS). *Arch Gerontol Geriatr*. 2013;57:360–368.
28. Woo J, Leung J, Morley JE. Comparison of frailty indicators based on clinical phenotype and the multiple deficit approach in predicting mortality and physical limitation. *J Am Geriatr Soc*. 2012;60:1478–1486.
29. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81:515–526.
30. Hanley JA, McNeil BJ. The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology*. 1982;143:29–36.
31. Antolini L, Nam BH, D'Agostino RB. Inference on correlated discrimination measures in survival analysis: a nonparametric approach. *Commun Stat-Theor M*. 2004;33:2117–2135.
32. Moreno-Ariño M, Torrente Jiménez I, Cartanyà Gutiérrez A, et al. Assessing the strengths and weaknesses of the clinical frailty scale through correlation with a frailty index. *Aging Clin Exp Res*. 2020;32:2225–2232.
33. Oviedo-Briones M, Laso ÁR, Carnicero JA, et al. A comparison of frailty assessment instruments in different clinical and social care settings: the frailtools project. *J Am Med Dir Assoc*. 2021;22:607.e607.
34. Johnson AA, Shokhirev MN, Shoshitaishvili B. Revamping the evolutionary theories of aging. *Ageing Res Rev*. 2019;55:100947.
35. O'Caoimh R, Sezgin D, O'Donovan MR, et al. Prevalence of frailty in 62 countries across the world: a systematic review and meta-analysis of population-level studies. *Age Ageing*. 2021;50:96–104.
36. Dzando G, Ward PR, Mwanri L, et al. Determining frailty index thresholds for older people across multiple countries in Sub-Saharan Africa. *Commun Med (Lond)*. 2025;5:237.
37. Lindgren LH, Hammer NM, Flurey CA, et al. Gender differences in illness acceptance and coping strategies among patients with rheumatoid arthritis, psoriatic arthritis, and axial spondyloarthritis: a cross-sectional survey study. *Rheumatol Int*. 2025;45:52.
38. Tirosh A, de Souza RJ, Sacks F, et al. Sex differences in the effects of weight loss diets on bone mineral density and body composition: POUNDS LOST trial. *J Clin Endocrinol Metab*. 2015;100:2463–2471.

39. Sharma M, Anand A, Chattopadhyay A, et al. Gender differentials in cognitive frailty among older adults in India: a multivariate decomposition approach. *Sci Rep*. 2024;14:24597.
40. Walker E, Winter R, Hodgson LE. Frailty in older adults admitted to hospital: a six-year dual-centre retrospective study of over 53 000 clinical frailty scale assessments. *Age Ageing*. 2025;54:afaf137.
41. Nicaise EH, Palmateer G, Schmeusser BN, et al. Differences in preoperative frailty assessment of surgical candidates by sex, age, and race. *Surg Open Sci*. 2024;19:172–177.
42. Hoogendijk EO, Stenholm S, Ferrucci L, et al. Operationalization of a frailty index among older adults in the InCHIANTI study: predictive ability for all-cause and cardiovascular disease mortality. *Aging Clin Exp Res*. 2020;32:1025–1034.
43. Kaskirbayeva D, West R, Jaafari H, et al. Progression of frailty as measured by a cumulative deficit index: a systematic review. *Ageing Res Rev*. 2023;84:101789.
44. Zucchelli A, Vetrano DL, Marengoni A, et al. Frailty predicts short-term survival even in older adults without multimorbidity. *Eur J Intern Med*. 2018;56:53–56.