ORIGINAL ARTICLE

Journal of Cachexia, Sarcopenia and Muscle 2022; **13**: 1487–1501
Published online 15 April 2022 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/jcsm.12990

The ability of eight frailty instruments to identify adverse outcomes across different settings: the FRAILTOOLS project

Myriam Oviedo-Briones[1,2], Ángel Rodríguez-Laso[3], José Antonio Carnicero[1], Barbara Gryglewska[4], Alan J. Sinclair[5], Francesco Landi[6], Bruno Vellas[7], Fernando Rodríguez Artalejo[8], Marta Checa-López[9] & Leocadio Rodriguez-Mañas[3,9]*

[1]*Fundación para la Investigación Biomédica del Hospital Universitario de Getafe, Madrid, Spain;* [2]*Facultad de Medicina, Universidad de las Américas, Quito, Ecuador;* [3]*CIBERFES: CIBER (Centers of the Network of Biomedical Research) thematic area of Frailty and Healthy Ageing, Instituto de Salud Carlos III, Madrid, Spain;* [4]*Department of Internal Medicine and Gerontology, Jagiellonian University Medical Collegium Medicum, Cracow, Poland;* [5]*Diabetes Frail and King's College, London, UK;* [6]*Hospital Centro Medicina dell'Invecchiamento, Università Cattolica del Sacro Cuore, Rome, Italy;* [7]*Gerontopole, Centre Hospitalier Universitaire de Toulouse, Toulouse, France;* [8]*Preventive Medicine and Public Health School of Medicine, Universidad Autónoma de Madrid, Madrid, Spain;* [9]*Jefe de Servicio de Geriatría, Hospital Universitario de Getafe, Madrid, Spain*

Abstract

Background To compare the performance of eight frailty instruments to identify relevant adverse outcomes for older people across different settings over a 12 month follow-up.

Methods Observational longitudinal prospective study of people aged 75 + years enrolled in different settings (acute geriatric wards, geriatric clinic, primary care clinics, and nursing homes) across five European cities. Frailty was assessed using the following: Frailty Phenotype, SHARE-FI, 5-item Frailty Trait Scale (FTS-5), 3-item FTS (FTS-3), FRAIL scale, 35-item Frailty Index (FI-35), Gérontopôle Frailty Screening Tool, and Clinical Frailty Scale. Adverse outcomes ascertained at follow-up were as follows: falls, hospitalization, increase in limitation in basic (BADL) and instrumental activities of daily living (IADL), and mortality. Sensitivity, specificity, and capacity to predict adverse outcomes in logistic regressions by each instrument above age, gender, and multimorbidity were calculated.

Results A total of 996 individuals were followed (mean age 82.2 SD 5.5 years, 61.3% female). In geriatric wards, the FI-35 (69.1%) and the FTS-5 (67.9%) showed good sensitivity to predict death and good specificity to predict BADL worsening (70.3% and 69.8%, respectively). The FI-35 also showed good sensitivity to predict BADL worsening (74.6%). In nursing homes, the FI-35 and the FTSs predicted mortality and BADL worsening with a sensitivity > 73.9%. In geriatric clinic, the FI-35, the FTS-5, and the FRAIL scale obtained specificities > 85% to predict BADL worsening. No instrument achieved high enough sensitivity nor specificity in primary care. All the instruments predict the risk for all the outcomes in the whole sample after adjusting for age, gender, and multimorbidity. The associations of these instruments that remained significant by setting were for BADL worsening in geriatric wards [FI-35 OR = 5.94 (2.69–13.14), FTS-3 = 3.87 (1.76–8.48)], nursing homes [FI-35 = 4.88 (1.54–15.44), FTS-5 = 3.20 (1.61–6.38), FTS-3 = 2.31 (1.27–4.21), FRAIL scale = 1.91 (1.05–3.48)], and geriatric clinic [FRAIL scale = 4.48 (1.73–11.58), FI-35 = 3.30 (1.55–7.00)]; for IADL worsening in primary care [FTS-5 = 3.99 (1.14–13.89)] and geriatric clinic [FI-35 = 3.42 (1.56–7.49), FRAIL scale = 3.27 (1.21–8.86)]; for hospitalizations in primary care [FI-35 = 3.04 (1.25–7.39)]; and for falls in geriatric clinic [FI-35 = 2.21 (1.01–4.84)].

Conclusions No single assessment instrument performs the best for all settings and outcomes. While in inpatients several commonly used frailty instruments showed good sensitivities (mainly for mortality and BADL worsening) but usually poor specificities, the contrary happened in geriatric clinic. None of the instruments showed a good performance in primary care. The FI-35 and the FTS-5 showed the best profile among the instruments assessed.

Keywords Frailty; Screening; Geriatric wards; Geriatric clinic; Nursing homes; Primary care

Received: 27 September 2021; Revised: 4 February 2022; Accepted: 7 March 2022

*Correspondence to: Prof Leocadio Rodriguez-Mañas, Jefe de Servicio de Geriatría, Hospital Universitario de Getafe, Ctra. de Toledo, Km. 12,500, 28905 Madrid, Spain. Phone: 00 34 916839360 (ext. 6412), Fax: 00 34 916839210, Email: leocadio.rodriguez@salud.madrid.org

Introduction

An inevitable challenge facing health systems across Europe is the trend of a growing population of older people, often characterized by multimorbidity profiles and declines in physical and mental function. Particularly problematical are the increasing number of frail older people.[1]

Frailty identification is important as it is associated with an increased risk of serious adverse events and declines in physiologic reserve and function across multiorgan systems, including disability.[2–5] In addition, frailty in contrast to disability, is more easily reversible using appropriate interventions.[6] Prevention of disability contributes to the maintenance of quality of life, promotes independence, and is expected to prevent institutional care.[7]

When older patients are screened in a systematic way, potential and additional geriatric problems may be identified and tackled at an early stage.[8]

Currently, there are several instruments for frailty identification, with different sets of items and scoring systems.[9] Nevertheless, most of the instruments have been used in epidemiological studies or single settings, and there is scarce information on how a broad set of frailty instruments performs in different settings simultaneously.[8,10–12] A previous publication from the FRAILTOOLS project showed that most of the instruments had a low inter-test agreement when evaluated with the Cohen's kappa index, which raised the possibility that these instruments approach frailty from different perspectives, which justifies the comparison of their diagnostic accuracy in each specific setting.[13]

The main aim of this study was to compare the performance (i.e. sensitivity and specificity) of eight frailty instruments in assessing relevant adverse health outcomes (falls, increase in dependency in basic and instrumental activities of daily living—BADLs and IADLs—, hospitalization, and mortality) in different settings over a 12 month follow-up period in a cohort of older adults aged 75+ from five European cities. As a secondary objective, we aimed to assess the capacity of the instruments to predict adverse health outcomes beyond what is achieved by above age, gender, and multimorbidity.

Methods

The full methodology for the development of the FRAILTOOLS project is reported elsewhere.[14] Briefly, this project began in 2016 with research teams from Birmingham (United Kingdom), Cracow (Poland), Getafe (Spain), Rome (Italy), and Toulouse (France). It was an observational, longitudinal, and prospective study with a follow-up of 18 months. In total, 1440 adults (aged 75 + years) were consecutively enrolled from three clinical settings (geriatric wards, geriatric clinic, and primary care clinics) and one type of social care setting (nursing homes) in each city. Exclusion criteria were as follows: a Mini-Mental State Examination ≤ 20 points, a terminal illness (life expectancy ≤ 6 months), and a Barthel Index < 90 (except for nursing home residents where, taking into account the usual condition of the patients, exclusion took place when the Barthel index was <40). Demographic data (i.e. gender and age), along with the eight frailty measures and their outcomes, were collected at baseline and after 6, 12, and 18 months.

The sample size was calculated to obtain at least 10 cases per independent variable of the most unlikely event (mortality) at 12 months of follow-up.[15] There would be four variables in complete models (age, gender, Charlson index, and each frailty scale). For an expected average annual death rate for people aged 75 + years in the five European participating countries of 10%, the sample size that yielded a 95% confidence interval (95% CI) lower limit of at least 40 events was 1420. This figure was slightly increased to allow an incomplete determination of mortality in some countries, rendering a final sample size of 1440 individuals. The results of this manuscript are focused on the 12 month follow-up period.

Frailty measures

Eight frailty instruments were used in this study:

The *frailty phenotype* (FP)[16] includes three self-reported components: exhaustion, physical activity level, and weight loss, and two objective measures: grip strength and gait speed.[17] These two last items were measured in a standardized way in all settings of all countries. Original cut-off points stratified by gender and body mass index (BMI) and gender and height were used to define low grip strength and low gait speed, respectively.[18] A patient was considered frail if ≥ 3 criteria were positive.

The *Survey of Health, Ageing and Retirement in Europe Frailty Index* (SHARE-FI) assesses handgrip strength measured on a hand-held dynamometer and four self-reported items: physical exhaustion, loss of appetite, difficulties climbing stairs, and/or walking 100 m and low physical activity. To classify a participant as frail, data are entered into the freely accessed web-based calculator (http://www.biomedcentral.com/1471-2318/10/57/additional). The cut-off used to identify frailty is <6 in female patients and <7 in men.[19]

The *5-item Frailty Trait Scale* (FTS-5) is a short version of the original 12-item FTS proposed by García-García *et al*. for diag-

nosis and evolution of frailty.[20] It includes nutritional status assessed with the BMI, balance assessed with the Romberg test, physical activity evaluated using the Physical Activity Scale for the Elderly (PASE), and handgrip strength and gait speed using the same methodology than for these items of the FP.[21] It ranges from 0 to 50. If the score is >25 a person is considered frail.[21]

The *3-item FTS* (FTS-3) is the shortest version of the FTS. It includes BMI, balance, and physical activity. It ranges from 0 to 30. If the score is >15 a person is considered frail.[21]

The *FRAIL scale* comprises five self-reported criteria exploring the presence/absence of the following signs or symptoms: Fatigue, Resistance, Ambulation, Illness, and Loss of weight.[22] It ranges from 0 (*the best*) to 5 (*the worst*).[23] The cut-off used to identify frailty is $\geq$3.[24]

The *35-item Frailty Index* (FI-35) is based on the accumulation of health deficits (i.e. symptoms, signs, chronic diseases, disability, and laboratory abnormalities).[18,25] The score is calculated adding the number of deficits present and dividing it by the total number of possible deficits. For this project, 35 items were obtained from medical records, self-reported by participants, or measured at the patient's evaluation (refer to Supporting Information, *Table S1*). The cut-off used to define frailty was set to $\geq$0.25.[26]

The *Gérontopôle Frailty Screening Tool* (GFST) is an 8-item questionnaire intended to identify frailty. The first six questions evaluate the patient's status (living alone, loss of weight, fatigue, mobility difficulties, memory problems, and gait speed). In a second section, the clinician expresses his/her personal view about the frailty status of the individual. The presence of frailty is based on the guided judgement of the clinician.[27]

The *Clinical Frailty Scale* (CFS) evolved from the Canadian Study of Health and Aging. It uses a visual chart and clinical descriptors to assess the overall level of fitness or frailty. The frailty status is based on the clinician's judgement.[28] A cut-point of $\geq$4 was used to consider a person as frail.[24]

Frailty status was assessed during the interview done at the baseline visit in all settings. In geriatric wards, scales were administered at 48 h after admission once the acute phase had passed and the patient was clinically stabilized. Items that were not objectively measured were referred to the time indicated in the scale (e.g. 'last week' in the exhaustion items of the Fried's scale) or, if a time reference was not indicated, to the usual situation previous to the acute event that led to the hospitalization (e.g. FI and CFS). Those who remained clinically unstable after this time-period were not finally included.

The research health care team (nurses/geriatricians) of each participating centre administered the instruments in all settings and received the same training on their administration.[14] They were systematically trained about the full protocol, including how to contact the participants, explain the study to them, collect the informed consent from them or if needed from the tutor/accompanying person, administer the scales, and collect the data. The procedures were routinely and timely verified by the most trained geriatrician of each site, including the accuracy of the research team in measuring the different subjective and objectives components of the frailty instruments.

In addition, in order to increase the accuracy and precision in measuring the different instruments, a brief explanation about every instrument was included in the heading of the sheet where each of the tools were printed to remind on how the information should be collected.

Handling of missing data

For some of the instruments used in the article (SHARE-FI, FTS-3, FTS-5, CFS, and GFST), it was necessary that all items be measured. Therefore, in these instruments, the absence of an item implied that the patient could not be classified as frail or non-frail and the patient was considered missing for that instrument. Other instruments such as the Frail and the FP allowed to classify a patient as frail or not frail even in the presence of missing items if the available information met certain criteria. For a patient to be considered frail, he/she must have at least three positive items, and for a patient to be considered non-frail, two constraints have to be met: first, that the number of positive items is less than 3 and, second, that the sum of positive items plus missing items is less than 3. Finally, according to the protocol, the FI-35 allowed a missingness up to 20% of items in order to calculate the score and be able to classify the patient as frail or not.

Outcome measures

We selected five adverse outcomes to test the predictive ability of the frailty indices: falls, hospitalization, increase in dependency in BADL and IADL, and mortality.

Falls and hospitalization were assessed every 6 months with the questions: 'Have you fallen down in the last 6 months?' and 'Have you been admitted to a hospital in the previous 6 months?', respectively. IADL worsening was defined as a loss of one point in the Lawton and Brody index, measured at baseline and at 12 months of follow-up.[29] The score comprises eight questions about the degree to which a person actually performs IADL independently. Taking into account cultural habits of the participants, the items food preparation, housekeeping, and laundry were omitted in men. Accordingly, it ranges from 0 (*fully dependent*) to 8 (*fully independent*) for women and from 0 to 5 for men. The items omitted for men were food preparation, housekeeping and laundry. To determine the BADL worsening, the Barthel index was measured at baseline and after 12 months of follow-up. It comprises 10 items describing BADL and mobility. It scores from 0 to 100, and a higher score is a reflection of greater

ability to function independently.[30] A reduction in ≥5 points at 12 months of follow-up compared with baseline was considered as worsening in dependency in BADL. Survival status was obtained at 6 and 12 months from phone calls to arrange the follow-up visits and, if no answer was obtained, hospital registries (plus death registries in Spain). We combined both follow-ups to define the outcome 'dead along 1 year'.

In order to minimize the lost to follow-up, home visits were conducted, and in case home visits failed, phone calls were made in order to evaluate subjective instruments such as the Frail scale, PASE, Charlson, as well as the outcomes (falls, hospitalization, increase in dependency in BADL and IADL, and mortality).

Covariates

These were age, gender, and the Charlson comorbidity index,[31] as predefined in the protocol of the study.[14] For descriptive purposes, the BMI and the degree of physical activity measured with the PASE[32] were also assessed.

Statistical analysis

Descriptive statistics were calculated to summarize the data of the frailty instruments, baseline characteristics, and adverse outcomes. Contingency tables were used to calculate sensitivity (proportion of those classified as frail by the instrument in relation to those experiencing the outcome) and specificity (proportion of those not classified as frail by the instrument in relation to those not experiencing the outcome), and their 95% CIs were calculated. Because even in the absence of a relationship, sensitivity is conditioned by the proportion of positive results and specificity by the proportion of negative results, and these indicators need to be high for a screening test be considered useful, we considered a test 'good enough/reasonable' when its sensitivity was close to or greater than 70% and its confidence interval did not include the point estimate of the prevalence of a positive result (and also the confidence interval of the prevalence did not include the point estimate of the sensitivity). The same criterion was applied to specificity and the complementary of prevalence. The association between frailty status and adverse outcomes was assessed using multivariate logistic regression models adjusted by age, gender, and Charlson index. The level of statistical significance was set at $P \leq 0.05$. Analyses were made using R for Windows Version 3.6.1.

Results

A total of 1440 persons were included in the study (mean age 82.2 SD 5.5 years, 61.3% female) and were screened for frailty. Follow-up data were available from 996 individuals (mean age 82.1, SD 5.4 years, and 61.6% female). During a follow-up period of 12 months, 113 (11.3%) participants died, 260 (26.1%) had at least one fall, 169 (17.0%) were admitted to a hospital at least once, and 212 (21.3%) participants increased their IADL dependency and 289 (29.0%) their BADL dependency. Some of them suffered from more than one outcome. Characteristics of the study population by setting are described in Table 1. Differences between settings reflect the characteristics of their users. Those in nursing homes were older and more dependent, more frequently women, exercised less and fell and experienced a worsening in BADL more often. On the contrary, those coming from primary care and geriatric clinic were younger, with less morbidity, exercised more and were more frequently married. Those coming from geriatric wards were characterized by a higher mortality, hospitalization and IADL worsening rates. Loss to follow-up in geriatric wards (44.9%) doubled that in primary care (22.3%). There were statistical differences between those followed-up or not, but these varied by setting. The only common pattern was a tendency of those loss to follow-up to exercise less and living alone more frequently.

Analysis of sensitivity and specificity per settings (Tables 2 and 3)

Table 2 presents the results for people who are cared for in an institution temporarily (hospitals) or permanently (nursing homes) and Table 3 for people who live at their homes but attend geriatric clinics at the hospital or primary care.

In geriatric wards, the FI-35 met our criterion for good enough sensitivity for BADL worsening (74.6%; 95% CI: 64.2–85), mortality (69.1%; 95% CI: 56.8–81.3), and hospital admissions (67.8%; 95% CI: 55.8–79.7). It showed a reasonable specificity for BADL worsening (70.3%; 95% CI: 59.1–81.5) too. In addition, two instruments showed good sensitivities for mortality, the FTS-5 (67.9%; 95% CI: 50.6–85.2) and the SHARE-FI (85.5%; 95% CI: 76.1–94.7). This last instrument had also a reasonable sensitivity to predict falls (88.9%; 95% CI: 80.5–97.2), like the GFST (79.2%; 95% CI: 68.3–90.1). The FTS-3 and FTS-5 showed good enough specificities to predict BADL worsening (74.6%; 95% CI: 63.8–85.3 and 69.8; 95% CI: 56.0–83.5, respectively).

In nursing homes, all instruments but the FP and the FRAIL scale showed reasonable sensitivities (above 78%) to predict mortality. The FI-35 (96.4%; 95% CI: 93.0–99.9), FTS-5 (80.2; 95% CI: 72.0–88.4), and the FTS-3 (73.9; 95% CI: 65.7–82.0) met our criterion of sensitivity to predict BADL worsening. The FRAIL scale had a reasonable specificity to predict BADL worsening (69.4%; 95% CI: 60.3–78.5).

In geriatric clinic, no instrument met our sensitivity criterion for any outcome. Regarding specificity, it was good enough to predict BADL worsening for the FRAIL scale (93.6%; 95% CI:

Table 1 Characteristics of the participants at baseline and at 12 months by setting

Variable	Geriatric wards			Nursing homes		
	Participants at baseline (n = 361)	Participants followed-up (n = 199)	Participants lost to follow-up (n = 162; 44.9%)	Participants at baseline (n = 358)	Participants followed-up (n = 269)	Participants lost to follow-up (n = 89; 24.9%)
Mean (SD)						
Age (years)	83.2 (5.4)	82.7 (5.0)*	83.9 (5.8)*	85.8 (5.9)	86.2 (5.6)*	84.5 (6.3)*
Barthel Index	96.0 (4.5)	96.2 (4.5)	95.7 (4.5)	76.8 (19.2)	79.8 (17.8)***	67.9 (20.6)***
Lawton and Brody index						
Men (5 items)	3.5 (1.5)	3.7 (1.4)*	3.2 (1.5)*	2.1 (1.1)	2.2 (1.2)	1.9 (0.8)
Women (8 items)	5.5 (2.5)	5.8 (2.4)	5.2 (2.5)	2.0 (1.8)	2.2 (1.9)**	1.5 (1.3)**
Age adjusted Charlson index	6.2 (2.4)	6.2 (2.6)	6.3 (2.1)	6.1 (1.9)	6.3 (1.9)*	5.8 (1.9)*
PASE	52.6 (53.1)	59.1 (55.3)*	44.7 (49.2)*	17.6 (26.5)	20.9 (28.2)***	7.6 (17.7)***
BMI	26.1 (4.5)	25.8 (4.4)	26.5 (4.5)	26.0 (5.4)	26.3 (5.4)	25.0 (5.4)
Number (%)						
Gender (Female)	197 (54.7)	108 (54.3)	89 (55.3)	250 (69.8)	191 (71.0)	59 (66.3)
Partner status						
Married	143 (39.7)	85 (42.7)	58 (36.0)	49 (13.8)	39 (14.6)	10 (11.2)
Divorced	11 (3.1)	4 (2.0)	7 (4.3)	17 (4.8)	14 (5.2)	3 (3.4)
Widowed	181 (50.3)	99 (49.7)	82 (50.9)	239 (67.1)	175 (65.5)	64 (71.9)
Single	25 (6.9)	11 (5.5)	14 (8.7)	51 (14.3)	39 (14.6)	12 (13.5)
Cohabitation						
Alone	141 (39.2)	66 (33.2)***	75 (46.6)***	12 (3.4)	9 (3.3)***	3 (3.4)***
Spouse	142 (39.4)	84 (42.2)***	58 (36.0)***	13 (3.6)	10 (3.7)***	3 (3.4)***
Family member	59 (16.4)	34 (17.1)***	25 (15.5)***	5 (1.4)	2 (0.7)***	3 (3.4)***
Professional caregiver	7 (1.9)	6 (3.0)***	1 (0.6)***	249 (69.6)	178 (66.2)***	71 (79.8)***
Others	11 (3.1)	9 (4.5)***	2 (1.2)***	79 (22.1)	70 (26.0)***	9 (10.1)***
Missing	1		1			
Number of events at follow-up						
Falls		55 (27.6)			101 (37.5)	
Hospital admission		60 (30.2)			44 (16.4)	
IADL worsening		58 (29.1)			63 (23.4)	
BADL worsening		68 (34.2)			116 (43.1)	
Mortality		60 (30.2)			42 (15.6)	

BADL, basic activities of daily living; BMI, body mass index; IADL, independent activities of daily living; PASE, Physical Activity Scale for the Elderly; SD, standard deviation.
*$P < 0.05$.
[1]**$P < 0.01$.
[2]***$P < 0.001$.

Table 1 (continued)

Variable	Geriatric clinic			Primary care		
	Participants at baseline (n = 340)	Participants followed-up (n = 232)	Participants lost to follow-up (n = 108; 31.8%)	Participants at baseline (n = 381)	Participants followed-up (n = 296)	Participants lost to follow-up (n = 85; 22.3%)
Mean (SD)						
Age (years)	79.9 (4.2)	79.9 (4.3)	80.0 (4.1)	80.0 (3.9)	79.7 (3.8)*	80.9 (4.3)*
Barthel Index	97.5 (3.4)	97.2 (3.7)	98.1 (2.8)	97.7 (3.1)	97.8 (3.1)	97.5 (3.4)
Lawton and Brody index						
Men (5 items)	4.7 (0.8)	4.7 (0.7)	4.7 (0.8)	4.8 (0.6)	4.8 (0.6)	4.7 (0.7)
Women (8 items)	7.4 (1.3)	7.3 (1.5)	7.6 (0.8)	7.7 (0.8)	7.7 (0.7)***	7.6 (0.8)**
Age adjusted Charlson index	4.8 (2.2)	4.9 (2.4)	4.5 (1.4)	4.6 (1.5)	4.5 (1.5)	4.8 (1.6)
PASE	98.7 (68.5)	103.0 (72.0)	89.6 (59.3)	94.7 (64.4)	98.7 (68.5)	80.7 (44.4)
BMI	27.2 (4.3)	27.6 (4.4)*	26.4 (4.0)*	27.0 (4.0)	26.9 (4.0)	27.2 (4.2)
Number (%)						
Gender (Female)	208 (61.4)	143 (61.6)	65 (60.7)	226 (59.3)	172 (58.1)	54 (63.5)
Partner status						
Married	177 (52.2)	117 (50.4)***	60 (56.1)***	212 (55.6)	167 (56.4)***	45 (52.9)***
Divorced	21 (6.2)	11 (4.7)***	10 (9.3)***	15 (3.9)	14 (4.7)***	1 (1.2)***
Widowed	125 (36.9)	90 (38.8)***	35 (32.7)***	129 (33.9)	96 (32.4)***	33 (38.8)***
Single	16 (4.7)	14 (6.0)***	2 (1.9)***	25 (6.6)	19 (6.4)***	6 (7.1)***
Cohabitation						
Alone	131 (38.6)	89 (38.4)***	42 (39.3)***	137 (36.0)	104 (35.1)	33 (38.8)
Spouse	174 (51.3)	115 (49.6)***	59 (55.1)***	209 (54.9)	166 (56.1)	43 (50.6)
Family member	27 (8.0)	23 (9.9)***	4 (3.7)***	29 (7.6)	22 (7.4)	7 (8.2)
Professional caregiver	1 (0.3)	1 (0.4)***	0 (0.0)***	1 (0.3)	0 (0.0)	1 (1.2)
Others	6 (1.8)	4 (1.7)***	2 (1.9)***	5 (1.3)	4 (1.4)	1 (1.2)
Missing	1		1			
Number of events at follow-up						
Falls		62 (26.7)			42 (14.2)	
Hospital admission		27 (11.6)			38 (12.8)	
IADL worsening		44 (19.0)			47 (15.9)	
BADL worsening		56 (24.1)			49 (16.6)	
Mortality		9 (3.9)			2 (0.7)	

BADL, basic activities of daily living; BMI, body mass index; IADL, independent activities of daily living; PASE, Physical Activity Scale for the Elderly; SD, standard deviation.
*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.

Table 2 Sensitivity and specificity of eight frailty screening instruments for the development of falls, hospital readmission, basic and instrumental activities of daily living disability, and mortality in geriatric wards and nursing homes

Instruments[a]	Prevalence of frailty (95% CI)	Falls		Hospital admission	
		Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)
Geriatric wards (n = 199)					
FP	63.1 (54.1, 72.0)	57.7 (38.7, 76.7)	35.7 (24.5, 46.9)	55.6 (39.3, 71.8)	34.7 (21.4, 48.0)
SHARE-FI	74.9 (68.7, 81.0)	**88.9 (80.5, 97.3)**	24.3 (14.0, 33.4)	80.7 (70.5, 90.1)	22.9 (13.9, 31.9)
FTS-5	47.4 (37.5, 57.4)	60.0 (38.5, 81.5)	64.8 (52.1, 77.6)	38.5 (19.8, 57.2)	57.9 (42.2, 73.6)
FTS-3	47.7 (40.7, 54.3)	58.5 (45.2, 71.8)	56.4 (47.1, 65.6)	54.2 (41.5, 67.0)	52.3 (41.8, 62.9)
Frail	46.6 (39.6, 53.7)	56.6 (48.9, 64.3)	56.1 (48.4, 63.8)	47.4 (39.1, 55.6)	58.3 (50.2, 66.5)
FI-35	56.8 (49.8, 63.8)	64.8 (57.5, 72.1)	47.3 (39.6, 54.9)	**67.8 (60.2, 75.4)**	46.5 (38.4, 54.6)
GFST	70.8 (64.4, 77.2)	**79.2 (68.3, 90.2)**	33.6 (24.8, 42.5)	72.9 (61.5, 84.2)	26.7 (17.4, 36.1)
CFS	52.6 (45.5, 59.7)	64.2 (56.8, 71.5)	53.6 (46.0, 61.3)	61.0 (53.1, 69.0)	53.5 (45.4, 61.6)
Nursing homes (n = 269)					
FP	56.8 (49.8, 63.8)	54.9 (43.4, 66.5)	43.3 (33.4, 53.2)	58.8 (42.3, 75.4)	45.2 (36.5, 53.9)
SHARE-FI	80.2 (75.3, 85.1)	81.7 (73.9, 89.6)	23.7 (16.1, 31.4)	79.5 (67.6, 91.5)	22.1 (15.5, 28.6)
FTS-5	69.6 (63.5–75.7)	75.0 (65.3, 84.7)	40.2 (30.9, 49.5)	61.8 (45.4, 78.1)	34.3 (26.3, 42.4)
FTS-3	66.9 (61.3, 72.5)	71.0 (61.7, 80.2)	42.1 (33.4, 50.9)	53.7 (38.4, 68.9)	35.0 (27.6, 42.5)
Frail	40.5 (34.5, 46.4)	41.7 (35.1, 48.2)	61.5 (55.0, 67.9)	46.5 (39.7, 53.4)	60.6 (53.9, 67.3)
FI-35	91.7 (88.4, 95.0)	93.9 (90.8, 97.1)	12.4 (8.0, 16.8)	88.6 (84.3, 93.0)	9.9 (5.8, 14.0)
GFST	82.5 (77.9, 87.0)	86.9 (80.2, 93.5)	24.6 (16.9, 32.2)	84.1 (73.3, 94.9)	21.6 (15.3, 27.9)
CFS	75.7 (70.6, 80.9)	78.0 (72.5, 83.5)	33.1 (26.9, 39.3)	75.0 (69.1, 80.9)	30.2 (24.0, 36.5)

BADL, basic activities of daily living; CFS, Clinical Frailty Scale; FI-35, 35-item Frailty Index; FP, frailty phenotype; FTS-3, 3-item Frailty Trait Scale; FTS-5, 5-item Frailty Trait Scale; GFST, Gérontopôle Frailty Screening Tool; IADL, independent activities of daily living; Sens, sensitivity; SHARE-FI, Survey of Health Ageing and Retirement in Europe Frailty Index; Spec, specificity.

In bold, sensitivities and specificities close to or greater than 70% whose confidence intervals do not include the prevalence of a positive result or its complementary, respectively.

[a]Cut-off scores: FP: ≥3. FRAIL: ≥3. FI-35: ≥0.25. CFS: ≥4. SHARE-FI: <6 female patients, <7 men. GFST: yes. FTS-3: ≥15. FTS-5: ≥25.

Table 2 (continued)

Instruments[a]	IADL worsening		BADL worsening		Mortality	
	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)
Geriatric wards (n = 199)						
FP	62.5 (45.7, 79.2)	44.7 (30.5, 58.9)	62.2 (46.5, 77.8)	45.2 (30.2, 60.3)	71.4 (54.7, 88.2)	36.0 (26.6, 45.4)
SHARE-FI	66.1 (53.7, 78.5)	23.9 (14.0, 33.9)	80.0 (70.1, 89.8)	37.1 (25.1, 49.1)	**85.5 (76.1, 94.8)**	20.8 (14.5, 27.1)
FTS-5	33.3 (14.5, 52.2)	60.5 (45.0, 76.1)	52.6 (30.2, 75.1)	**69.8 (56.0, 83.5)**	**67.9 (50.6, 85.2)**	58.0 (47.3, 68.8)
FTS-3	42.9 (29.9, 55.8)	55.4 (44.1, 66.7)	61.2 (49.5, 72.9)	**74.6 (63.9, 85.4)**	61.4 (48.7, 74.0)	54.7 (46.9, 62.3)
Frail	36.8 (28.5, 45.2)	60.0 (51.5, 68.5)	33.8 (25.6, 42.1)	56.5 (47.8, 65.1)	62.5 (56.0, 69.0)	51.9 (45.2, 58.6)
FI-35	57.9 (49.4, 66.3)	51.4 (42.8, 59.9)	**74.6 (67.2, 82.1)**	**70.3 (62.5, 78.1)**	**69.1 (62.9, 75.2)**	44.4 (37.8, 51.1)
GFST	69.1 (56.9, 81.3)	35.6 (24.6, 46.6)	77.6 (67.6, 87.6)	45.9 (33.4, 58.4)	73.2 (61.6, 84.8)	29.2 (22.2, 36.2)
CFS	45.3 (36.6, 54.0)	56.2 (47.5, 64.8)	50.7 (42.0, 59.5)	62.7 (54.3, 71.2)	63.0 (56.5, 69.4)	49.1 (42.4, 55.8)
Nursing homes (n = 269)						
FP	47.8 (33.4, 62.3)	42.2 (33.3, 51.2)	63.8 (53.2, 74.3)	53.7 (42.9, 64.5)	60.9 (40.0, 80.8)	44.5 (36.7, 52.3)
SHARE-FI	76.3 (65.4, 87.1)	21.1 (14.5, 27.7)	82.6 (75.4, 89.7)	27.1 (18.2, 36.0)	**88.9 (78.6, 99.2)**	20.3 (14.7, 25.9)
FTS-5	56.0 (42.2, 69.8)	29.4 (21.4, 37.3)	**80.2 (72.0, 88.4)**	48.2 (37.6, 58.9)	**86.2 (73.7, 98.8)**	33.7 (26.6, 40.9)
FTS-3	52.6 (39.7, 65.6)	30.9 (23.5, 38.3)	**73.9 (65.7, 82.0)**	46.8 (36.7, 56.9)	**78.9 (66.0, 91.9)**	35.9 (29.2, 42.)
Frail	37.7 (31.1, 44.3)	61.1 (54.5, 67.7)	45.9 (39.2, 52.7)	**69.4 (63.1, 75.6)**	51.3 (45.0, 57.6)	59.9 (53.7, 66.1)
FI-35	83.6 (78.6, 88.6)	6.6 (3.3, 10.0)	**96.4 (93.9, 98.9)**	16.2 (11.2, 21.1)	**97.6 (95.6, 99.5)**	9.8 (6.1, 13.5)
GFST	69.4 (57.9, 80.8)	15.9 (10.1, 21.7)	85.1 (78.5, 91.6)	26.5 (17.8, 35.3)	**92.7 (84.7, 1.0)**	19.5 (14.1, 24.9)
CFS	53.2 (46.5, 59.9)	19.2 (13.9, 24.5)	76.3 (70.6, 82.0)	31.6 (25.4, 37.9)	**87.8 (83.7, 91.9)**	28.3 (22.7, 33.9)

BADL, basic activities of daily living; CFS, Clinical Frailty Scale; FI-35, 35-item Frailty Index; FP, frailty phenotype; FTS-3, 3-item Frailty Trait Scale; FTS-5, 5-item Frailty Trait Scale; GFST, Gérontopôle Frailty Screening Tool; IADL, independent activities of daily living; Sens, sensitivity; SHARE-FI, Survey of Health Ageing and Retirement in Europe Frailty Index; Spec, specificity.

In bold, sensitivities and specificities close to or greater than 70% whose confidence intervals do not include the prevalence of a positive result or its complementary, respectively.

[a]Cut-off scores: FP: ≥3. FRAIL: ≥3. FI-35: ≥0.25. CFS: ≥4. SHARE-FI: <6 female patients, <7 men. GFST: yes. FTS-3: ≥15. FTS-5: ≥25.

Table 3 Sensitivity and specificity of eight frailty screening instruments for the development of falls, hospital readmission, basic and instrumental activities of daily living disability, and mortality in primary care and geriatric clinic

Instruments[a]	Prevalence of frailty (95% CI)	Falls		Hospital admission	
		Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)
Primary care (n = 296)					
FP	14.3 (9.8, 18.8)	22.0 (9.3, 34.6)	87.2 (82.1, 92.3)	30.6 (15.5, 45.6)	89.3 (84.4, 94.3)
SHARE-FI	29.1 (23.7, 34.5)	38.1 (23.4, 52.8)	71.7 (65.4, 78.0)	36.1 (20.4, 51.8)	72.3 (65.8, 78.8)
FTS-5	5.9 (3.1, 8.7)	2.7 (0, 7.9)	94.0 (90.8, 97.1)	5.6 (0, 13.0)	94.0 (90.7, 97.3)
FTS-3	6.1(3.3, 8.8)	9.5(0.6, 18.4)	94.7 (91.7, 97.6)	5.3 (0, 12.4)	93.8 (90.5, 97.1)
Frail	7.9 (4.8, 11.0)	9.8 (0.7, 18.8)	91.8 (88.2, 95.4)	7.9 (0.7, 16.5)	91.1 (87.2, 95.0)
FI-35	13.6 (9.7, 17.5)	21.4 (9.0, 33.8)	87.1 (82.7, 91.5)	28.9 (14.5, 43.4)	88.5 (84.1, 92.8)
GFST	31.4 (26.1, 36.7)	45.2 (30.2, 60.3)	68.9 (62.8, 74.9)	50.0 (34.1, 65.9)	68.9 (62.6, 75.2)
CFS	15.2 (11.1, 19.3)	14.3 (3.7, 24.9)	84.0 (79.2, 88.8)	23.7 (10.2, 37.2)	85.2 (80.3, 90.0)
Geriatric clinic (n = 232)					
FP	23.5 (17.2, 29.9)	33.3 (20.8, 45.9)	**84.4 (77.0, 91.9)**	26.1 (8.1, 44.0)	77.8 (70.2, 85.3)
SHARE-FI	44.2 (37.7, 50.7)	54.2 (41.5, 66.9)	63.2 (55.0, 71.4)	44.4 (25.7, 63.2)	58.6 (51.1, 66.2)
FTS-5	14.0 (9.4, 18.5)	12.3 (3.8, 20.8)	88.1 (82.6, 93.6)	18.5 (3.9, 33.2)	87.5 (82.4, 92.6)
FTS-3	9.5 (5.7, 13.3)	8.3 (1.3, 15.3)	93.5 (89.4, 97.6)	11.1 (0, 23.0)	93.5 (89.7, 97.2)
Frail	10.9 (6.9, 15.0)	15.0 (6.0, 24.0)	92.0 (87.5, 96.5)	14.8 (1.4, 28.2)	90.4 (86.0, 94.9)
FI-35	19.9 (14.7, 25.1)	28.3 (16.9, 39.7)	84.2 (78.1, 90.2)	33.3 (15.6, 51.1)	82.1 (76.4, 87.9)
GFST	37.9 (31.7, 44.2)	48.3 (35.7, 61.0)	66.2 (58.3, 74.1)	48.1 (29.3, 67.0)	63.1 (55.8, 70.4)
CFS	19.8 (14.7, 25.0)	25.0 (14.0, 36.0)	82.7 (76.5, 89.0)	25.9 (9.4, 42.5)	81.5 (75.7, 87.4)

BADL, basic activities of daily living; CFS, Clinical Frailty Scale; FI-35, 35-item Frailty Index; FP, frailty phenotype; FTS-3, 3-item Frailty Trait Scale; FTS-5, 5-item Frailty Trait Scale; GFST, Gérontopôle Frailty Screening Tool; IADL, independent activities of daily living; Sens, sensitivity; SHARE-FI, Survey of Health Ageing and Retirement in Europe Frailty Index; Spec, specificity.

In bold, sensitivities and specificities close to or greater than 70% whose confidence intervals do not include the prevalence of a positive result or its complementary, respectively.

[a]Cut-off scores: FP: ≥ 3. FRAIL: ≥ 3. FI-35: ≥0.25. CFS: ≥4. SHARE-FI: <6 female patients, <7 men. GFST: yes. FTS-3: ≥15. FTS-5: ≥25.

Table 3 (continued)

Instruments[a]	IADL worsening		BADL worsening		Mortality	
	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)
Primary care (n = 296)						
FP	22.7 (10.3, 35.1)	87.8 (83.0, 92.6)	29.0 (13.1, 45.0)	88.1 (83.5, 92.7)	0.0 (0.0, 0.0)	84.8 (79.9, 89.7)
SHARE-FI	26.1 (13.4, 38.8)	71.2 (65.1, 77.2)	29.5 (16.1, 43.0)	71.9 (65.9, 77.9)	50.0 (0, 100)	70.3 (64.5, 76.1)
FTS-5	12.5 (2.3, 22.7)	95.5 (92.8, 98.2)	8.7(6.0, 16.8)	95.0 (92.0, 97.9)	50.0 (0, 100)	94.5 (91.6, 97.3)
FTS-3	8.5 (0.5, 16.5)	94.2 (91.2, 97.1)	10.2 (1.7, 18.7)	94.6 (91.7, 97.4)	0.0(0.0, 0.0)	94.4 (91.6, 97.1)
Frail	12.8 (3.2, 22.3)	93.2 (90.0, 96.4)	12.5 (3.1, 21.9)	93.2 (89.9 96.4)	0.0 (0.0, 0.0)	91.5 (88.1, 94.9)
FI-35	21.3 (9.6, 33.0)	87.5 (83.3, 91.7)	10.2 (1.7, 18.7)	85.3 (80.8, 89.8)	00.0 (0.0, 0.0)	85.7 (81.5, 89.9)
GFST	48.9 (34.6, 63.2)	72.2 (66.5, 77.9)	38.8 (25.1, 52.4)	70.3 (64.5, 76.1)	50.0 (0, 100)	65.4 (59.7, 71.1)
CFS	29.8 (16.7, 42.9)	87.6 (83.4, 91.7)	20.4 (9.1, 31.7)	85.8 (81.3, 90.2)	0.0 (0.0, 0.0)	83.8 (79.4, 88.3)
Geriatric clinic (n = 232)						
FP	27.8 (13.1, 42.4)	77.5 (70.0, 85.0)	38.5 (23.2, 53.7)	82.6 (75.7, 89.5)	0.0 (0.0, 0.0)	76.5 (70.0, 83.1)
SHARE-FI	46.5 (31.6, 61.4)	57.5 (49.8, 65.2)	57.1 (43.3, 71.0)	61.8 (54.1, 69.6)	44.4 (12.0, 76.9)	57.2 (50.5, 63.9)
FTS-5	21.4 (9.0, 33.8)	89.4 (84.6, 94.1)	23.1 (11.6, 34.5)	**90.7 (86.0, 95.3)**	22.2 (0, 49.4)	87.0 (82.5, 91.6)
FTS-3	14.0 (3.6, 24.3)	93.5 (89.7, 97.2)	11.8 (2.9, 20.6)	93.7 (89.9, 97.5)	22.2 (16.8, 27.6)	91.7 (88.0, 95.4)
Frail	20.9 (8.8, 33.1)	91.6 (87.3, 95.8)	23.5 (11.9, 35.2)	**93.6 (89.8, 97.5)**	12.5 (0, 35.4)	89.3 (85.2, 93.4)
FI-35	41.9 (27.1, 56.6)	84.4 (78.9, 89.9)	39.2 (25.8, 52.6)	**85.4 (79.8, 90.9)**	11.1 (0, 31.6)	79.2 (73.8, 84.6)
GFST	51.2 (36.2, 66.1)	65.5 (58.3, 72.7)	51.0 (37.3, 64.7)	67.1 (59.8, 74.4)	33.3 (2.5, 64.1)	61.8 (55.3, 68.2)
CFS	23.3 (10.6, 35.9)	81.5 (75.7, 87.4)	27.5 (15.2, 39.7)	84.2 (78.5, 89.9)	33.3 (2.5, 64.1)	81.1 (75.9, 86.3)

BADL, basic activities of daily living; CFS, Clinical Frailty Scale; FI-35, 35-item Frailty Index; FP, frailty phenotype; FTS-3, 3-item Frailty Trait Scale; FTS-5, 5-item Frailty Trait Scale; GFST, Gérontopôle Frailty Screening Tool; IADL, independent activities of daily living; Sens, sensitivity; SHARE-FI, Survey of Health Ageing and Retirement in Europe Frailty Index; Spec, specificity.

In bold, sensitivities and specificities close to or greater than 70% whose confidence intervals do not include the prevalence of a positive result or its complementary, respectively.

[a]Cut-off scores: FP: ≥ 3. FRAIL: ≥ 3. FI-35: ≥0.25. CFS: ≥4. SHARE-FI: <6 female patients, <7 men. GFST: yes. FTS-3: ≥15. FTS-5: ≥25.

Table 4 Adjusted odds ratios (age, gender, and Charlson index) of the eight frailty instruments per outcome measure and per setting

Instruments[a]	Falls OR (95% CI)	Hospital admission OR (95% CI)	IADL worsening OR (95% CI)	BADL worsening OR (95% CI)	Mortality OR (95% CI)
All settings					
Number of events	260	169	212	289	113
FP	1.16 (0.79–1.69)	1.57 (1.02–2.4)*	1.32 (0.88–1.97)	2.48 (1.69–3-65)***	2.29 (1.24–4.23)**
SHARE-FI	1.68 (1.17–2.4)**	1.64 (1.08–2.48)*	1.19 (0.82–1.73)	2.09 (1.45–3.00)***	3.98 (2.20–7.22)***
FTS-5	1.58 (1.05–2.38)*	1.03 (0.62–1.7)	1.28 (0.81–2.02)	2.99 (1.98–4.51)***	4.76 (2.50–9.06)***
FTS-3	1.70 (1.20–2.42)**	1.18 (0.78–1.77)	1.04 (0.70–1.54)	3.19 (2.23–4.56)***	3.07 (1.91–4.94)***
Frail	1.36 (0.96–1.94)	1.45 (0.97–2.17)	1.43 (0.97–2.10)	1.92 (1.32–2.77)***	2.77 (1.78–4.32)***
FI-35	1.90 (1.35–2.68)***	1.81 (1.22–2.68)***	1.58 (1.10–2.28)*	3.32 (2.35–4.68)***	2.41 (1.45–4.02)***
GFST	1.90 (1.36–2.64)***	1.67 (1.36–2.78)**	1.52 (1.07–2.15)*	2.25 (1.62–3.11)***	2.02 (1.22–3.33)**
CFS	1.56 (1.12–2.19)*	1.58 (1.07–2.33)*	0.93 (0.64–1.34)	1.79 (1.27–2.50)***	2.57 (1.59–4.17)***
Geriatric wards					
Number of events	55	60	58	68	60
FP	0.41 (0.14–1.23)	0.52 (0.20–1.39)	1.10 (0.40–2.98)	0.82 (0.29–2.29)	1.04 (0.36–3.02)
SHARE-FI	2.11 (0.70–6.34)	1.14 (0.43–2.96)	0.50 (0.20–1.27)	2.26 (0.87–5.88)	1.60 (0.62–4.11)
FTS-5	2.28 (0.74–7.01)	0.77 (0.25–2.40)	0.65 (0.20–2.17)	2.58 (0.70–9.51)	2.51 (0.90–7.06)
FTS-3	1.48 (0.73–2.99)	1.20 (0.60–2.40)	0.79 (0.37–1.66)	3.87 (1.76–8.48)***	1.56 (0.81–3.02)
Frail	1.40 (0.69–2.82)	1.23 (0.61–2.46)	0.79 (0.37–1.67)	0.49 (0.22–1.09)	1.65 (0.87–3.15)
FI-35	1.30 (0.64–2.65)	1.70 (0.82–3.53)	1.22 (0.59–2.54)	5.94 (2.69–13.14)***	1.35 (0.67–2.70)
GFST	1.80 (0.79–4.09)	0.95 (0.44–2.06)	1.19 (0.55–2.58)	2.81 (1.25–6.33)*	1.01 (0.49–2.07)
CFS	1.75 (0.86–3.54)	1.69 (0.84–3.40)	0.92 (0.43–1.94)	1.33 (0.62–2.86)	1.38 (0.71–2.70)
Nursing homes					
Number of events	101	44	63	116	42
FP	0.89 (0.48–1.68)	1.19 (0.55–2.60)	0.74 (0.36–1.51)	1.95 (1.03–3.68)*	1.13 (0.44–2.85)
SHARE-FI	1.49 (0.68–3.28)	1.15 (0.44–3.02)	0.60 (0.25–1.44)	1.60 (0.73–3.49)	3.49 (1.02–12.00)*
FTS-5	1.71 (0.87–3.34)	0.80 (0.35–1.83)	0.56 (0.27–1.14)	3.20 (1.61–6.38)***	2.85 (0.88–9.20)
FTS-3	1.71 (0.95–3.10)	0.58 (0.28–1.20)	0.56 (0.29–1.07)	2.31 (1.27–4.21)**	1.89 (0.80–4.49)
Frail	1.01 (0.57–1.79)	1.23 (0.61–2.49)	0.95 (0.49–1.85)	1.91 (1.05–3.48)*	1.82 (0.88–3.73)
FI-35	1.93 (0.70–5.28)	0.75 (0.25–2.25)	0.41 (0.16–1.09)	4.88 (1.54–15.44)**	4.15 (0.53–32.53)
GFST	2.06 (0.98–4.34)	1.35 (0.53–3.43)	0.49 (0.23–1.05)	1.80 (0.87–3.69)	2.98 (0.84–10.59)
CFS	1.52 (0.79–2.92)	1.14 (0.50–2.61)	0.25 (0.12–0.51)***	1.30 (0.68–2.50)	3.25 (1.16–9.12)*
Primary care					
Number of events	42	38	47	49	2
FP	1.57 (0.64–3.88)	4.32 (1.72–10.88)**	2.67 (1.04–6.83)*	2.65 (1.06–6.63)*	—
SHARE-FI	1.37 (0.63–2.96)	1.98 (0.85–4.62)	1.51 (0.65–3.51)	0.97 (0.45–2.09)	—
FTS-5	0.32 (0.04–2.62)	0.99 (0.21–4.80)	3.99 (1.14–13.89)*	1.57 (0.46–5.35)	—
FTS-3	1.33 (0.38–4.61)	0.82 (0.17–3.86)	1.50 (0.44–5.16)	1.61 (0.52–4.95)	—
Frail	0.83 (0.25–2.77)	0.78 (0.21–2.95)	1.78 (0.59–5.39)	1.71 (0.6–4.82)	—
FI-35	1.52 (0.61–3.79)	3.04 (1.25–7.39)*	1.60 (0.66–3.91)	0.59 (0.21–1.70)	—
GFST	1.55 (0.77–3.11)	2.19 (1.06–4.50)*	2.34 (1.18–4.63)*	1.38 (0.72–2.67)	—
CFS	0.62 (0.23–1.66)	1.70 (0.71–4.07)	2.47 (1.13–5.42)*	1.37 (0.61–3.09)	—
Geriatric clinic					
Number of events	62	27	44	56	9
FP	2.29 (0.96–5.45)	1.08 (0.35–3.32)	1.16 (0.43–3.13)	2.13 (0.90–5.07)	—
SHARE-FI	1.46 (0.73–2.89)	0.88 (0.35–2.19)	1.22 (0.55–2.69)	1.82 (0.87–3.81)	0.97 (0.21–4.40)
FTS-5	0.72 (0.27–1.93)	1.21 (0.38–3.88)	2.30 (0.85–6.19)	2.23 (0.90–5.52)	1.74 (0.28–10.86)
FTS-3	1.08 (0.33–3.52)	1.37 (0.33–5.73)	2.11 (0.68–6.50)	1.52 (0.50–4.62)	2.91 (0.46–18.34)
Frail	1.90 (0.72–5.01)	1.95 (0.56–6.72)	3.27 (1.21–8.86)*	4.48 (1.73–11.58)**	1.14 (0.12–10.44)
FI-35	2.21 (1.01–4.84)*	1.94 (0.76–4.95)	3.42 (1.56–7.49)**	3.30 (1.55–7.00)**	0.34 (0.04–3.02)
GFST	1.63 (0.8–3.08)	1.49 (0.64–3.47)	2.14 (1.04–4.41)*	1.94 (1.00–3.78)*	0.78 (0.18–3.36)
CFS	1.42 (0.66–3.05)	1.19 (0.45–3.17)	1.06 (0.45–2.50)	1.64 (0.76–3.55)	1.77 (0.40–7.91)

BADL, basic activities of daily living; CFS, Clinical Frailty Scale; FI-35, 35-item Frailty Index; FP, frailty phenotype; FTS-3, 3-item Frailty Trait Scale; FTS-5, 5-item Frailty Trait Scale; GFST, Gérontopôle Frailty Screening Tool; IADL, independent activities of daily living; Sens, sensitivity; SHARE-FI, Survey of Health Ageing and Retirement in Europe Frailty Index; Spec, specificity.

[a]Cut-off scores: FP: ≥ 3. FRAIL: ≥ 3. FI-35: ≥0.25. CFS: ≥4. SHARE-FI: <6 female patients, <7 men. GFST: yes. FTS-3: ≥15. FTS-5: ≥25.

*$P < 0.05$.

**$P < 0.01$.

***$P < 0.001$.

89.8–97.5), FTS-5 (90.7%; 95% CI: 86.0–95.3), and FI-35 (85.4%; 95% CI: 79.8–90.9). Specificity of the FP was adequate to predict falls (84.4%; 95% CI: 77.0–91.9).

No instrument met our sensitivity nor specificity criteria for any outcome in primary care.

Logistic regression

Bivariate analyses showed significant associations of almost all frailty scales with all the outcomes (fewer with IADL worsening) in the combined sample (refer to *Table S2*). In logistic regressions adjusted for age, gender, and Charlson index (*Table 4*), associations remained significant in the combined sample for all instruments with mortality, BADL worsening, and falls (but for the FP and the FRAIL scale), although with different strengths. The FTSs and the FRAIL scale were not able to predict hospital admissions, and only the FI-35 and the GFST predicted IADL dependency worsening.

In geriatric wards, no scale was able to predict mortality and only the FI-35, the FTS-3, and the GFST BADL worsening. In nursing homes, the SHARE-FI and the CFS could predict mortality and the FI-35, FTSs, FRAIL, and FP BADL worsening. The inverse association with IADL worsening remained significant only for the CFS. In primary care, hospital admission was predicted by FP, FI-35 and GFST, and IADL worsening by FP, FTS-5, GFST, and CFS; BADL worsening only by FP. In geriatric clinic, only the FI-35 remained associated with falls, and FRAIL, FI-35, and GFST with BADL and IADL worsening.

Discussion

To our knowledge, this study compares for the first time the performance of well-known and broadly used frailty instruments to assess various adverse outcomes (falls, hospitalization, increase in dependency in BADL and IADL, and mortality) in several specific clinical and social care settings in a very-old population (+75). Specificities for BADL worsening were good enough for the FTS-5 in geriatric inpatient and outpatient services. The FI-35 showed fine sensitivities for mortality, BADL worsening, and hospitalization and good specificity for BADL worsening in geriatric wards. Several instruments offered reasonable sensitivities for mortality and BADL worsening in nursing homes, but none showed good specificities. Conversely, only some specificities were high in geriatric clinic, those for BADL worsening of the already mentioned FTS-5 and the FI-35, and the FRAIL scale and that of the FP for the prediction of falls. Under the light of our results, no instrument may be recommended for screening in primary care. The ability of some instruments to predict outcomes over age, gender, and multimorbidity was limited to BADL worsening in geriatric wards and nursing homes, mortality in nursing homes, BADL and IADL worsening in primary care and geriatric clinic, hospitalizations in primary care and falls in geriatric clinic.

It may seem surprising that the sensitivity and specificity of the instruments depended on the setting in which they were applied. This is due to two reasons. The most determinant one is the prevalence of the condition as detected by a positive result in the test. The higher the prevalence, the highest the sensitivity, and the lowest the prevalence, the highest the specificity, even in the absence of an association between the test and the condition. This explains why sensitivities are high in inpatient settings (and more in nursing homes than in geriatric wards) and specificities are high in outpatient settings. This also justifies that, despite obtaining high specificities in primary care, no instrument could be recommended because none of them significantly exceeded the complementary of the prevalence of frailty according to that instrument. Taking into account the relationship between the prevalence of frailty detected by the instruments, their sensitivity and specificity, is not done in evaluations of the diagnostic accuracy of frailty instruments, generally speaking. However, there is something else. The difference between sensitivity and prevalence was greater, and the difference between specificity and complementary prevalence was less, for most of the instruments in geriatric wards than in nursing homes.

This may be due to the fact that in geriatric wards the patients were acutely ill and for that reason the assessment could have overestimated the actual frailty status, producing high values of sensitivity at the expense of low specificities. The consequence of these two reasons is that frailty instruments were only useful for screening in settings that require admission due to their high sensitivity, and only for excluding diagnosis in geriatric clinic, where their specificities were high.

The instruments that met our criteria for good enough sensitivity or specificity did not always obtain significant odds ratio after adjusting for age, gender, and multimorbidity. This could be because the other indicator, specificity or sensitivity, was not high enough, like in the SHARE-FI and the FI-35 and mortality, and the FI-35 and hospitalizations in geriatric wards; the FI-35 and the GFST in mortality in nursing homes; and the FTS-5 in BADL worsening in geriatric clinic. Another explanation is reduced power due to a low number of events. This could be the case for the difficulties in predicting mortality and hospitalizations, but not for falls, which was a common event in our sample. These instruments seem genuinely not suited to predict falls so well above sociodemographic and comorbidity variables.

Comparison with previous publications using different instruments is challenging due to the fact that the timeframe,[12,33] the instruments evaluated,[7] and the settings and characteristics of the participants[12] are different. The closest study to ours is the one performed by Bongue et al.[33] regarding the comparison of FP with other instruments at the community setting; they found comparable specificities but much higher sensitivities for mortality and new development

of disability for IADL or BADL for the FP; nevertheless, their definition of the disability outcome is different, and the follow-up period is twice ours.

The strengths of our study include the large number of participants aged 75 and older; the longitudinal design of the study; the comprehensive set of measurements performed blinded to outcome status[34]; and the diversity of frailty tools and settings, enabling the construction of a broad and detailed comparison of frailty instruments. We did not expect to find differences in the performance of the instruments by city, so the inclusion of participants from different European cities was done to increase the generalizability of our results, not comparing them.

However, there are limitations to this study. Firstly, the number of participants lost to follow-up is one of the main flaws, with 31% overall at 12 months. While studies in older adults frequently accrue relatively large numbers of drop-outs and lost to follow-up data, we believe that and additional reason accounting for this finding in our study was due to the long administration time to perform the full frailty assessment with all instruments, which caused tiredness and lack of motivation among the participants to respond to the follow-up assessment 12 months later. This happened more frequently among those with a worse health status, what can be guessed from the comparison of age, multimorbidity, dependency, and physical activity between those lost and not lost to follow-up. Lost to follow-up was disproportionately high among participants recruited in geriatric wards and was associated with living alone. This situation probably provoked that these patients were discharged to a place different to the one they dwelled before hospitalization, making difficult to contact them at follow-up at the telephone number provided during the assessment. A reduced sample size may have diminished our power to obtain narrower confidence intervals for the outcomes. It has also prevented us from stratifying analyses by gender and city that could have shown interesting differences and could be the target of new research.

Secondly, we have assessed instruments that were designed under different paradigms, such as the CFS, which was developed for screening purposes, or the FI-35, which was built as a general geriatric assessment tool. However, they are all considered frailty screening/diagnostic tools and are commonly used[35]; many other tools could not be assessed due to limitations in the administration time of the scales.

Thirdly, the study population was a convenience sample of a fairly active group of volunteer elderly adults. Participants with severe cognitive impairment and moderate to high dependency were excluded.

Fourthly, hospitalized patients incapable of making decisions due to delirium or critical illness were excluded. Including a broader range of functional level status might have powered the results.[36,37] In addition, we found more participants with an impairment in their BADL (289–29.0%) than in their IADL dependency (212–21.3%). Two main reasons can account for this apparently paradoxical finding: first, the criteria for defining 'impairment' for BADL and IADL. While the criterion for the impairment in BADL was very tight, in an effort to be as sensitive as possible to detect it but with a lower specificity, the one used for IADL is broadly accepted and probably reflects the true changes. A second potential explanation is the characteristics of the participants in some of the settings of care. Around 50% of the participants were recruited in non-community settings, and indeed many of those coming from community settings attended geriatric clinics. This distribution can explain the high incidence of severe deterioration of function along time, with a more prominent ground effect for IADL dependency impairment. In fact, the difference in the impairment in basic BADL in comparison with the impairment in IADL can be accounted for the participants in geriatric wards and, specially, in nursing homes, while those participants from the outpatient settings did not show such difference.

Lastly, a longer follow-up would have rendered more events and increased power, but in our case meant more missing data.

Conclusions

There is no single frailty instrument that always achieves good enough sensitivity and specificity to detect all adverse outcomes across all settings where older patients are cared for, although some outperform the rest. Mortality and BADL worsening are the outcomes for which some instruments have good enough diagnostic accuracy, although this is mainly restricted to sensitivity in inpatient settings and specificity in geriatric clinic. Stemming from the results of our study, we suggest that patients detected as frail as an inpatient, what may be accomplished with the FI-35, the SHARE-FI, or the FTS-5, should have this suspicion confirmed after discharge.

Future research should focus on exploring different cut-offs for the existing instruments to increase specificity in inpatient settings (the FTS-5 seems a good candidate) and sensitivity in geriatric clinic (the SHARE-FI being a good candidate). In primary care, with the limitations in resources and time of this level of attention, maybe altering the cut-off of the CFS or FRAIL scale could increase sensitivity, but positive patients should be derived to geriatric clinic to confirm the diagnostic suspicion.

Acknowledgements

All authors thank Matteo Tosato, Tania Guevara-Guevara, Jimmy Gonzáles-Turín, Shirley Álamo-Ascencio, and Armando Pardo-Gómez for their contribution to the project and Rich-

ard Mier and Winster Ceballos, who contributed to professional editing in the creation of the manuscript.

Funding

This work was supported by the European Commission Directorate General for Health and Consumer Affairs (DG SANTE) – Third Health Programme. Founding Health Initiatives (2014–2020), and Spanish Ministry of Economy, Industry and Competitiveness, cofinanced by FEDER (RD120001/0043) and CIBERFES (CB16/10/00464). The funding body did not have any role in the design, methods, subject recruitment, data collections, analysis, and preparation of paper.

Conflicts of interest

All authors declare no conflicts of interest. All participants gave informed, written consent prior to their inclusion in the study. The study was approved by the Ethics committee of each participating centre. It has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments.[38]

Trial registration

Comprehensive validation of frailty assessment tools in older adults in different clinical and social settings (FRAILTOOLS), NCT02637518 (date of registration: 12/18/2015). Available from: https://clinicaltrials.gov/ct2/show/NCT02637518

Online supplementary material

Additional supporting information may be found online in the Supporting Information section at the end of the article.

References

1. Reeves D, Pye S, Ashcroft DM, Clegg A, Kontopantelis E, Blakeman T, et al. The challenge of ageing populations and patient frailty: can primary care adapt? *BMJ* 2018;**362**:k3349.
2. Ambagtsheer RC, Shafiabady N, Dent E, Seiboth C, Beilby J. The application of artificial intelligence (AI) techniques to identify frailty within a residential aged care administrative data set. *Int J Med Inform* 2020;**136**:104094.
3. Chen X, Mao G, Leng SX. Frailty syndrome: an overview. *Clin Interv Aging* 2014;**9**:433–441.
4. Abizanda P, Romero L, Sanchez-Jurado PM, Atienzar-Nunez P, Esquinas-Requena JL, Garcia-Nogueras I. Association between functional assessment instruments and frailty in older adults: the FRADEA study. *J Frailty Aging* 2012;**1**:162–168.
5. Romero Rizos L, Abizanda Soler P. Fragilidad como predictor de episodios adversos en estudios epidemiológicos: Revisión de la literatura. *Rev Esp Geriatr Gerontol* 2013;**48**:285–289.
6. Tavassoli N, Guyonnet S, Van Kan GA, Sourdet S, Krams T, Soto M, et al. Description of 1, 108 older patients referred by their physician to the "Geriatric Frailty clinic (G.F.C) for assessment of frailty and prevention of disability" at the gerontopole. *J Nutr Heal Aging* 2014;**18**:457–464.
7. Daniels R, Van Rossum E, Beurskens A, Van Den Heuvel W, De Witte L. The predictive validity of three self-report screening instruments for identifying frail older people in the community. *BMC Public Health* 2012;**12**:69.
8. Warnier RMJ, van Rossum E, van Kuijk SMJ, Mulder WJ, Schols JMGA, Kempen GIJM. The Maastricht frailty screening tool for hospitalised patients (MFST-HP) to identify non-frail patients. *Int J Clin Pract* 2017;**71**:e13003.
9. Pérez-Zepeda MU, Cesari M, García-Peña C. Predictive value of frailty indices for adverse outcomes in older adults. *Rev Invest Clin* 2016;**68**:92–98.
10. Hogan DB, Freiheit EA, Strain LA, Patten S, Schmaltz H, Rolfson D, et al. Comparing frailty measures in their ability to predict adverse outcome among older residents of assisted living. *BMC Geriatr* 2012;**12**: https://doi.org/10.1186/1471-2318-12-56
11. Coelho T, Santos R, Paúl C, Gobbens RJJ, Fernandes L. Portuguese version of the Tilburg frailty indicator: transcultural adaptation and psychometric validation. *Geriatr Gerontol Int* 2015;**15**:951–960.
12. Op het Veld L, Beurskens A, de Vet H, van Kuijk S, Hajema K, Kempen G, 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.
13. Oviedo-Briones M, Laso ÁR, Carnicero JA, Cesari M, Grodzicki T, Gryglewska B, 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.e7–607.e12.
14. Checa-López M, Oviedo-Briones M, Pardo-Gómez A, Gonzales Turin J, Guevara-Guevara T, Carnicero J, et al. FRAILTOOLS study protocol: a comprehensive validation of frailty assessment tools to screen and diagnose frailty in different clinical and social settings and to provide instruments for integrated care in older adults. *BMC Geriatr* 2019;**19**:86.
15. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstem AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;**49**:1373–1379.
16. Fried LP, Tangen CM, Walston J, Newman A, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol Ser A Biol Sci Med Sci* 2001;**56**:M146–M156.
17. Cesari M, Gambassi G, Van Kan GA, Vellas B. The frailty phenotype and the frailty index: different instruments for different purposes. *Age Ageing* 2014;**43**:10–12.
18. Hoogendijk EO, van Kan GA, Guyonnet S, Vellas B, Cesari M. Components of the frailty phenotype in relation to the frailty index: results from the Toulouse frailty platform. *J Am Med Dir Assoc* 2015;**16**:855–859.
19. Romero-Ortuno R, Walsh CD, Lawlor BA, Kenny RA. A frailty instrument for primary care: findings from the Survey of Health, Ageing and Retirement in Europe (SHARE). *BMC Geriatr* 2010;**10**:57.
20. García-García FJ, Carcaillon L, Fernandez-Tresguerres J, Alfaro A, Larrion J, Castillo C, et al. A new operational definition of frailty: the frailty trait scale. *J Am Med Dir Assoc* 2014;**15**:371.e7–371.e13.
21. García-García FJ, Carnicero JA, Losa-Reyna J, Alfaro-Acha A, Castillo-Gallego C, Rosado Artalejo C, et al. Frailty trait scale–short form: a frailty instrument for clinical practice. *J Am Med Dir Assoc* 2020;Jan 28:pii: S1525-8610(19)30868-0;**21**:1260–1266.e2.

22. Abellan Van Kan G, Rolland Y, Bergman H, Morley JE, Kritchevsky SB, Vellas B. The I. A.N.A. task force on frailty assessment of older people in clinical practice. *J Nutr Heal Aging.* 2008;**12**:29–37.
23. Valdatta L, Perletti G, Maggiulli F, Tamborini F, Pellegatta I, Cherubino M. FRAIL scale as a predictor of complications and mortality in older patients undergoing reconstructive surgery for non-melanoma skin cancer. *Oncol Lett* 2019;**17**:263–269.
24. Dent E, Kowal P, Hoogendijk EO. Frailty measurement in research and clinical practice: a review. *Eur J Intern Med* 2016;**31**: 3–10.
25. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. *BMC Geriatr* 2008;**8**:24.
26. Theou O, Rockwood MRH, Mitnitski A, Rockwood K. Disability and co-morbidity in relation to frailty: how much do they overlap? *Arch Gerontol Geriatr* 2012;**55**: e1–e8.
27. Vellas B, Balardy L, Gillette-Guyonnet S, Abellan Van Kan G, Ghisolfi-Marque A, Subra J, et al. Looking for frailty in community-dwelling older persons: the Gerontopole Frailty Screening Tool (GFST). *J Nutr Heal Aging.* 2013;**17**:629–631.
28. Rockwood K, Song X, MacKnight C, Bergman H, Hogan D, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005;**30** (173:489–495.
29. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;**3**:179–186.
30. Mahoney FI, Barthel DW. Funtional evaluation: the Barthel index. *Md State Med J* 1965;**14**:61–65.
31. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–383.
32. Washburn RA, Smith KW, Jette AM, Janney CA. The physical activity scale for the elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993;**46**:153–162.
33. Bongue B, Buisson A, Dupre C, Beland F, Gonthier R, Crawford-Achour É. Predictive performance of four frailty screening tools in community-dwelling elderly. *BMC Geriatr* 2017;**17**:262.
34. Ensrud KE, Ewing SK, Taylor BC, Fink H, Stone K, Cauley J, et al. Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women. *Arch Intern Med* 2008;**168**: 382–389.
35. Rodríguez-Laso Á, Caballero-Mora MA, García Sanchez I, Rodríguez Manas L, Bernabei R, Gabrovec B, Rodríguez Mañas L, García-Sánchez I, Hendry A, Roller-Wirnsberger R, Liew A, Carriazo AM, Redon J, Galluzzo L, Viña J, Antoniadou E, Targowski T, di Furia L, Lattanzio F, Bozdog E & Telo M State of the art report on the prevention and management of frailty. http://www.advantageja.eu/images/SoAR-AdvantageJA_Full-text.pdf. Published 2018. Accessed March 22, 2020. 892 897
36. 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.
37. Pritchard J, Kennedy C, Karampatos S, Loannidis G, Misiaszek B, Marr S, et al. Measuring frailty in clinical practice: a comparison of physical frailty assessment methods in a geriatric out-patient clinic. *BMC Geriatr* 2017;**17**:264.
38. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2021. *J Cachexia Sarcopenia Muscle* 2021;**12**:2259–2261.