

# Diagnostic accuracy of the FRAIL scale plus functional measures for frailty screening: a cross-sectional study

Ángel Rodríguez-Laso<sup>1</sup>, Iñaki Martín-Lesende<sup>2</sup>, Alan Sinclair<sup>3,4</sup>, Sandrine Sourdet<sup>5</sup>, Matteo Tosato<sup>6</sup>, Leocadio Rodríguez-Mañas<sup>1\*</sup>

<sup>1</sup>CIBERFES (Network-based Biomedical Research Consortium, area of Frailty and Healthy Ageing), Instituto de Salud Carlos III, Madrid, Spain; <sup>2</sup>Indautxu Primary Health Centre, Bilbao-Basurto Integrated Health Organisation, Basque Health Service (Osakidetza), Bilbao, Spain; <sup>3</sup>Foundation for Diabetes Research in Older People (fDROP), London, UK; <sup>4</sup>Medical Sciences, Nursing and Midwifery, Kings College, London, UK; <sup>5</sup>Department of Internal Medicine and Geriatrics, Toulouse University Hospital, Toulouse, France; <sup>6</sup>Fondazione Policlinico Universitario Agostino Gemelli IRCSS, Rome, Italy

## Abstract

**Background:** There is little knowledge of the diagnostic accuracy of screening programmes for frailty in primary care settings.

**Aim:** To assess a two-step strategy consisting of the administration of the FRAIL scale to those who are non-dependent and aged  $\geq 75$  years, followed-up by measurement of the Short Physical Performance Battery (SPPB) or gait speed in those who are positive.

**Design & setting:** Cross-sectional and longitudinal cohort study. Analysis of primary care data from the FRAILTOOLS project at five European cities.

**Method:** All primary care patients consecutively attending were enrolled. They received the index tests, plus the Fried frailty phenotype (FP) and the frailty index to assess their frailty status. Mortality and worsening of dependency in basic and instrumental activities of daily living (BADL and IADL) over 1 year were ascertained.

**Results:** Prevalence of frailty based on FP was 14.9% in the 362 participants. A FRAIL scale score  $\geq 1$  had a sensitivity of 83.3% (95% confidence interval [CI] = 73.1 to 93.6) to detect frailty. A positive result and an SPPB score  $< 11$  had a sensitivity of 72.2% (95% CI = 59.9 to 84.6); when combined with a gait speed  $< 1.1$  m/s, the sensitivity was 80.0% (95% CI = 68.5 to 91.5). Two-thirds of those screened as positive were not frail. In the best scenario, sensitivities of this last combination to detect IADL and BADL worsening were 69.4% (95% CI = 59.4 to 79.4) and 63.6% (95% CI = 53.4 to 73.9), respectively.

**Conclusion:** Combining the FRAIL scale with other functional measures offers an acceptable screening approach for frailty. Accurate prediction of worsening dependency and death need to be confirmed through the piloting of a frailty screening programme.

\*For correspondence: leocadio.rodriguez@salud.madrid.org

**Competing interest:** The authors declare that no competing interests exist.

**Received:** 18 November 2021

**Accepted:** 24 January 2022

**Published:** 24 August 2022

©This article is Open Access: CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

**Author Keywords:** primary health care, frailty, frailty phenotype, mass screening

Copyright © 2022, The Authors; DOI:10.3399/BJGPO.2021.0220

## How this fits in

Busy primary care clinicians need a quick and easy-to-administer tool for frailty screening in older patients, whose positive results may ideally be confirmed with functional performance measures. There is no information on the diagnostic accuracy of such a strategy in primary care. The administration of the FRAIL scale plus the SPPB or the measurement of gait speed to non-dependent individuals aged

≥75 years had a good sensitivity to detect frailty, with an acceptable rate of false positives. Results suggest that it can also predict worsening of dependency in a 1-year period.

## Introduction

Frailty in older people is a progressive age-related decline in physiological systems resulting in decreased reserves of intrinsic capacity, extreme vulnerability to stressors, and increased risk of adverse health outcomes.<sup>1</sup> Screening for this very common condition (35% of patients aged ≥70 years attend primary care in Europe)<sup>2</sup> allows early detection and intervention before consequences occur, such as disability, which is more difficult to reverse. There is evidence on the validity, reliability, and feasibility of several tools to perform the screening and on the efficacy of interventions to reverse frailty, mainly multicomponent exercise.<sup>3</sup>

Several countries and regions have deployed screening programmes in primary care with different instruments that generate variable workloads for primary care teams.<sup>3,4</sup> The usual limitation of attention time in these practices combined with possible limitations to face-to-face contact, like those brought about by the COVID-19 pandemic, suggest the need for a screening instrument that can be administered quickly and on the phone. The FRAIL scale<sup>5</sup> meets these requirements and could be combined with performance tests, such as the SPPB<sup>6</sup> or the measurement of gait speed, to confirm positive results. These are tools recommended by the ADVANTAGE Joint Action.<sup>3</sup>

To the authors' knowledge, there is no published evidence on the diagnostic accuracy of this strategy for frailty, worsening of dependency, or death screening. The objective of this study was to evaluate the sensitivity, specificity, and predictive values for this strategy. Different cut-offs for the three instruments were explored.

## Method

This article adheres to the STARD guidelines for reporting diagnostic accuracy studies.<sup>7</sup>

The design and rationale of the FRAILTOOLS project have been previously published in more detail.<sup>8</sup> It was an observational, prospective, and longitudinal study planned to explore the diagnostic accuracy of several frailty instruments. It enrolled consecutively 1440 adults (aged ≥75 years) from primary care clinics, geriatric medicine services, and nursing homes from France (Toulouse), Italy (Rome), Poland (Cracow), Spain (Getafe), and the UK (Birmingham). Exclusion criteria were as follows: a Mini-Mental State Examination ≤20 points; a terminal illness (life expectancy ≤6 months); and a Barthel Index <90. Variables were collected at baseline in 2016 and at 6-, 12-, and 18-months' follow-up. This article is limited to the 381 primary care patients and their 1-year follow-up. The first five patients with no exclusion criteria attending primary care practices each morning were selected. These practices were mainly those that volunteered to participate among those that referred patients to the principal investigators' affiliation hospitals.

Information on age, sex, multimorbidity,<sup>9</sup> and several frailty instruments was collected at baseline. This article focuses on the following index tests and frailty measures.

The FP<sup>10</sup> constitutes one of the reference standards because of its general acceptance as a measure of frailty.<sup>3</sup> It consists of three self-reported components (exhaustion, physical activity, and weight loss), and two objective measures (grip strength and gait speed). Exhaustion was considered present when the responder answered at least 3–4 days during the past week to any of the following two questions from the Center for Epidemiological Studies Depression (CES-D) scale:<sup>11</sup> 'I felt that anything I did was a big effort' and 'I felt that I could not keep on doing things'. The physical activity item was considered present when males referred <2.5 hours walking per week (equivalent to <383 kcal) and women <2 hours (<270 kcal) usually. The weight loss item was considered positive if there was an unintentional loss of at least 4.5 kg in a year. Grip strength was measured as the best of three trials with a Jamar hydraulic dynamometer in the dominant hand. Gait speed was measured as the best of two trials at usual pace in a 4.5-metre distance from a standing position without using assisting devices. Both items were considered positive when the individual was in the worst quintile of strata of sex and body mass index for grip strength, and sex and height for gait speed.<sup>12</sup> A patient was considered frail if ≥3 criteria were positive, even if the rest of the items were not measured. No imputation of missing items was performed.

The 35-item Frailty Index (FI-35) was the second reference standard because it belongs to the second recognised conceptualisation of frailty as an accumulation of deficits.<sup>13</sup> It was calculated as the proportion of a list of health deficits (that is, symptoms, signs, chronic diseases, disability, and laboratory abnormalities) the patient suffered from,<sup>14</sup> obtained from medical records, self-reported, or measured at the patient's evaluation. The cut-off used to identify frailty was set to  $\geq 0.25$ .<sup>15</sup> According to the original protocol, the FI-35 allowed a missingness up to 20% of items to calculate the score and be able to classify the patient as frail.

The FRAIL scale, one of the index tests, comprises five self-reported items: Fatigue, Resistance, Ambulation, Illness, and Loss of weight.<sup>5</sup> Although the recommended cut-off is  $\geq 3$ , this article also explores lower ones. Any individual with any item lost was excluded from analyses.

Two performance measures, the two other index tests, were administered: the SPPB,<sup>6</sup> which is a scale that ranges from 0–12 and combines three tests, gait speed, time to perform five chair stands, and balance assessment (in three positions: feet together, semi-tandem, and tandem); and gait speed (measured like the FP's item).

For assessing predictive validity, the following frailty adverse outcomes were employed: 1) IADL and BADL dependency worsening at follow-up, which were defined as a loss of one point in the Lawton and Brody<sup>16</sup> or a loss  $\geq 5$  points in the Barthel<sup>17</sup> indexes, respectively; 2) death, which was ascertained through phone calls to arrange follow-up visits at 6 and 12 months and, when no answer was obtained, hospital registries (plus the death registry of the Ministry of Health in Spain).

All members of the research healthcare team (nurses and geriatricians) of all countries received the same training on the administration of the scales.<sup>8</sup> All participants gave informed, written consent.

## Statistical analysis

Description of variables was carried out with medians and interquartile ranges (IQRs) or absolute and relative frequencies. Sensitivity, specificity, and percentage of positives who are not frail (false positives, understood as the complementary of the positive predictive value, not of the specificity) for frailty defined by the FP and the FI-35, and the three adverse outcomes were calculated for different thresholds of the FRAIL scale, and for different scores of the SPPB and increments of 0.1 m/s of gait speed in individuals with a FRAIL score  $\geq 1$ .

The R package (version 4.0.2) was used for all analyses.

## Results

Out of the 381 primary care patients, 19 did not provide information on the index tests or frailty measures. There were more females, although the difference with participants was not statistically significant ( $P = 0.407$ ). There were no differences in age ( $P = 0.49$ ), Charlson Comorbidity Index ( $P = 0.323$ ), nor Lawton and Brody scale ( $P = 0.437$ ) and Barthel index ( $P = 0.326$ ). The final sample size for the diagnostic accuracy of frailty was 362.

Characteristics of the sample for analyses are presented in **Table 1**. Median age was 79 years (IQR 77–82) and 58.8% were female. Three cities (Getafe, Toulouse, and Rome) contributed 79.2% of the sample. The median Charlson Comorbidity Index was high (median = 4; IQR 4–5). Dependency at baseline was infrequent, but around 17% became dependent after 1 year (loss to follow-up for this variable amounted to 24.0%). The prevalence of frailty was 14.9% (95% CI = 11.2 to 18.6) and 15.2% (95% CI = 11.5 to 18.9) according to the FP and the FI-35, respectively. Median SPPB score and gait speed were 10 and 1 m/s, respectively. Deaths were an extremely infrequent outcome, although missing information was the highest for this variable.

**Table 2** presents the prevalence and diagnostic accuracy of the FRAIL scale for three cut-off points. The traditional cut-off point ( $\geq 3$ ) had a very low sensitivity for detecting frailty according to the FP (37.0%, 95% CI = 23.7 to 50.3). Decreasing the cut-off by one point rendered a higher sensitivity of 66.7% (95% CI = 53.7 to 79.7). Scoring any item of the FRAIL scale had a sensitivity of 83.3% (95% CI = 73.1 to 93.6). Indicators were slightly worse for frailty operationalised as the FI-35. Of the sample, 42.3% (95% CI = 37.2 to 47.4) scored at least one item of the FRAIL scale and therefore would be offered to be screened with the performance tests.

**Table 3** presents the prevalence and diagnostic accuracy of the different scores of the SPPB among those with a score in the FRAIL scale  $\geq 1$ . Percentages refer to the whole sample, not only those with

**Table 1** Characteristics of the sample for the cross-sectional analysis (*n* = 362)

Category	Variable	<i>n</i> (%) <sup>a</sup>
Age, years, median (IQR)		79 (77–82)
Sex, female		213 (58.8)
City	Getafe (Spain)	116 (32.0)
	Toulouse (France)	93 (25.7)
	Rome (Italy)	78 (21.5)
	Cracow (Poland)	48 (13.3)
	Birmingham (UK)	27 (7.5)
Charlson Comorbidity Index score, median (IQR)		4 (4–5)
Barthel index score, median (IQR)		100 (95–100)
Lawton and Brody scale score, median (IQR)		8 (7–8)
Frail according to frailty phenotype <sup>b</sup>		54 (14.9)
Frail or prefrail according to frailty phenotype <sup>b</sup>	Total	276 (76.2)
	Missing	4 (1.1)
Number of items of the frailty phenotype <sup>b</sup>	0	82 (22.7)
	1	112 (30.9)
	2	104 (28.7)
	3	33 (9.1)
	4	17 (4.7)
	5	1 (0.3)
	Missing	13 (3.6)
Frailty index score, median (IQR)		0.16 (0.12–0.22)
Frail according to frailty index		55 (15.2)
Number of items of the FRAIL scale	0	209 (57.7)
	1	85 (23.5)
	2	38 (10.5)
	3	25 (6.9)
	4	4 (1.1)
	5	1 (0.3)
SPPB score, median (IQR)		10 (9–11)
Gait speed, m/s	Total, median (IQR)	1 (0.9–1.3)
	Missing	18 (5.0)
Worsening of dependence in basic activities of daily living	Total	46 (16.7) <sup>c</sup>
	Missing	87 (24.0)
Worsening of dependence in instrumental activities of daily living	Total	47 (17.1) <sup>c</sup>
	Missing	87 (24.0)
Deaths	Total	2 (0.8) <sup>d</sup>
	Missing	107 (29.6)

<sup>a</sup>Unless otherwise stated. <sup>b</sup>These numbers do not match because of the possibility of assigning an individual with missing items to the frail or frail+ prefrail categories. <sup>c</sup>*N* = 275. <sup>d</sup>*N* = 255. IQR = interquartile range. SPPB = Short Physical Performance Battery.

**Table 2** Prevalence and diagnostic accuracy for frailty of different cut-offs of the FRAIL scale

FRAIL scale	Prevalence (95% CI)	Sensitivity (95% CI)		Specificity (95% CI)		% of positives who were not frail (95% CI)	
		Fried's frailty phenotype	Frailty index	Fried's frailty phenotype	Frailty index	Fried's frailty phenotype	Frailty index
		≥3	8.3 (5.4 to 11.1)	37.0 (23.7 to 50.3)	23.6 (12.0 to 35.2)	96.8 (94.8 to 98.7)	94.5 (91.9 to 97.0)
≥2	18.8 (14.7 to 22.8)	66.7 (53.7 to 79.7)	40.0 (26.6 to 53.4)	89.6 (86.2 to 93.0)	85.0 (81.0 to 89.0)	47.1 (34.9 to 59.2)	67.6 (56.2 to 79.1)
≥1	42.3 (37.2 to 47.4)	83.3 (73.1 to 93.6)	74.5 (62.7 to 86.4)	64.9 (59.6 to 70.3)	63.5 (58.1 to 68.9)	70.6 (63.3 to 77.9)	73.2 (66.1 to 80.3)

a FRAIL score ≥1. A cut-off score of 10 yielded a sensitivity of 68.5% (95% CI = 55.7 to 81.3) to detect frailty operationalised with the FP. Raising the cut-off by one point increased the sensitivity to 72.2% (95% CI = 59.9 to 84.6) at the cost of increasing the proportion of false positives from 58.4% (95% CI = 48.0 to 68.9) to 66.7% (95% CI = 58.0 to 75.3). When the condition to be screened was defined as frailty according to the FI-35, the cut-off to obtain a similar sensitivity was <12. Using a cut-off of <11, 32.3% (95% CI = 27.5 to 37.2) of the total eligible population would be referred to a multidimensional evaluation.

**Table 4** presents the same structure as **Table 3** but refers to different gait speeds. Sample size was smaller (N = 342) because in some cases gait speed was ascertained for the SPPB scoring but was not recorded in m/s. A FRAIL score of ≥1 plus a gait speed <0.8 m/s showed a sensitivity for frailty of only 52.0% (95% CI = 37.7 to 66.3) for the FP and 34.7% (95% CI = 20.9 to 48.5) for the FI-35. Sensitivity reached 74.0% at a cut-off <1 m/s and got higher at the expense of a little increase of false positives

**Table 3** Prevalence and diagnostic accuracy for frailty of different SPPB scores in individuals with a FRAIL score ≥1

SPPB score	Prevalence (95% CI)	Sensitivity (95% CI)		Specificity (95% CI)		% of positives who were not frail (95% CI)	
		Fried's frailty phenotype	Frailty index	Fried's frailty phenotype	Frailty index	Fried's frailty phenotype	Frailty index
		<2	0.6 (0.0 to 1.3)	3.7 (0.0 to 8.9)	3.6 (0.0 to 8.7)	100.0	100.0
<3	0.8 (0.0 to 1.8)	5.6 (0.0 to 11.9)	3.6 (0.0 to 8.7)	100.0	99.7 (99.0 to 100.0)	0.0	33.3 (0.0 to 100.0)
<4	1.1 (0.0 to 2.2)	7.4 (0.2 to 14.6)	5.5 (0.0 to 11.7)	100.0	99.7 (99.0 to 100.0)	0.0	25.0 (0.0 to 100.0)
<5	1.9 (0.5 to 3.4)	11.1 (2.5 to 19.8)	7.3 (0.2 to 14.4)	99.7 (99.0 to 100.0)	99.0 (97.9 to 100.0)	14.3 (0.0 to 49.2)	42.9 (0.0 to 92.3)
<6	3.0 (1.3 to 4.8)	18.5 (7.8 to 29.2)	10.9 (2.4 to 19.4)	99.7 (99.0 to 100.0)	98.4 (96.9 to 99.8)	9.1 (0.0 to 29.3)	45.5 (10.4 to 80.5)
<7	5.5 (3.2 to 7.9)	31.5 (18.7 to 44.3)	21.8 (10.6 to 33.1)	99.0 (97.9 to 100.0)	97.4 (95.6 to 99.2)	15.0 (0.0 to 32.1)	40.0 (16.5 to 63.5)
<8	7.2 (4.5 to 9.9)	38.9 (25.5 to 52.3)	23.6 (12.0 to 35.2)	98.4 (97 to 99.8)	95.8 (93.5 to 98.0)	19.2 (3.0 to 35.5)	50.0 (29.4 to 70.6)
<9	10.2 (7.1 to 13.4)	46.3 (32.6 to 60.0)	30.9 (18.3 to 43.5)	96.1 (93.9 to 98.3)	93.5 (90.7 to 96.3)	32.4 (16.6 to 48.3)	54.1 (37.2 to 70.9)
<10	15.7 (12.0 to 19.5)	55.6 (41.9 to 69.2)	40.0 (26.6 to 53.4)	91.2 (88.1 to 94.4)	88.6 (85.0 to 92.2)	47.4 (34.0 to 60.7)	61.4 (48.4 to 74.4)
<11	24.6 (20.1 to 29.0)	68.5 (55.7 to 81.3)	56.4 (42.8 to 69.9)	83.1 (78.9 to 87.3)	81.1 (76.7 to 85.5)	58.4 (48.0 to 68.9)	65.2 (55.1 to 75.3)
<12	32.3 (27.5 to 37.2)	72.2 (59.9 to 84.6)	67.3 (54.5 to 80.1)	74.7 (69.8 to 79.6)	73.9 (69.0 to 78.9)	66.7 (58.0 to 75.3)	68.4 (59.8 to 76.9)
<13	38.4 (33.4 to 43.4)	75.9 (64.1 to 87.7)	72.7 (60.6 to 84.9)	68.2 (63.0 to 73.4)	67.8 (62.5 to 73.0)	70.5 (62.8 to 78.2)	71.2 (63.6 to 78.8)

Percentages refer to the total sample (n = 362), not only to individuals with a FRAIL score ≥1. SPPB = Short Physical Performance Battery.

**Table 4** Prevalence and diagnostic accuracy for frailty of different gait speeds in individuals with a FRAIL score  $\geq 1$ 

Gait speed, m/s	Prevalence (95% CI)	Sensitivity (95% CI)		Specificity (95% CI)		% of positives who were not frail (95% CI)	
		Fried's frailty phenotype	Frailty index	Fried's frailty phenotype	Frailty index	Fried's frailty phenotype	Frailty index
<0.6	6.7 (4.1 to 9.4)	36.0 (22.2 to 49.8)	24.5 (12.0 to 37.0)	98.3 (96.8 to 99.8)	96.2 (94.1 to 98.4)	21.7 (3.5 to 40.0)	47.8 (25.7 to 69.9)
<0.7	9.6 (6.5 to 12.8)	40.0 (25.9 to 54.1)	24.5 (12.0 to 37.0)	95.5 (93.2 to 97.9)	92.8 (89.9 to 95.8)	39.4 (21.8 to 57.0)	63.6 (46.3 to 81.0)
<0.8	16.4 (12.4 to 20.3)	52.0 (37.7 to 66.3)	34.7 (20.9 to 48.5)	89.7 (86.2 to 93.2)	86.7 (82.8 to 90.6)	53.6 (40.1 to 67.0)	69.6 (57.2 to 82.1)
<0.9	23.7 (19.2 to 28.2)	62.0 (48.1 to 75.9)	57.1 (42.8 to 71.5)	82.9 (78.5 to 87.2)	81.9 (77.5 to 86.3)	61.7 (50.9 to 72.5)	65.4 (54.9 to 76.0)
<1	33.0 (28.0 to 38.1)	74.0 (61.4 to 86.6)	69.4 (56.0 to 82.8)	74.0 (68.9 to 79.0)	73.0 (67.9 to 78.1)	67.3 (58.5 to 76.0)	69.9 (61.3 to 78.5)
<1.1	36.3 (31.1 to 41.4)	80.0 (68.5 to 91.5)	71.4 (58.3 to 84.5)	71.2 (66.0 to 76.5)	69.6 (64.3 to 74.9)	67.7 (59.4 to 76.1)	71.8 (63.7 to 79.8)
<1.2	39.5 (34.3 to 44.7)	82.0 (71.0 to 93.0)	73.5 (60.7 to 86.3)	67.8 (62.4 to 73.2)	66.2 (60.8 to 71.7)	69.6 (61.8 to 77.5)	73.3 (65.8 to 80.9)
<1.3	41.5 (36.3 to 46.8)	82.0 (71.0 to 93.0)	75.5 (63.0 to 88.0)	65.4 (59.9 to 70.9)	64.2 (58.6 to 69.7)	71.1 (63.6 to 78.7)	73.9 (66.6 to 81.3)
<1.4	42.1 (36.8 to 47.4)	82.0 (71.0 to 93.0)	75.5 (63.0 to 88.0)	64.7 (59.2 to 70.2)	63.5 (57.9 to 69.0)	71.5 (64.1 to 79.0)	74.3 (67.1 to 81.5)
<1.5	42.4 (37.1 to 47.7)	82.0 (71.0 to 93.0)	75.5 (63.0 to 88.0)	64.4 (58.9 to 69.9)	63.1 (57.6 to 68.7)	71.7 (64.3 to 79.1)	74.5 (67.3 to 81.7)

Percentages refer to the total sample ( $n = 342$ ), not only to individuals with a FRAIL score  $\geq 1$ .

at <1.1 m/s (sensitivity 80.0%, 95% CI = 68.5 to 91.5; false positives 67.7%, 95% CI = 59.4 to 76.1). Of the sample, 36.3% (95% CI = 31.1 to 41.4) walked at this speed. A cut-off of <1.2 m/s increased sensitivity and false positives by 2%. For the FI-35, starting from <1 m/s increments in sensitivity and false positives were parallel.

Out of the 362 individuals in the cross-sectional analyses, 87 could not be assessed for their final dependency status, rendering a sample size of 275. Compared with those assessed, those lost were 1-year older ( $P = 0.037$ ). Their Charlson Comorbidity Index ( $P = 0.124$ ), Lawton and Brody scale ( $P = 0.319$ ), and Barthel ( $P = 0.735$ ) baseline indexes were similar, but frailty according to the FP was more common (21% versus 13%,  $P = 0.049$ ) and the average number of positive items of FRAIL scale mildly higher (0.23;  $P = 0.059$ ).

A score of the FRAIL scale  $\geq 1$  had a sensitivity of 52.2% (95% CI = 37.2 to 67.2) and 46.8% (95% CI = 32.0 to 61.6) to detect a 1-year worsening of BADLs and IADLs, respectively. Combining it with the SPPB or gait speed, sensitivities would get even lower.

Sensitivity analysis was performed by handling missing data in the worsening dependency variables by the best-case imputation, where missing cases were considered to have worsened their dependency if their FRAIL scale score was >0 at baseline, and not considered to have worsened their dependence if their FRAIL scale score was 0. Under this assumption, FRAIL score  $\geq 1$  sensitivities increased to 76.1% (95% CI = 67.2 to 85.0) for BADL and 73.1% (95% CI = 63.9 to 82.3) for IADL. A FRAIL score  $\geq 1$  plus SPPB <11 had a sensitivity of 60.9% (95% CI = 50.7 to 71.0) and false positives of 52.1% (95% CI = 43.0 to 61.3) to predict worsening of BADL, and of 59.1% (95% CI = 49.0 to 69.3) and 53.0% (95% CI = 43.8 to 62.2) to predict worsening of IADL. A FRAIL score  $\geq 1$  plus gait speed <1.1 m/s had a sensitivity of 63.6% (95% CI = 53.4 to 73.9) and false positives of 54.8% (95% CI = 46.0 to 63.7) to predict worsening of BADL, and of 69.4% (95% CI = 59.4 to 79.4) and 52.4% (95% CI = 43.5 to 61.3) to predict worsening of IADL. Supplementary Tables S1 and S2 present results for other cut-offs.

Death could not be ascertained in 107 people of the cross-sectional sample, which left a sample size for the diagnostic accuracy of mortality of 255 individuals. Two deaths occurred during the follow-up, both with a FRAIL score of 1, SPPB scores of 6 and 9, and gait speeds of 0.79 m/s and 0.67 m/s, respectively. That means that the sensitivity for death of a FRAIL score  $\geq 1$  and SPPB <11 or any of

the proposed cut-offs for gait speed was 100%. Nevertheless, 97.2% (95% CI = 93.3 to 100) of those considered positive under the former criteria did not die.

## Discussion

### Summary

This article shows that a strategy, which screens all non-dependent adults aged  $\geq 75$  years in primary care with the FRAIL scale and follows-up positive results with the SPPB or measurement of gait speed, has a reasonable diagnostic accuracy for frailty detection when the following thresholds are applied: a positive answer to any of the items of the FRAIL scale (instead of the recommended cut-off of  $\geq 3$  items), plus an SPPB score  $< 11$  or gait speed  $< 1.1$  m/s (instead of the usual threshold of 0.8 m/s). The results suggest that this strategy may also predict those who worsen their dependency level in 1-year's time, although cautiousness in the interpretation is warranted because of losses to follow-up.

### Strengths and limitations

The article has the strength of presenting the diagnostic accuracy results of a previously untested frailty screening strategy carried out in a multi-country sample of primary care patients. Its main limitation is loss to follow-up, which did not allow the authors to obtain conclusive results on prediction of dependency worsening because of the discrepancies in results when non-imputing and imputing by the best-case approach. It is known that lost-to-follow-up individuals were older and marginally frailer, which limits the capacity of generalising the results to all eligible users. The authors believe most of these individuals probably dropped out from the study because of tiredness and loss of motivation owing to the long administration time to perform the full frailty assessment in FRAILTOOLS with seven instruments. Another limitation is a short follow-up to detect deaths, but extending analyses to 18 months, as established in the FRAILTOOLS protocol, would have increased missingness.

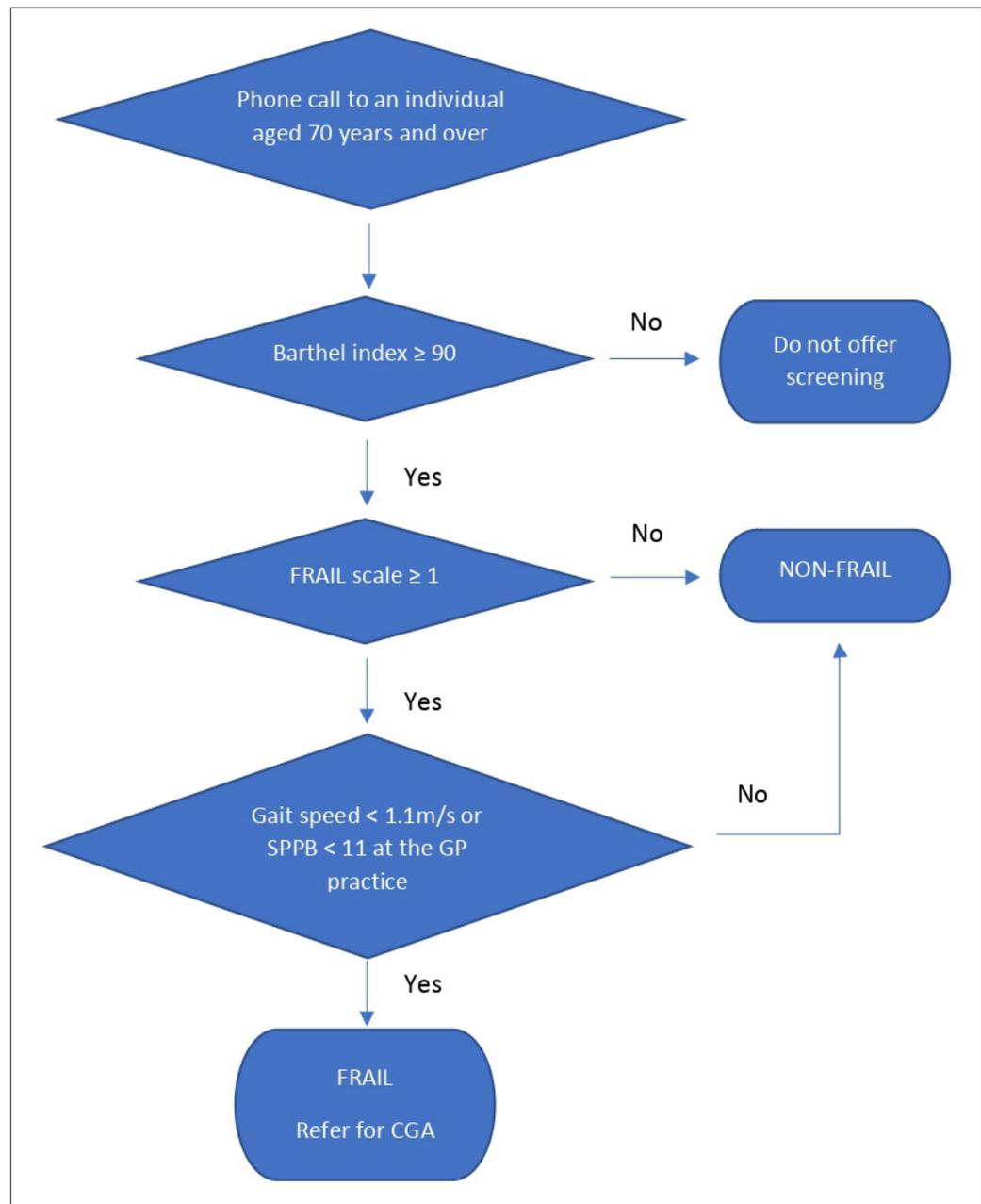
### Comparison with existing literature

Our results of the sensitivity of the FRAIL scale are comparable to those published by Ambagtsheer *et al*<sup>18</sup> from a study in primary care patients aged  $\geq 75$  years in southern Australia, where they reported a sensitivity of 30% (95% CI = 16.6 to 46.5) and 19.8% (95% CI = 12.9 to 28.5) for the FP and the frailty index, respectively, with a cut-off of  $\geq 3$ . They are lower than those reported in eastern China, where sensitivity results for the FP in community dwellers aged  $\geq 60$  years were 52.2% for a score  $\geq 3$ , 87% for a score  $\geq 2$ , and 97.8% for a score  $\geq 1$ .<sup>19</sup> Similar results to Dong *et al*<sup>19</sup> were reported by Thompson *et al*<sup>20</sup> for the FP in community dwellers from the northwest of Australia aged  $\geq 65$  years. In relation to prediction of disability worsening, Si *et al*<sup>21</sup> found a sensitivity of a score of the FRAIL scale  $\geq 3$  at 1-year follow-up of 11.7% for BADL and 9.9% for IADL in community dwellers aged  $\geq 60$  years from a Chinese city. They did not offer results for lower cut-off points.

Although the SPPB has been used for screening in primary care,<sup>22</sup> data have not been recorded about its ability to detect frailty nor adverse outcomes in this level of attention. Ambagtsheer *et al*<sup>18</sup> also studied the diagnostic accuracy of gait speed  $\leq 0.8$  m/s (at four-metre distance). Sensitivity against the FP and the frailty index was 70.0% (95% CI = 53.5 to 83.4) and 47.8% (95% CI = 38.2 to 57.4), respectively, and specificity 77.1% (95% CI = 70.5 to 82.9) and 84.6% (95% CI = 76.8 to 90.6), respectively. Our sensitivities are lower and specificities higher because we added the requirement of having a positive answer in any of the items of the FRAIL scale.

### Implications for research and practice

The use of the FRAIL scale as a screening tool for frailty has many advantages for busy primary care clinics; for example, it has a short administration time (less than a minute and a half in most cases),<sup>23</sup> requires little training or instruction for the assessor, and can be delivered over the phone. The authors recommend using it in the context of the algorithm presented in **Figure 1**. Although the results were limited to individuals aged  $\geq 75$  years, existing recommendations have been adhered to<sup>3,4</sup> and all individuals aged  $\geq 70$  years have been considered. Any positive answer to the items of the FRAIL scale over the phone in a non-dependent patient should elicit an in-person consultation where either the SPPB or gait speed would be measured. The authors predict that less than half of the screened population will require to be referred to functional assessment, which would certainly reduce primary



**Figure 1** Recommended frailty screening algorithm in primary care. CGA = comprehensive geriatric assessment. SPPB = Short Physical Performance Battery.

care teams' workload compared with the assessment of all individuals with functional measures (as recommended in other screening programmes).<sup>24</sup> Positive results in functional assessments should be confirmed through a comprehensive geriatric assessment (CGA) carried out in the primary or secondary level of attention, something that will be required by around one-third of the eligible population. The CGA should encompass the prescription of a multicomponent exercise intervention in confirmed cases.<sup>3</sup> Cut-offs that the authors consider acceptable considering the sensitivity, proportion of false positives, and workload that they would produce are suggested here. These decisions should be tempered by the resources available to carry out functional measurements and CGAs.

To increase certainty on the ability of this strategy to predict death and dependency worsening, a pilot programme with usual primary care users followed for a longer period is warranted.

NHS England opted for detecting frailty in primary care following the accumulation of deficits paradigm using electronic medical records.<sup>25</sup> This would be equivalent in the present study to just

administering the FI-35 to the whole sample. This is not a screening strategy, but a diagnostic one, because all cases according to this definition of frailty would be detected. Curiously enough, NHS England states nevertheless that '*confirmation of frailty in an individual should be undertaken using a validated tool such as: [the] Gait Speed Test*'.<sup>26</sup> They also state, in another document:<sup>25</sup> '*... a clinician from the primary care team should verify the frailty diagnosis by direct assessment using the Clinical Frailty Scale (CFS) or similar validated tool*'. Both instruments are considered screening tests for frailty, not diagnostic tools.<sup>3</sup> In any case, the FP and index approaches have different purposes and are to be considered complementary in the evaluation of the older person.<sup>27</sup> One of their main differences is that the frailty index includes diseases, disability, and dependency items, while the FP was conceived as a measure of a condition that usually precedes disability, and because of that it is based on assessing performance-based tasks, which are different from disability. It has been shown that the algorithm is more sensitive to detect frailty according to the FP rather than the frailty index, and therefore more suitable to identify patients at risk of developing disability.

### Funding

This study was supported by the European Commission Directorate General for Health and Consumer Affairs — Third Health Programme, Funding Health Initiatives (2014–2020); Consorcio Centro de Investigación Biomédica en Red, Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación; and FEDER funds.

### Ethical approval

This study was approved by the ethics committee of each participating centre. The comprehensive validation of frailty assessment tools in older adults in different clinical and social settings (FRAIL-TOOLS) project was registered at <http://www.clinicaltrials.org> (reference: NCT02637518; date of registration: 18 December 2015).

### Provenance

Freely submitted; externally peer reviewed.

### Acknowledgements

The authors would like to acknowledge all patients who provided information for this study.

## References

1. World Health Organization. World report on ageing and health. 2015. [https://apps.who.int/iris/bitstream/handle/10665/186463/9789240694811\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/186463/9789240694811_eng.pdf) (accessed 22 Jul 2022).
2. O'Caioimh R, Galluzzo L, Rodríguez-Laso Á, et al. Prevalence of frailty at population level in European ADVANTAGE Joint Action Member States: a systematic review and meta-analysis. *Ann Ist Super Sanita* 2018; **54**(3): 226–238. DOI: [https://doi.org/10.4415/ANN\\_18\\_03\\_10](https://doi.org/10.4415/ANN_18_03_10)
3. Rodríguez-Laso A, Caballero Mora MA, García Sánchez I, et al. Updated state of the art report on the prevention and management of frailty. 2019. [https://www.advantageja.eu/images/SoAR-AdvantageJA\\_Fulltext.pdf](https://www.advantageja.eu/images/SoAR-AdvantageJA_Fulltext.pdf) (accessed 12 May 2022).
4. Rodríguez-Laso Á, O'Caioimh R, Galluzzo L, et al. Population screening, monitoring and surveillance for frailty: three systematic reviews and a grey literature review. *Ann Ist Super Sanita* 2018; **54**(3): 253–262. DOI: [https://doi.org/10.4415/ANN\\_18\\_03\\_13](https://doi.org/10.4415/ANN_18_03_13)
5. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. *J Nutr Health Aging* 2012; **16**(7): 601–608. DOI: <https://doi.org/10.1007/s12603-012-0084-2>
6. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol* 1994; **49**(2): M85–M94. DOI: <https://doi.org/10.1093/geronj/49.2.m85>
7. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 2016; **6**(11): e012799. DOI: <https://doi.org/10.1136/bmjopen-2016-012799>
8. Checa-López M, Oviedo-Briones M, Pardo-Gómez A, 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**(1): 86. DOI: <https://doi.org/10.1186/s12877-019-1042-1>
9. 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**(5): 373–383. DOI: [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)

10. 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**(3): M146–M156. DOI: <https://doi.org/10.1093/gerona/56.3.m146>
11. Orme JG, Reis J, Herz EJ. Factorial and discriminant validity of the Center for Epidemiological Studies Depression (CES-D) scale. *J Clin Psychol* 1986; **42**(1): 28–33. DOI: [https://doi.org/10.1002/1097-4679\(198601\)42:1<28::aid-jclp2270420104>3.0.co;2-t](https://doi.org/10.1002/1097-4679(198601)42:1<28::aid-jclp2270420104>3.0.co;2-t)
12. Hoogendijk EO, van Kan GA, Guyonnet S, et al. Components of the frailty phenotype in relation to the frailty index: results from the Toulouse Frailty Platform. *J Am Med Dir Assoc* 2015; **16**(10): 855–859. DOI: <https://doi.org/10.1016/j.jamda.2015.04.007>
13. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. *ScientificWorldJournal* 2001; **1**: 323–336. DOI: <https://doi.org/10.1100/tsw.2001.58>
14. Searle SD, Mitnitski A, Gahbauer EA, et al. A standard procedure for creating A frailty index. *BMC Geriatr* 2008; **8**: 24. DOI: <https://doi.org/10.1186/1471-2318-8-24>
15. 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**(2): e1–e8. DOI: <https://doi.org/10.1016/j.archger.2012.03.001>
16. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969; **9**(3): 179–186. DOI: [https://doi.org/10.1093/geront/9.3\\_Part\\_1.179](https://doi.org/10.1093/geront/9.3_Part_1.179)
17. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. *Md State Med J* 1965; **14**: 61–65.
18. Ambagtsheer RC, Visvanathan R, Dent E, et al. Commonly used screening instruments to identify frailty among community-dwelling older people in a general practice (primary care) setting: a study of diagnostic test accuracy. *J Gerontol A Biol Sci Med Sci* 2020; **75**(6): 1134–1142. DOI: <https://doi.org/10.1093/gerona/glz260>
19. Dong L, Qiao X, Tian X, et al. Cross-cultural adaptation and validation of the FRAIL scale in chinese community-dwelling older adults. *J Am Med Dir Assoc* 2018; **19**(1): 12–17. DOI: <https://doi.org/10.1016/j.jamda.2017.06.011>
20. Thompson MQ, Theou O, Tucker GR, et al. FRAIL scale: predictive validity and diagnostic test accuracy. *Australas J Ageing* 2020; **39**(4): e529–e536. DOI: <https://doi.org/10.1111/ajag.12829>
21. Si H, Jin Y, Qiao X, et al. Predictive performance of 7 frailty instruments for short-term disability, falls and hospitalization among Chinese community-dwelling older adults: a prospective cohort study. *Int J Nurs Stud* 2021; **117**: 103875. DOI: <https://doi.org/10.1016/j.ijnurstu.2021.103875>
22. Bandinelli S, Lauretani F, Boscherini V, et al. A randomized, controlled trial of disability prevention in frail older patients screened in primary care: the FRASI study. design and baseline evaluation. *Ageing Clin Exp Res* 2006; **18**(5): 359–366. DOI: <https://doi.org/10.1007/BF03324831>
23. Oviedo-Briones M, Laso ÁR, Carnicero JA, Cesari M, 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**(3): 607.e7–607.e12. DOI: <https://doi.org/10.1016/j.jamda.2020.09.024>
24. Ministry of Health of Spain. Consensus document on frailty and falls prevention among the elderly. The prevention and health promotion strategy of the Spanish NHS Madrid. 2014. [https://www.sanidad.gob.es/profesionales/saludPublica/prevPromocion/Estrategia/EnvejecimientoSaludable\\_Fragilidad/docs/Frailtyandfalls\\_Elderly.pdf](https://www.sanidad.gob.es/profesionales/saludPublica/prevPromocion/Estrategia/EnvejecimientoSaludable_Fragilidad/docs/Frailtyandfalls_Elderly.pdf) (accessed 29 Jul 2022).
25. Moody D, Lyndon H, Stevens G. Toolkit for general practice in supporting older people living with frailty. 2017. <https://www.england.nhs.uk/publication/toolkit-for-general-practice-in-supporting-older-people-living-with-frailty> (accessed 22 Jul 2022).
26. NHS England. Identifying frailty. <https://www.england.nhs.uk/ourwork/clinical-policy/older-people/frailty/frailty-risk-identification> (accessed 22 Jul 2022).
27. 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**(1): 10–12. DOI: <https://doi.org/10.1093/ageing/aft160>