Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Can a single question replace the Cumberland Ankle Instability Tool? Diagnostic accuracy of the Simple Foot and Ankle Value for screening functional ankle instability: a multicentre cross-sectional study

Marine Deba
Centre Hospitalier Universitaire de Toulouse
Ronny Lopes
Centre Orthopédique Santy
Alexandre Hardy
Clinique du Sport
David Ancelin
d.ancelin@yahoo.fr

Centre Hospitalier Universitaire de Toulouse

Research Article

Keywords: Ankle instability, Cumberland Ankle Instability Tool, Single-item measure, Patient-reported outcome measures, Screening, Diagnostic accuracy, Foot and Ankle Ability Measure, Simple Foot and Ankle Value

Posted Date: August 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10485067/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background.

Functional ankle instability (FAI) is common, largely undetected in the community, and amenable to low-cost neuromuscular intervention, yet the instruments used to identify it are too long for routine consultations. The Cumberland Ankle Instability Tool (CAIT) requires nine items per ankle and more than five minutes to complete bilaterally. The Simple Foot and Ankle Value (SFAV) is a single question that has been validated as a measure of ankle function but has never been evaluated as a screening test. We examined whether the SFAV can substitute for the CAIT as a first-line screen for FAI.

Methods.

In a multicentre cross-sectional study conducted between January and March 2025 in three French orthopaedic centres, 717 adults accompanying patients to consultation completed a self-administered questionnaire comprising sociodemographic data, the Ankle Activity Score, the Foot and Ankle Ability Measure (FAAM), the CAIT and the SFAV. FAI was defined a priori as a CAIT score of 23 or less in at least one ankle, the threshold established in the French validation of the instrument. Diagnostic accuracy of the SFAV was assessed by receiver operating characteristic (ROC) analysis, with the FAAM analysed as a comparator index test. Confidence intervals were obtained by bootstrap resampling and Wilson's method, and the derived decision rule was internally validated by bootstrapping. Pre-specified sensitivity analyses used alternative CAIT thresholds of 24, 25 and 27.

Results.

FAI was present in 196 of 716 participants (27.4%), bilaterally in 111 (56.6% of cases). In the 704 participants with complete index and reference data, the SFAV discriminated FAI with an area under the curve (AUC) of 0.898 (95% CI 0.869–0.923). The Youden-optimal threshold was 87; the pragmatic threshold of SFAV below 90 performed almost identically (Youden index 0.728 versus 0.729) and yielded a sensitivity of 91.3% (95% CI 86.5–94.5), specificity of 81.5% (77.9–84.6), positive predictive value of 65.6% (59.8–71.0), negative predictive value of 96.1% (93.8–97.5), and likelihood ratios of 4.94 and 0.11. Applying this rule would have avoided 61.2% of CAIT administrations. The SFAV outperformed both FAAM subscales as a screening test (AUC 0.898 versus 0.772 for FAAM–ADL, difference 0.126, 95% CI 0.082–0.169; and versus 0.755 for FAAM–Sport, difference 0.143, 95% CI 0.099–0.185). Accuracy was stable across CAIT thresholds, age, sex and activity level, and bootstrap internal validation showed negligible optimism.

Conclusions.

A single question performed better than a 21-item region-specific questionnaire at identifying CAIT-defined functional ankle instability. An SFAV below 90 is an efficient first-line triage rule that reserves the CAIT for the minority of individuals who screen positive. These findings concern functional instability as operationalised by the CAIT and should not be extrapolated to chronic ankle instability as defined by the International Ankle Consortium.

Background

Lateral ankle sprain is among the most frequent musculoskeletal injuries worldwide and imposes a burden that extends well beyond the acute episode [1, 2]. A substantial proportion of individuals who sustain a sprain never fully recover: prospective cohort data indicate that a third to a half continue to report symptoms, recurrent giving way or reduced function a year later, and the 2016 consensus statement of the International Ankle Consortium identified persistent instability, reduced physical activity and post-traumatic ankle osteoarthritis as the principal long-term consequences [2, 3, 4]. Because a large share of sprains are managed without medical contact, much of this burden remains invisible to the health system, and the individuals concerned are rarely offered the balance and neuromuscular training programmes that reduce recurrence at low cost and negligible risk [2, 5].

Two related but distinct constructs are used to describe the residual problem, and the distinction matters for the present study. Chronic ankle instability (CAI), as operationalised by the International Ankle Consortium, is a composite condition requiring a documented history of a significant ankle sprain, at least one subsequent episode of giving way, and impaired self-reported function on validated instruments, with a minimum interval since injury [6]. Functional ankle instability (FAI), by contrast, denotes the self-perceived component of that picture — the subjective sense that the ankle is unreliable during daily and sporting activities — irrespective of documented injury history. This distinction is not merely semantic. Perceived instability is reported by people who cannot recall a significant sprain, and Hiller and colleagues have argued that perceived instability, mechanical laxity and recurrent sprain overlap only partially and should be modelled as separate, intersecting entities [7, 8]. In a population-based sample of 900 healthy individuals aged 8 to 101 years, Baldwin and colleagues found perceived ankle instability in 31%, most often bilaterally, a pattern difficult to attribute solely to residual effects of unilateral trauma [9]. FAI is therefore a legitimate target in its own right, and one that can be identified without any injury history being available.

The clinical argument for detecting FAI in unselected adults rests on three observations. First, it is common: pooled estimates place the prevalence of perceived instability at approximately one in four in active populations, with wide variation according to criteria and setting [10]. Second, it is consequential: individuals with self-reported instability show altered postural control and gait organisation, lower objectively measured physical activity, and impaired ankle-related quality of life, and repeated instability episodes are implicated in the development of ligamentous post-traumatic ankle osteoarthritis, which accounts for the majority of ankle arthritis [4, 11, 12]. Third, and decisively for a screening argument, it is actionable: supervised balance and neuromuscular training reduces the risk of recurrent sprain and

improves patient-reported function, requires no equipment, and carries essentially no harm, so the cost of a false-positive screen is a short conversation and an exercise leaflet rather than an invasive investigation [5, 13].

What is missing is not a measurement instrument but a triage instrument. The Cumberland Ankle Instability Tool (CAIT) is the most widely used measure of FAI and has excellent psychometric properties, but it comprises nine items per ankle and must be completed bilaterally, which takes more than five minutes [14]. Median consultation duration is under fifteen minutes in most developed health systems, and often under ten in primary care [15], so an instrument of this length competes directly with the clinical encounter itself and is unlikely to be adopted for opportunistic screening. Single-item patient-reported measures have been proposed as a solution to precisely this problem in other joints. The Subjective Shoulder Value and the Single Assessment Numeric Evaluation correlate strongly with multi-item instruments while taking seconds to administer [16, 17], and the same logic has recently been applied to the ankle. The Simple Foot and Ankle Value (SFAV) asks a single question — how the respondent rates the function of their ankle as a percentage of normal — and has been shown to correlate strongly with the FAOS and AOFAS scores in surgical patients and to discriminate symptomatic from asymptomatic individuals in a mixed clinical population [18, 19].

Both existing SFAV studies evaluated it as a measure of ankle function in symptomatic or postoperative patients. Neither examined whether it can be used as a screening test, and neither addressed instability specifically. Whether a single global question about function can identify people whose ankles feel unreliable is not self-evident: perceived function and perceived stability are related but not identical, and a person may consider their everyday function entirely normal while still experiencing episodic giving way. The present study was therefore designed to answer a deliberately narrow and directly testable question: can the SFAV substitute for the CAIT as a first-line screen for functional ankle instability in unselected adults? Our primary objective was to quantify the criterion validity of the SFAV against CAIT-defined FAI and to derive a usable threshold. Secondary objectives were to compare the SFAV with the Foot and Ankle Ability Measure (FAAM) analysed as a comparator index test — that is, to ask whether a longer region-specific questionnaire does the screening job any better — and to characterise the correlation between the three instruments. We hypothesised that the SFAV would discriminate CAIT-defined instability with clinically useful accuracy at a threshold suitable for routine use.

We emphasise at the outset that the CAIT is used here as the reference standard by design and not as a proxy for the International Ankle Consortium definition of CAI. The criteria required for that definition, in particular a documented history of ankle sprain and of subsequent giving way episodes, were not collected in this study. All findings reported below therefore concern functional ankle instability as operationalised by the CAIT, and the term CAI is not used to describe our results.

Methods

Design, setting and reporting

This was a multicentre, cross-sectional diagnostic accuracy study conducted between 1 January and 31 March 2025 in the orthopaedic outpatient departments of three French centres: Pierre-Paul Riquet Hospital (Toulouse University Hospital), Clinique du Sport (Paris) and Centre Orthopédique Santy, Hôpital Privé Jean Mermoz (Lyon). The study is reported in accordance with the STARD 2015 statement for diagnostic accuracy studies [20]; the completed checklist is provided as Additional file 1.

The study involved no intervention on participants and consisted solely of a self-administered questionnaire. Under French law (articles L.1121-1 and R.1121-1 of the French Public Health Code, as amended by Loi n° 2012 – 300 of 5 March 2012), research of this type is classified as research not involving the human person (*recherche n'impliquant pas la personne humaine*). Such research falls outside the remit of the Comités de Protection des Personnes, the committees designated by French law to review research involving the human person, and requires no ethics committee opinion; the need for ethics approval was therefore waived in accordance with this national legislation, and the study was not submitted to an ethics committee or institutional review board. The collection and computerised processing of personal and medical data were carried out under reference methodology MR-004 of the French National Commission for Informatics and Liberties (CNIL), the framework governing this category of research in France. Toulouse University Hospital, the sponsoring institution, signed a commitment of compliance with this methodology (CNIL number 2407763 v 0). The study was evaluated and validated by the data protection officer of Toulouse University Hospital in accordance with Regulation (EU) 2016/679, and is registered in the register of studies not involving the human person of Toulouse University Hospital under reference number RnIPH 2024-59. All participants received written information about the study, including its purpose, the voluntary nature of participation and their right to decline, and consent was implied by their completion and return of the questionnaire.

Participants

Individuals accompanying a patient to an orthopaedic consultation were invited consecutively to take part during the recruitment period. This sampling frame was chosen deliberately: companions attend for reasons unrelated to their own musculoskeletal health, which avoids the spectrum bias inherent in recruiting from patients presenting with ankle complaints, while remaining practical to implement. Adults aged 18 years or older who provided informed consent were eligible. Participants were excluded if they were under 18 years of age, reported any history of lower-limb surgery, were under active treatment or follow-up for an ankle condition at the time of inclusion, or were unable to complete the questionnaire unaided.

Index test

The index test was the SFAV, a single self-administered question phrased, in the validated French wording, as: "How would you rate the function of your ankle as a percentage of normal (0% to 100%, where 100% is normal)?" [19]. Participants recorded a single value between 0 and 100. The question

refers to the respondent's ankle function as a whole rather than to a designated side, and was therefore treated throughout as a person-level measure. It was positioned at the end of the questionnaire, after the CAIT and the FAAM; the implications of this ordering are considered in the Discussion.

Comparator index test

The FAAM was administered as a comparator index test in both its Activities of Daily Living (FAAM–ADL, 21 items) and Sport (FAAM–Sport, 8 items) subscales, using the validated French version and scored as a percentage of the maximum attainable score [21, 22]. The FAAM was included with a specific purpose: it is the most widely used region-specific patient-reported instrument for foot and ankle conditions and is recommended within the International Ankle Consortium criteria, so it represents the realistic alternative to the SFAV for a clinician who wants a shorter instrument than the CAIT. Comparing the two directly answers whether the brevity of a single item is bought at the cost of accuracy.

Reference standard

The reference standard was the French version of the CAIT, completed for both ankles [14, 23]. FAI was defined a priori as a CAIT score of 23 or less in at least one ankle. This threshold was taken from the French translation and validation study of Geerinck and colleagues, who derived it by Youden index in 102 French-speaking participants and reported a sensitivity of 92.2% and a specificity of 100% [23]. We chose it because all participants completed the French instrument and because a cut-off derived in the same language and cultural context is the most defensible choice; it should not be assumed to transfer to other CAIT versions. The threshold is not universal: the original English instrument used 27 or below [14], Wright and colleagues recalibrated the English cut-off to 25 [24], and translations have produced values ranging from 11 to 25 [23]. Because the choice of threshold materially affects apparent prevalence and therefore predictive values, we pre-specified sensitivity analyses repeating the entire primary analysis with the reference standard set at 24, 25 and 27.

Both ankles were assessed because instability is frequently bilateral, and a person-level definition based on the more affected ankle is the definition a clinician would use when deciding whether to investigate further. For correlation analyses the lower of the two CAIT scores was therefore used as the person-level value, with side-specific correlations reported alongside.

Covariates

Age, sex, body mass index, smoking status and the Ankle Activity Score (AAS) [25] were recorded. Smoking was collected because it is an established correlate of impaired connective-tissue healing and of poorer patient-reported outcomes after musculoskeletal injury and was therefore considered a plausible confounder of self-reported function at the design stage. We acknowledge, and return to in the Discussion, that history of previous ankle sprain, of giving way episodes and of past sport participation were not recorded; their absence is the principal limitation of the dataset.

Data collection

The questionnaire administered was a compilation of previously published and validated instruments, and no new instrument was developed for this study. It comprised a sociodemographic section written for the present study (age, sex, height, weight and smoking status), the Ankle Activity Score [25], the French version of the Foot and Ankle Ability Measure in its Activities of Daily Living and Sport subscales [21, 22], the French version of the Cumberland Ankle Instability Tool, completed for each ankle [14, 23], and the Simple Foot and Ankle Value [19]. An English-language version of the questionnaire as administered is provided as Additional file 2.

Participants chose between a paper and an online version of the questionnaire according to preference. The online version required each field to be completed before submission; the paper version did not. The two versions were otherwise identical in wording and item order.

Statistical analysis

The primary analysis quantified the ability of the SFAV to discriminate participants meeting the FAI definition. Discrimination was summarised by the area under the receiver operating characteristic curve (AUC), with 95% confidence intervals obtained by percentile bootstrap using 5000 resamples. The optimal threshold was identified by the Youden index (J = sensitivity + specificity − 1); because the Youden-optimal value fell on a non-round number, we additionally evaluated the nearest clinically convenient threshold and report both. At each candidate threshold we calculated sensitivity, specificity, positive and negative predictive values with Wilson score confidence intervals, positive and negative likelihood ratios, and the proportion of participants screening positive. Predictive values are conditional on the prevalence observed in this sample and are interpreted accordingly.

The decision rule was internally validated by bootstrap resampling with 2000 replications, re-deriving the Youden-optimal threshold in each resample to quantify threshold stability and estimating optimism in sensitivity, specificity and predictive values. AUCs of the SFAV and of the two FAAM subscales were compared on the subset with complete data for all three instruments, using the bootstrap distribution of paired AUC differences to obtain confidence intervals and two-sided p values.

Distributional normality was assessed by the Shapiro–Wilk test, which rejected normality for all scores; correlations were therefore estimated by Spearman's rank coefficient with confidence intervals derived by Fisher's z transformation. Correlation magnitude was interpreted according to the conventions summarised by Schober and colleagues, in which coefficients of 0.40 to 0.69 are considered moderate, 0.70 to 0.89 strong and 0.90 or above very strong [26]. Continuous variables were compared between groups with the Mann–Whitney U test and categorical variables with the chi-squared test.

Missing index-test data were handled by complete-case analysis for the primary estimate. Because the pattern of missingness proved to be informative (see Results), we additionally computed best-case and worst-case bounds on specificity by assigning all participants with a missing SFAV in turn to the screen-

negative and screen-positive categories. Pre-specified subgroup analyses examined accuracy by age tertile, sex and activity level. The alpha level was set at 0.05 and all tests were two-sided. Analyses were performed in Python 3.11 using scikit-learn 1.8, SciPy 1.17 and pandas.

No formal sample size calculation was performed, the sample being determined by the recruitment window. Post hoc, the 196 observed cases give a half-width of approximately 4 percentage points on the 95% confidence interval for sensitivity, and the 508 non-cases a half-width of approximately 3.4 points for specificity, which we considered adequate precision for a screening application.

Results

Participants

Of 762 individuals assessed for eligibility, 45 were excluded (4 aged under 18, 32 with a history of lower-limb surgery and 9 under active follow-up for an ankle condition), leaving 717 participants (Fig. 1). Three hundred and thirty-two (46.3%) completed the online version and 385 (53.7%) the paper version. Complete data were available for the CAIT in 716 participants (99.9%), for the FAAM–ADL in 714 (99.6%), for the FAAM–Sport in 711 (99.2%) and for the SFAV in 705 (98.3%). The primary analysis set, comprising participants with both index and reference standard data, contained 704 individuals.

Participants had a mean age of 40.6 years (SD 15.0, range 18 to 85) and a mean body mass index of 24.6 kg/m² (SD 4.3); 403 (56.2%) were women and 124 (17.3%) were current smokers. The median Ankle Activity Score was 4 (interquartile range 4 to 7), indicating a predominantly moderate level of physical activity. Characteristics of the whole sample and of the two reference-standard groups are given in Table 1. Participants meeting the FAI definition had a slightly higher body mass index (median 25 versus 23 kg/m², $p < 0.001$) and were more often women (64.3% versus 53.1%, $p = 0.009$), but did not differ in age, activity level or smoking status.

Table 1
Characteristics of participants, overall and by reference-standard group

Characteristic	All participants (n = 717)	Functional ankle instability (n = 196)	No functional ankle instability (n = 520)	P value
Age, years, mean (SD)	40.6 (15.0)	39.5 (14.7)	41.0 (15.1)	0.212
Age, years, median (IQR)	37 (28–53)	37 (27–52)	38 (28–53)	0.212
Body mass index, kg/m², mean (SD)	24.6 (4.3)	25.6 (5.0)	24.2 (3.9)	< 0.001
Ankle Activity Score, median (IQR)	4 (4–7)	5 (4–7)	4 (4–6)	0.757
Women, n (%)	403 (56.2)	126 (64.3)	276 (53.1)	0.009
Current smokers, n (%)	124 (17.3)	34 (17.3)	90 (17.3)	1.000

Functional ankle instability was defined as a Cumberland Ankle Instability Tool score of 23 or less in at least one ankle. One participant with missing Cumberland Ankle Instability Tool data could not be classified and is included only in the first column. Continuous variables were compared with the Mann–Whitney U test and categorical variables with the chi-squared test. IQR, interquartile range; SD, standard deviation.

Prevalence and laterality of functional ankle instability

A CAIT score of 23 or less in at least one ankle was present in 196 of 716 participants (27.4%). Instability was bilateral in 111 participants (15.5% of the cohort, 56.6% of cases) and unilateral in 85 (11.9% of the cohort). The prevalence in the 704 participants forming the primary analysis set was 27.8%. Functional scores in the two groups are reported in Table 2: participants with FAI scored lower on every instrument, with the largest separation on the SFAV (median 70, IQR 52 to 80, versus median 98, IQR 90 to 100; $p < 0.001$).

Table 2
Functional scores, overall and by reference-standard group

Instrument	All participants	Functional ankle instability	No functional ankle instability	P value
FAAM–ADL (%)	96.6 (8.1), n = 714	92.7 (9.7), n = 196	98.0 (6.9), n = 517	< 0.001
FAAM–Sport (%)	91.3 (17.2), n = 711	83.2 (19.6), n = 195	94.3 (15.1), n = 515	< 0.001
CAIT, left ankle (/30)	26.3 (4.8), n = 716	20.7 (5.6), n = 196	28.4 (1.7), n = 520	< 0.001
CAIT, right ankle (/30)	26.2 (4.7), n = 716	20.5 (5.4), n = 196	28.3 (1.8), n = 520	< 0.001
CAIT, more affected ankle (/30)	25.4 (5.3), n = 716	18.2 (4.4), n = 196	28.2 (1.9), n = 520	< 0.001
SFAV (%)	84.6 (19.9), n = 705	66.4 (18.0), n = 196	91.6 (15.8), n = 508	< 0.001

Values are mean (standard deviation). Groups were compared with the Mann–Whitney U test. CAIT, Cumberland Ankle Instability Tool; FAAM–ADL, Foot and Ankle Ability Measure, Activities of Daily Living subscale; FAAM–Sport, Foot and Ankle Ability Measure, Sport subscale; SFAV, Simple Foot and Ankle Value.

Diagnostic accuracy of the SFAV

The SFAV discriminated CAIT-defined FAI with an AUC of 0.898 (95% CI 0.869 to 0.923) (Fig. 2a). The Youden index was maximised at a threshold of 87 (sensitivity 90.8%, specificity 82.1%, J = 0.729). Because SFAV responses cluster heavily on round values — 74.8% were multiples of ten and 87.7% multiples of five — we also evaluated the adjacent round threshold of 90. A cut-off of SFAV below 90 gave an essentially identical Youden index of 0.728 and is far easier to apply at the point of care; it is therefore the threshold we recommend and report in full.

Performance across candidate thresholds is shown in Table 3. At SFAV below 90, sensitivity was 91.3% (95% CI 86.5 to 94.5) and specificity 81.5% (77.9 to 84.6), with a positive predictive value of 65.6% (59.8 to 71.0) and a negative predictive value of 96.1% (93.8 to 97.5) at the observed prevalence of 27.8%. The positive likelihood ratio was 4.94 and the negative likelihood ratio 0.11. Of 704 participants, 273 (38.8%) screened positive, so applying the rule would have avoided administering the CAIT to 61.2% of the sample while missing 17 of 196 cases (8.7%).

Table 3
Diagnostic performance of the Simple Foot and Ankle Value across candidate thresholds

SFAV threshold	Screen-positive n (%)	False neg.	Sensitivity % (95% CI)	Specificity % (95% CI)	PPV % (95% CI)	NPV % (95% CI)	LR+ / LR−	Youden
< 70	107 (15.2)	112	42.9 (36.1–49.9)	95.5 (93.3–97.0)	78.5 (69.8–85.2)	81.2 (77.9–84.2)	9.47 / 0.60	0.383
< 75	156 (22.2)	72	63.3 (56.3–69.7)	93.7 (91.2–95.5)	79.5 (72.5–85.1)	86.9 (83.8–89.4)	10.04 / 0.39	0.570
< 80	174 (24.7)	61	68.9 (62.1–74.9)	92.3 (89.7–94.3)	77.6 (70.8–83.1)	88.5 (85.5–90.9)	8.97 / 0.34	0.612
< 85	247 (35.1)	28	85.7 (80.1–89.9)	84.4 (81.0–87.3)	68.0 (62.0–73.5)	93.9 (91.3–95.7)	5.51 / 0.17	0.702
≤ 87	269 (38.2)	18	90.8 (86.0–94.1)	82.1 (78.5–85.2)	66.2 (60.3–71.6)	95.9 (93.6–97.4)	5.07 / 0.11	0.729
< 90	273 (38.8)	17	91.3 (86.5–94.5)	81.5 (77.9–84.6)	65.6 (59.8–71.0)	96.1 (93.8–97.5)	4.94 / 0.11	0.728
≤ 90	370 (52.6)	10	94.9 (90.9–97.2)	63.8 (59.5–67.8)	50.3 (45.2–55.3)	97.0 (94.6–98.4)	2.62 / 0.08	0.587
< 95	377 (53.6)	10	94.9 (90.9–97.2)	62.4 (58.1–66.5)	49.3 (44.3–54.4)	96.9 (94.5–98.3)	2.52 / 0.08	0.573

Analysis set n = 704, of whom 196 (27.8%) met the reference-standard definition of functional ankle instability (Cumberland Ankle Instability Tool score ≤ 23 in at least one ankle). Confidence intervals for proportions were calculated by Wilson's method. The threshold of SFAV ≤ 87 is the Youden-optimal value; SFAV < 90 is the recommended pragmatic threshold. CI, confidence interval; LR+, positive likelihood ratio; LR−, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; SFAV, Simple Foot and Ankle Value.

Table 3 also makes explicit a distinction that matters in practice. Defining screen-positive as SFAV of 90 or below rather than strictly below 90 raises sensitivity to 94.9% and reduces false negatives to ten, but specificity falls to 63.8% and positive predictive value to 50.3%, because a large number of participants scored exactly 90. The two rules are not interchangeable, and the strict inequality is the better-balanced choice.

The 17 participants misclassified as screen-negative had less severe instability than cases overall (mean CAIT score in the more affected ankle 20.4 versus 18.2). Seven scored exactly 90 and eight rated their ankle function at 100% despite CAIT scores as low as 10, illustrating that a minority of individuals regard their everyday function as entirely normal while still reporting the situational apprehension that the CAIT captures.

Comparison with the FAAM

In the 700 participants with complete data for all three instruments, the SFAV discriminated FAI substantially better than either FAAM subscale (Fig. 2b). The AUC was 0.898 (95% CI 0.869 to 0.923) for the SFAV, 0.772 (0.734 to 0.809) for the FAAM–ADL and 0.755 (0.716 to 0.794) for the FAAM–Sport. The paired differences were 0.126 (95% CI 0.082 to 0.169; $p < 0.001$) in favour of the SFAV over the FAAM–ADL and 0.143 (0.099 to 0.185; $p < 0.001$) over the FAAM–Sport. Ceiling effects differed markedly between instruments: 55.7% of participants achieved the maximum FAAM–ADL score and 56.5% the maximum FAAM–Sport score, compared with 33.9% for the SFAV.

Internal validation and robustness

Bootstrap internal validation of the SFAV below 90 rule showed negligible optimism: bootstrap-corrected estimates were 91.4% for sensitivity, 81.5% for specificity, 65.5% for positive predictive value and 96.1% for negative predictive value, differing from the apparent values by less than 0.1 percentage points in every case. Re-deriving the Youden-optimal threshold in each of 2000 resamples produced a median value of 87 with an interquartile range of 85 to 89 and a 95% interval of 82 to 89, confirming that the threshold is stable and that 90 lies at the upper edge of the plausible range.

Accuracy was consistent across pre-specified subgroups. The AUC was 0.884 in participants aged 18 to 34 years, 0.888 in those aged 35 to 49 and 0.931 in those aged 50 or above; 0.904 in women and 0.888 in men; and 0.894 and 0.901 in participants with lower and higher Ankle Activity Scores respectively. Sensitivity at the recommended threshold ranged from 89.0% to 96.5% and specificity from 79.0% to 83.6% across these strata.

The sensitivity analyses using alternative CAIT thresholds are summarised in Table 4. Discrimination remained good as the reference standard was relaxed, with AUCs of 0.898, 0.874, 0.850 and 0.821 at CAIT cut-offs of 23, 24, 25 and 27 respectively, and the Youden-optimal SFAV threshold remained between 87 and 93. As expected, sensitivity at SFAV below 90 fell and specificity rose as the reference definition became more inclusive, since the additional cases captured by a higher CAIT threshold have milder instability and correspondingly higher self-rated function.

Table 4
Sensitivity analyses using alternative Cumberland Ankle Instability Tool thresholds

Reference standard	Cases n (%)	AUC (95% CI)	Youden SFAV cut-off	Sensitivity %	Specificity %	PPV %	NPV %
CAIT ≤ 23	196 (27.8)	0.898 (0.869–0.924)	≤ 87	91.3	81.5	65.6	96.1
CAIT ≤ 24	223 (31.7)	0.874 (0.845–0.903)	≤ 89	85.2	82.7	69.6	92.3
CAIT ≤ 25	261 (37.1)	0.850 (0.820–0.878)	≤ 89	78.2	84.4	74.7	86.8
CAIT ≤ 27	369 (52.4)	0.821 (0.790–0.850)	≤ 93	63.1	88.1	85.3	68.4

Sensitivity, specificity and predictive values are given at the recommended screening threshold of SFAV < 90. All analyses were performed in the primary analysis set (n = 704). The threshold of 23 was pre-specified as the primary reference standard on the basis of the French validation of the Cumberland Ankle Instability Tool; 25 corresponds to the recalibrated English threshold and 27 to the original English instrument. Confidence intervals for the area under the curve were obtained by percentile bootstrap with 2000 resamples. AUC, area under the receiver operating characteristic curve; CAIT, Cumberland Ankle Instability Tool; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; SFAV, Simple Foot and Ankle Value.

Missing data

The SFAV was missing for 12 participants (1.7%). None of them met the FAI definition, the lowest CAIT score in this group being 25, and their mean CAIT score in the more affected ankle was higher than that of participants who answered (28.6 versus 25.4, p = 0.02); they did not differ in age, body mass index or activity level. Sensitivity is therefore unaffected by these exclusions, since none of the participants with missing index data could have contributed a false negative. Bounding the effect on specificity by assigning all 12 in turn to the screen-negative and screen-positive categories gives a range of 79.6% to 81.9% around the complete-case estimate of 81.5%, so no plausible pattern of missingness would materially alter the conclusions.

Correlations between instruments

The SFAV correlated moderately with the CAIT score of the more affected ankle, at the upper end of that range (ρ = 0.66, 95% CI 0.61 to 0.70; p < 0.001; Fig. 3a), and somewhat less closely with each side considered separately (ρ = 0.57 for the left and 0.55 for the right ankle). Correlations with the FAAM were lower: ρ = 0.50 (95% CI 0.43 to 0.55) with the FAAM–ADL and ρ = 0.50 (0.44 to 0.56) with the FAAM–

Sport. The CAIT of the more affected ankle correlated with the FAAM–ADL at ρ = 0.44. The distribution of SFAV scores by reference-standard group is shown in Fig. 3b.

Discussion

The principal finding of this study is that a single question performed better than a 21-item region-specific questionnaire at identifying functional ankle instability as defined by the CAIT. In 704 adults recruited outside any ankle-related care pathway, the SFAV discriminated CAIT-defined instability with an area under the curve of 0.898, and a threshold of below 90 combined a sensitivity of 91.3% with a specificity of 81.5% and a negative predictive value of 96.1%. Applied prospectively, this rule would have restricted administration of the CAIT to 39% of the sample while identifying more than nine in ten cases. The same reference standard was discriminated far less well by the FAAM–ADL and FAAM–Sport, whose areas under the curve were 0.13 and 0.14 lower.

That a one-item measure should outperform a longer one is counter-intuitive only if length is equated with information. Two explanations are consistent with our data. The first is a matter of construct alignment. The FAAM enumerates specific tasks — walking on uneven ground, descending stairs, running — and asks how difficult each is; a person whose ankle gives way unpredictably during otherwise easy activities can answer that none of these tasks is difficult and still feel that the joint is unreliable. The SFAV instead asks for a single global judgement of function relative to normal, and that judgement appears to incorporate the sense of unreliability that the CAIT is designed to detect. The second explanation is distributional. More than half of our participants achieved the maximum FAAM–ADL and FAAM–Sport scores, against a third for the SFAV, so the FAAM had far less room to separate mild instability from none in a non-clinical sample. Ceiling effects of this magnitude in unselected populations are a known limitation of region-specific instruments developed in patient cohorts, and they bear directly on the suitability of such instruments for screening. It is worth noting that this comparison concerns screening performance and not measurement quality: the FAAM remains a valid and responsive outcome measure in symptomatic populations, and nothing here argues against its use for that purpose.

Our results extend rather than contradict the two existing validations of the SFAV. Barnavon and colleagues reported strong correlations between a single ankle value and the FAOS and AOFAS scores in patients undergoing ankle and hindfoot surgery, and demonstrated responsiveness to postoperative change [18]; Marot and colleagues showed that the SFAV distinguished patients with acute or arthritic ankle conditions from healthy controls [19]. Both studies established the SFAV as a measure of ankle function in clinical populations. Neither addressed instability, and neither evaluated the instrument as a test with a threshold, sensitivity and predictive values. The present study answers that different question, and does so in a population in which screening would actually be deployed. The magnitude of association we observed between the SFAV and the CAIT of the more affected ankle (ρ = 0.66) is of the same order as that reported between single-item measures and multi-item instruments at other joints, such as the Subjective Shoulder Value against the Constant score and the Single Assessment Numeric Evaluation against the revised Foot Function Index [16, 17], suggesting that the utility of single-item

global ratings is a general phenomenon rather than a peculiarity of the ankle. It is also notable that the SFAV correlated more closely with the more affected ankle than with either side taken individually, which supports treating the question as a person-level rather than a side-specific measure.

The prevalence of CAIT-defined instability we observed, 27.4%, is consistent with the pooled figure of approximately 25% reported by Doherty and colleagues for active populations [10] and with the 31% of perceived instability found by Baldwin and colleagues in 900 healthy individuals [9]. The distribution of laterality is also concordant: bilateral involvement accounted for 56.6% of our cases, against roughly three-quarters in the Baldwin cohort. A preponderance of bilateral involvement is difficult to reconcile with a model in which perceived instability is simply the residue of a previous unilateral injury, and is more compatible with the contribution of person-level factors — proprioceptive acuity, postural control strategy, generalised laxity — to which both ankles are equally exposed. This interpretation aligns with the updated model of chronic ankle instability proposed by Hertel and Corbett, in which pathomechanical, sensory-perceptual and motor-behavioural impairments interact rather than following a single injury-driven pathway [8]. It also supports the case for screening in individuals with no injury history at all, since under such a model the absence of a recalled sprain does not exclude clinically meaningful instability.

The clinical value of a screening test depends on what happens after a positive result, and here the arithmetic is favourable. At the prevalence observed in our sample, a positive screen raised the probability of CAIT-defined instability from 28% to approximately 66%, and a negative screen lowered it to 4%. In a lower-prevalence setting — say a general practice population with 10% prevalence — the same likelihood ratios of 4.94 and 0.11 would move a positive result to approximately 35% and a negative result to approximately 1%. In other words, the SFAV performs well as a rule-out test across the plausible range of settings, which is exactly the property required of a first-line screen whose purpose is to decide who needs the longer instrument. Because the intervention that follows confirmation is balance and neuromuscular training, which is inexpensive and carries no meaningful harm, the consequences of the residual false-positive rate are modest. A practical pathway follows directly: ask the single question opportunistically; if the answer is 90 or above, no further ankle assessment is indicated on this basis; if it is below 90, administer the CAIT and act on the result.

Several features strengthen confidence in these findings. The sample was large, drawn from three geographically distinct centres, and recruited from individuals attending for reasons unrelated to their own ankles, which limits the spectrum bias that affects diagnostic accuracy studies conducted within specialist ankle clinics. Completeness of the reference standard was high, at 99.9%. The threshold was internally validated by bootstrapping, which showed negligible optimism and a stable optimal cut-off, and accuracy was consistent across age, sex and activity strata and robust to substantial changes in the definition of the reference standard.

The limitations are equally important and constrain what may be concluded. First and most fundamentally, history of ankle sprain and of giving way episodes was not collected. This means that the

criteria of the International Ankle Consortium could not be applied, that our findings cannot be extrapolated to chronic ankle instability as that consortium defines it, and that we cannot determine what proportion of the instability we detected was post-traumatic. A screening instrument for CAI would need to be validated against the full consortium criteria, and that study remains to be done. Second, the reference standard is itself a patient-reported instrument administered in the same sitting as the index test, so the two share method variance; agreement between them reflects concordance between two self-reports rather than against an external clinical or biomechanical criterion, and the estimates reported here should be understood as criterion validity against the CAIT rather than as diagnostic accuracy against a physical truth. The SFAV was moreover positioned after the CAIT in the questionnaire, so responses may have been anchored by the preceding items; this ordering could plausibly inflate agreement, and a randomised-order administration would be a useful refinement. Third, although index-test data were missing for only 1.7% of participants, none of those participants met the reference-standard definition, so the missingness was not entirely unrelated to ankle status; the bounding analysis shows the effect on specificity to be confined to a range of about two percentage points, but the estimate is not free of this assumption. Fourth, recruiting among the companions of orthopaedic patients yields a sample that is convenient and free of ankle-related selection, but it is not a random sample of the general population, and the observed prevalence should be read with that in mind. Fifth, the SFAV asks about “your ankle” without specifying a side, whereas instability is a side-specific phenomenon; we handled this by defining the reference standard at person level, but a side-specific version of the question might perform differently and deserves evaluation. Finally, the study was cross-sectional and used the French versions of all instruments with a French-derived CAIT threshold, so neither the reproducibility of the SFAV over time nor the transferability of the threshold to other languages can be inferred from these data.

Three lines of further work follow. The most pressing is prospective validation against the complete International Ankle Consortium criteria, including injury and giving-way history, which would establish whether the SFAV screens for chronic ankle instability as well as it screens for its perceived component. The second is a test–retest study of the SFAV itself, since no reproducibility data exist for the instrument in an asymptomatic population and a screening tool must be stable to be useful. The third, and ultimately the question that matters, is whether screening changes anything: a trial in which SFAV-positive individuals are randomised to neuromuscular training or usual advice would determine whether earlier detection of functional instability translates into fewer recurrent sprains and better long-term ankle health, which is the premise on which the entire screening argument rests.

Conclusions

In a multicentre sample of 717 adults recruited outside any ankle care pathway, a single question about perceived ankle function identified CAIT-defined functional ankle instability with high accuracy and outperformed a 21-item region-specific questionnaire. A Simple Foot and Ankle Value below 90 offers a practical first-line triage rule that captures more than nine in ten cases while sparing 61% of individuals the longer instrument, and its high negative predictive value makes it well suited to ruling out instability

in low-prevalence settings. These conclusions apply to functional ankle instability as operationalised by the CAIT; establishing whether the SFAV performs comparably against the full International Ankle Consortium definition of chronic ankle instability will require a study in which injury and giving-way history are recorded.

Abbreviations

AAS
Ankle Activity Score
AUC
area under the receiver operating characteristic curve
CAI
chronic ankle instability
CAIT
Cumberland Ankle Instability Tool
CI
confidence interval
FAAM
Foot and Ankle Ability Measure
FAAM–ADL
Foot and Ankle Ability Measure,Activities of Daily Living subscale
FAAM–Sport
Foot and Ankle Ability Measure,Sport subscale
FAI
functional ankle instability
ROC
receiver operating characteristic
SFAV
Simple Foot and Ankle Value
STARD
Standards for Reporting of Diagnostic Accuracy Studies.

Declarations

Ethics approval and consent to participate.

The study involved no intervention on participants and consisted solely of a self-administered questionnaire. Under French law (articles L.1121-1 and R.1121-1 of the French Public Health Code, as amended by Loi n° 2012 – 300 of 5 March 2012), research of this type is classified as research not involving the human person (*recherche n'impliquant pas la personne humaine*). Such research falls outside the remit of the Comités de Protection des Personnes, the committees designated by French law

to review research involving the human person, and requires no ethics committee opinion; the need for ethics approval was therefore waived in accordance with this national legislation, and the study was not submitted to an ethics committee or institutional review board. The collection and computerised processing of personal and medical data were carried out under reference methodology MR-004 of the French National Commission for Informatics and Liberties (CNIL), the framework governing this category of research in France. Toulouse University Hospital, the sponsoring institution, signed a commitment of compliance with this methodology (CNIL number 2407763 v 0). The study was evaluated and validated by the data protection officer of Toulouse University Hospital in accordance with Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, and is registered in the register of studies not involving the human person of Toulouse University Hospital under reference number RnIPH 2024-59. All participants received written information about the study, including its purpose, the voluntary nature of participation and their right to decline, and consent was implied by their completion and return of the questionnaire. All methods were performed in accordance with the Declaration of Helsinki.

Consent for publication.

Not applicable.

Competing interests.

The authors declare that they have no competing interests.

Funding.

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Author Contribution

MD contributed to conception and design, data acquisition, analysis and interpretation, statistical analysis, critical revision of the manuscript and study supervision. RL contributed to data acquisition and critical revision of the manuscript. AH contributed to data acquisition and critical revision of the manuscript. DA contributed to conception and design, critical revision of the manuscript, administration and study supervision, and approved the final version submitted. All authors read and approved the final manuscript.

Acknowledgements.

Not applicable.

Data Availability

The datasets generated and analysed during the current study are not publicly available because participants did not consent to public data sharing, but are available from the corresponding author on reasonable request.

References

1. Herzog MM, Kerr ZY, Marshall SW, Wikstrom EA. Epidemiology of ankle sprains and chronic ankle instability. J Athl Train. 2019;54(6):603–10.
2. Gribble PA, Bleakley CM, Caulfield BM, Docherty CL, Fourchet F, Fong DT, et al. 2016 consensus statement of the International Ankle Consortium: prevalence, impact and long-term consequences of lateral ankle sprains. Br J Sports Med. 2016;50(24):1493–5.
3. Doherty C, Bleakley C, Hertel J, Caulfield B, Ryan J, Delahunt E. Recovery from a first-time lateral ankle sprain and the predictors of chronic ankle instability: a prospective cohort analysis. Am J Sports Med. 2016;44(4):995–1003.
4. Valderrabano V, Hintermann B, Horisberger M, Fung TS. Ligamentous posttraumatic ankle osteoarthritis. Am J Sports Med. 2006;34(4):612–20.
5. Doherty C, Bleakley C, Delahunt E, Holden S. Treatment and prevention of acute and recurrent ankle sprain: an overview of systematic reviews with meta-analysis. Br J Sports Med. 2017;51(2):113–25.
6. Gribble PA, Delahunt E, Bleakley C, Caulfield B, Docherty CL, Fourchet F, et al. Selection criteria for patients with chronic ankle instability in controlled research: a position statement of the International Ankle Consortium. Br J Sports Med. 2014;48(13):1014–8.
7. Hiller CE, Kilbreath SL, Refshauge KM. Chronic ankle instability: evolution of the model. J Athl Train. 2011;46(2):133–41.
8. Hertel J, Corbett RO. An updated model of chronic ankle instability. J Athl Train. 2019;54(6):572–88.
9. Baldwin JN, McKay MJ, Hiller CE, Nightingale EJ, Moloney N, Burns J, et al. Correlates of perceived ankle instability in healthy individuals aged 8 to 101 years. Arch Phys Med Rehabil. 2017;98(1):72–9.
10. Doherty C, Delahunt E, Caulfield B, Hertel J, Ryan J, Bleakley C. The incidence and prevalence of ankle sprain injury: a systematic review and meta-analysis of prospective epidemiological studies. Sports Med. 2014;44(1):123–40.
11. Hubbard-Turner T, Turner MJ. Physical activity levels in college students with chronic ankle instability. J Athl Train. 2015;50(7):742–7.
12. Alexandre É, Monteiro D, SottoMayor R, Jacinto M, Silva FM, Cid L, et al. Assessing functional ankle instability in sport: a critical review and bibliometric analysis. Healthcare. 2024;12(17):1733.

13. Delahunt E, Bleakley CM, Bossard DS, Caulfield BM, Docherty CL, Doherty C, et al. Clinical assessment of acute lateral ankle sprain injuries (ROAST): 2019 consensus statement and recommendations of the International Ankle Consortium. Br J Sports Med. 2018;52(20):1304–10.
14. Hiller CE, Refshauge KM, Bundy AC, Herbert RD, Kilbreath SL. The Cumberland Ankle Instability Tool: a report of validity and reliability testing. Arch Phys Med Rehabil. 2006;87(9):1235–41.
15. Irving G, Neves AL, Dambha-Miller H, Oishi A, Tagashira H, Verho A, et al. International variations in primary care physician consultation time: a systematic review of 67 countries. BMJ Open. 2017;7(10):e017902.
16. Gilbart MK, Gerber C. Comparison of the subjective shoulder value and the Constant score. J Shoulder Elb Surg. 2007;16(6):717–21.
17. Bellas N, Cirino C, Cote MP, Sathe V, Geaney L. Validation of the Single Assessment Numeric Evaluation (SANE) score as an outcome measure by comparison to the revised Foot Function Index (rFFI). Foot Ankle Orthop. 2019;4(4):2473011419868953.
18. Barnavon T, Hardy A, Duguay T, Bouche PA, Lopes R. Description of the Simple Ankle Value: a simplified patient-reported outcome measure for the assessment of ankle and hindfoot function. Orthop J Sports Med. 2023;11(10):23259671231200498.
19. Marot V, Justo A, Guenego E, Klein D, Reina N, Cavaignac E, et al. Simple foot and ankle value: a simple evaluation correlated to the existing PROMs. Eur J Orthop Surg Traumatol. 2023;33(7):3011–7.
20. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.
21. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). Foot Ankle Int. 2005;26(11):968–83.
22. Goulart Neto AM, Maffulli N, Migliorini F, De Menezes FS, Okubo R. Validation of Foot and Ankle Ability Measure (FAAM) and the Foot and Ankle Outcome Score (FAOS) in individuals with chronic ankle instability: a cross-sectional observational study. J Orthop Surg Res. 2022;17(1):38.
23. Geerinck A, Beaudart C, Salvan Q, Van Beveren J, D'Hooghe P, Bruyère O, et al. French translation and validation of the Cumberland Ankle Instability Tool, an instrument for measuring functional ankle instability. Foot Ankle Surg. 2020;26(4):391–7.
24. Wright CJ, Arnold BL, Ross SE, Linens SW. Recalibration and validation of the Cumberland Ankle Instability Tool cutoff score for individuals with chronic ankle instability. Arch Phys Med Rehabil. 2014;95(10):1853–9.
25. Halasi T, Kynsburg A, Tállay A, Berkes I. Development of a new activity score for the evaluation of ankle instability. Am J Sports Med. 2004;32(4):899–908.
26. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 2018;126(5):1763–8.

Figures

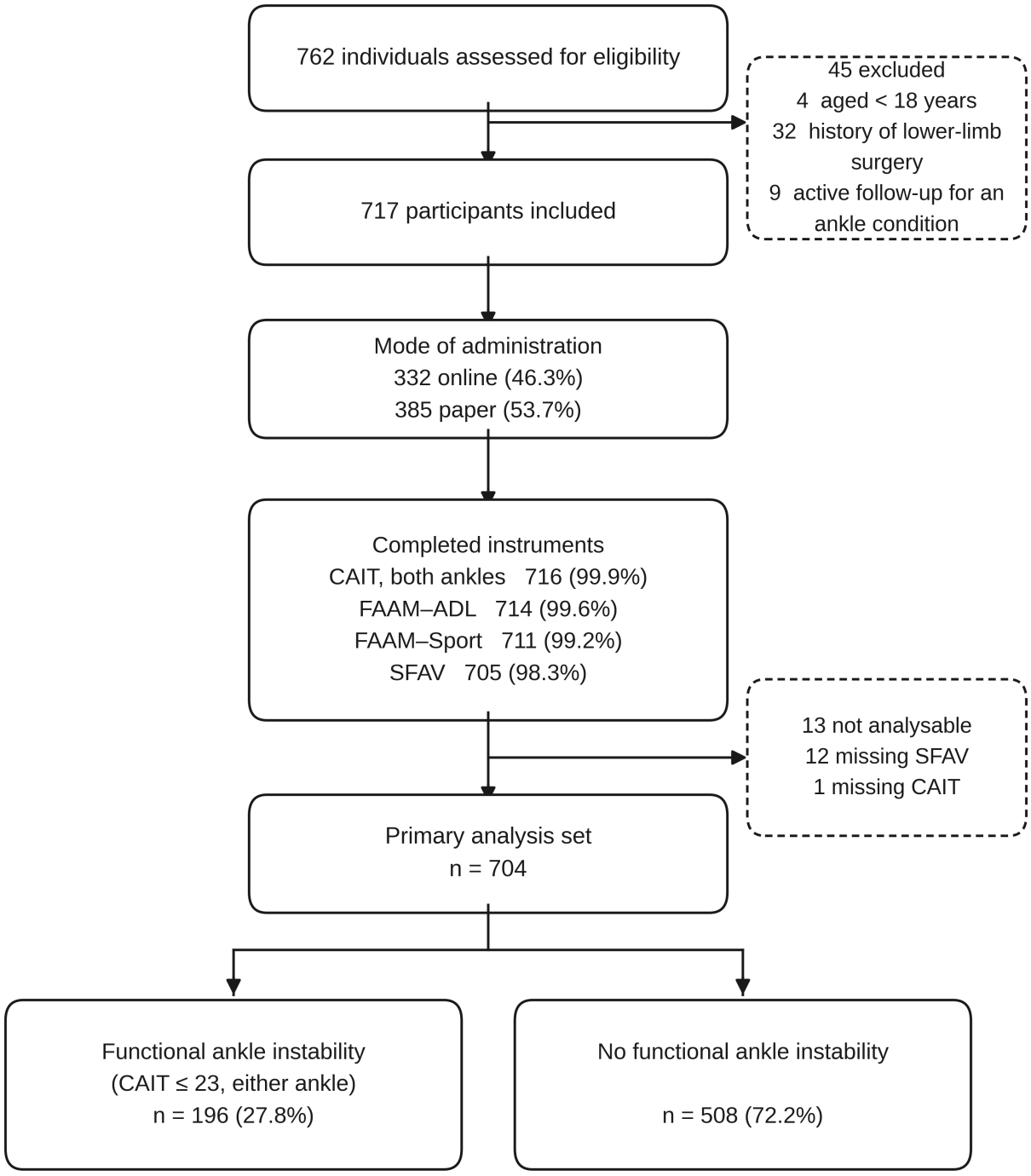


Figure 1

Flow of participants through the study.

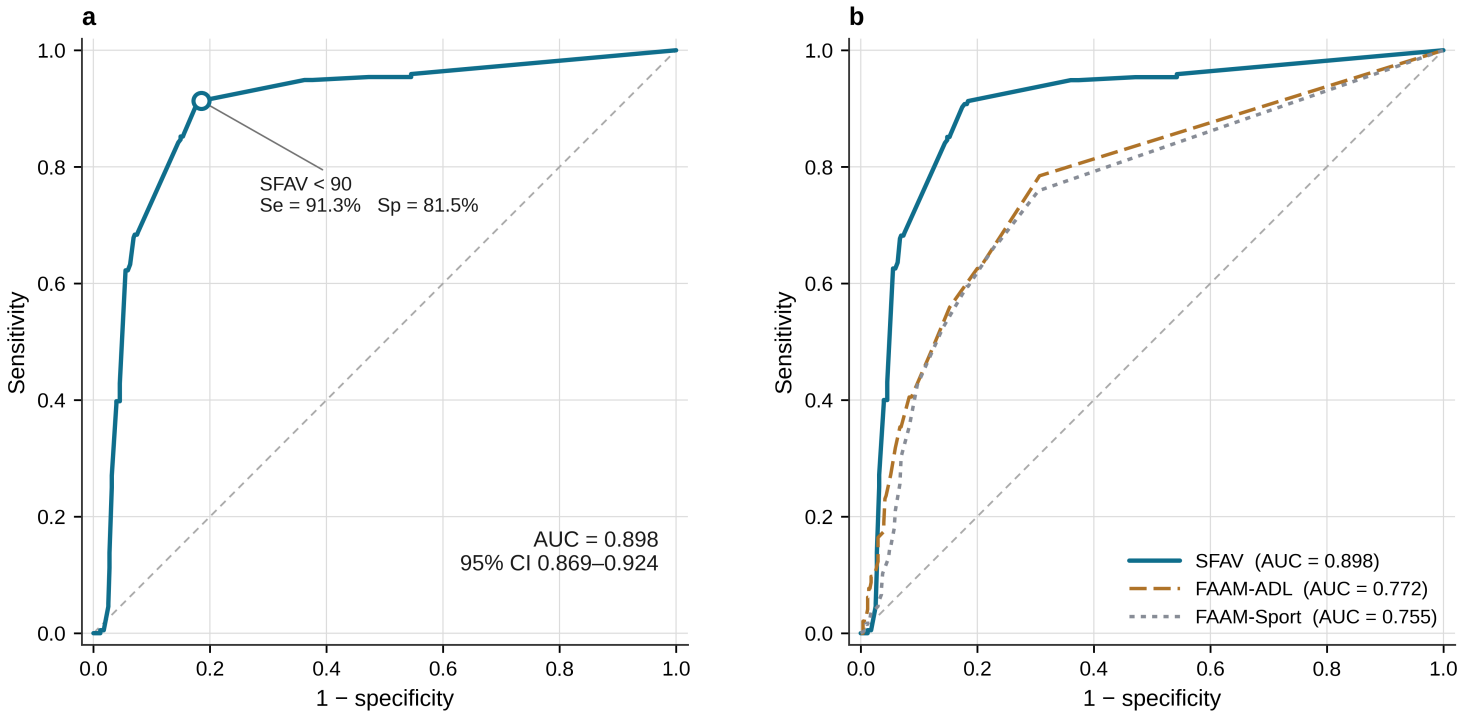


Figure 2

Discrimination of functional ankle instability. **a**Receiver operating characteristic curve for the Simple Foot and Ankle Value against the reference standard of a Cumberland Ankle Instability Tool score of 23 or less in at least one ankle, with the recommended operating point marked. **b**Comparison with the two Foot and Ankle Ability Measure subscales in the 700 participants with complete data for all three instruments.

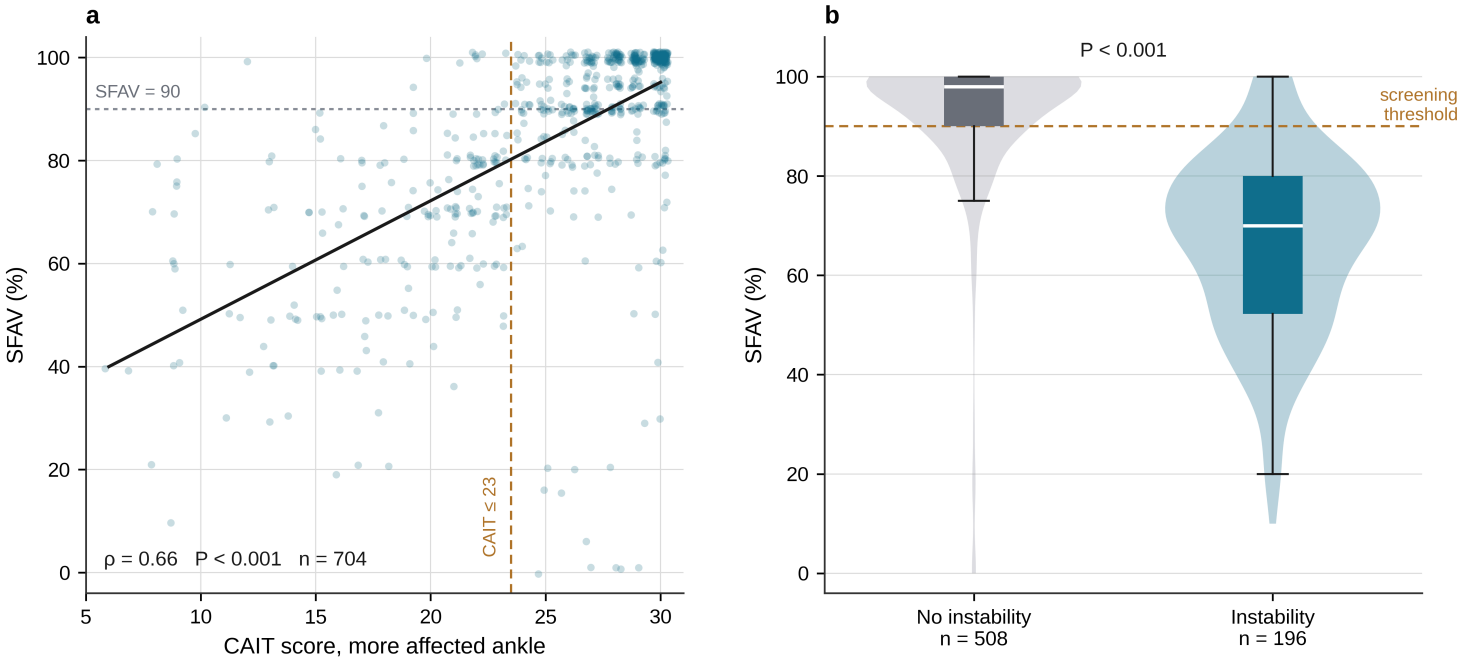


Figure 3

Relationship between the Simple Foot and Ankle Value and the reference standard. **a** Scatter plot of the Simple Foot and Ankle Value against the Cumberland Ankle Instability Tool score of the more affected ankle, with points jittered to reduce overplotting and the least-squares line shown; dashed lines mark the reference-standard and screening thresholds. **b**Distribution of Simple Foot and Ankle Value scores by reference-standard group, shown as violin plots with superimposed box plots.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 04Additionalfile1STARDchecklist.docx
- SFAVAdditionalfile2questionnaire.docx