Author/s:
Voelker, SN;Michalopoulos, N;Maier, AB;Reijnierse, EM

Title:
Reliability and Concurrent Validity of the SARC-F and Its Modified Versions: A Systematic Review and Meta-Analysis

Date:
2021-09-01

Citation:
Voelker, S. N., Michalopoulos, N., Maier, A. B. & Reijnierse, E. M. (2021). Reliability and Concurrent Validity of the SARC-F and Its Modified Versions: A Systematic Review and Meta-Analysis. Journal of the American Medical Directors Association, 22 (9), pp.1864-1876.e16. https://doi.org/10.1016/j.jamda.2021.05.011.

Persistent Link:
https://hdl.handle.net/11343/281046

JAMDA xxx (2021) 1–13

JAMDA

journal homepage: www.jamda.com

Review Article

Reliability and Concurrent Validity of the SARC-F and Its Modified Versions: A Systematic Review and Meta-Analysis

Stefanie N. Voelker BBMED (Hons.) [a], Nikolaos Michalopoulos MSc [b], Andrea B. Maier MD, PhD [a,b,c,d,*], Esmee M. Reijnierse PhD [a,e]

[a] *Department of Medicine and Aged Care, @AgeMelbourne, The Royal Melbourne Hospital, The University of Melbourne, Parkville, Melbourne, Victoria, Australia*
[b] *Department of Human Movement Sciences, @AgeAmsterdam, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, the Netherlands*
[c] *Healthy Longevity Program, Yong Loo Lin School of Medicine, National University of Singapore, Singapore*
[d] *Centre for Healthy Longevity, @AgeSingapore, National University Health System, Singapore*
[e] *Department of Rehabilitation Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, Amsterdam, The Netherlands*

Keywords:
Sarcopenia
muscular atrophy
SARC-F
screening
aged

A B S T R A C T

Objectives: Sarcopenia, being prevalent in up to 40% of older adults, is associated with adverse health outcomes. The international sarcopenia guidelines recommend screening for sarcopenia using the SARC-F. A previous meta-analysis (2017) reported poor validity of the SARC-F among community-dwelling older adults. Since then, modified SARC-F versions were developed and new sarcopenia definitions were published, including the SARC-F for case-finding. This systematic review and meta-analysis aimed to assess the reliability of the SARC-F and its concurrent validity to identify sarcopenia.
Design: Systematic review and meta-analyses.
Setting and Participants: Adults (all ages) from any study population.
Methods: A systematic search was conducted in MEDLINE, EMBASE, Cochrane, and CINAHL (January 1, 2013, to April 6, 2020). Articles were included if they reported on the reliability and/or concurrent validity of the (modified) SARC-F. No restrictions were applied for sex, age, study population, or sarcopenia definition. Reliability measures included inter-rater reliability, test-retest reliability, and internal consistency. Meta-analyses were performed for concurrent validity.
Results: The 29 included articles included 21,855 individuals (mean age of 63.3±14.6 years, 61.3% females) among community-dwelling (n = 16), geriatric inpatient (n = 5), geriatric outpatient (n = 2), nursing home (n = 2), and long-term care (n = 1) populations. The SARC-F had good (2/4 articles) to excellent (2/4 articles) inter-rater reliability, moderate (1/6 articles) to good (5/6 articles) test-retest reliability, and low (4/8 articles) to high (4/8 articles) internal consistency. The SARC-F had low to moderate sensitivity (28.9%-55.3%) and moderate to high specificity (68.9%-88.9%) according to the European Working Group on Sarcopenia in Older People (EWGSOP; n = 13), revised EWGSOP definition (EWGSOP2; n = 6), Asian Working Group for Sarcopenia (AWGS; n = 13), Foundation for the National Institutes of Health (FNIH; n = 8), International Working Group on Sarcopenia (IWGS; n = 9), and Society on Sarcopenia, Cachexia and Wasting Disorders (n = 2). The SARC-CalF had low to moderate sensitivity (45.9%-57.2%) and high specificity (87.7%-91.3%) according to the EWGSOP (n = 5), AWGS (n = 4), FNIH (n = 3), and IWGS (n = 3).
Conclusions and Implications: Despite the good reliability of the SARC-F, its low to moderate sensitivity and moderate to high specificity make it nonoptimal to use for sarcopenia screening. It is recommended to apply the diagnostic criteria for sarcopenia without screening.

S.N.V. and N.M. contributed equally.
Funding sources: This research was funded by an unrestricted grant of the University of Melbourne, Australia, received by Prof. Andrea B. Maier.
The authors declare no conflicts of interest.
* Address correspondence to Andrea B. Maier, MD, PhD, Department of Medicine and Aged Care, @Age, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, van der Boechorststraat 7, 1081 BT, Amsterdam, the Netherlands.
E-mail address: a.b.maier@vu.nl (A.B. Maier).

https://doi.org/10.1016/j.jamda.2021.05.011

Sarcopenia is a disease characterized by low skeletal muscle mass, muscle strength, and physical performance.[1,2] Its prevalence has been reported to be 1% to 29% in community-dwelling populations, 14% to 33% in long-term care populations, 2% to 34% in geriatric outpatients, and 40% in acute hospitalized geriatric inpatients.[3–5] Sarcopenia is associated with functional impairment, physical disability,[6,7] falls, fractures,[8] comorbidity,[9] hospitalization, and mortality.[10] Knowledge among health care professionals on how to diagnose sarcopenia and the availability of diagnostic equipment are currently limited,[11,12] despite the willingness of older adults to counteract sarcopenia.[13] Screening of older adults for sarcopenia might be an appropriate step to identify individuals at risk of sarcopenia.[14]

The SARC-F screening tool was introduced to identify older adults at risk of sarcopenia. It encompasses self-reported questions regarding strength, assistance with walking, rising from a chair, climbing stairs, and falls history.[15] The international sarcopenia guidelines recommend screening for sarcopenia using the SARC-F.[16] A previous meta-analysis reported poor concurrent validity of the SARC-F, however, only including 7 articles in community-dwelling populations.[17] Since then, the literature has significantly expanded, modified SARC-F versions were developed, new sarcopenia definitions were published, and the SARC-F was included for case-finding in the European Working Group on Sarcopenia in Older People revised definition (EWGSOP2).[1]

This systematic review and meta-analysis aimed to assess the reliability of the SARC-F and its concurrent validity as a screening tool for identifying adults with sarcopenia.

Methods

Data Sources and Searches

The protocol was registered at PROSPERO international prospective register of systematic reviews (CRD42020182062). The review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[18] A systematic search was conducted in MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials and CINAHL, and included articles published from January 1, 2013, to April 6, 2020. The SARC-F (0-10 points) was first published in 2013[15]; thus, the search was restricted to articles published from 2013 onwards. The search strategy (Supplementary Table 1) was developed in consultation with a senior tertiary librarian from a biomedical university library.

Article Selection

Articles were included if they were published in English, if reliability and/or concurrent validity data of the (modified) SARC-F were reported, and sarcopenia according to any definition (for articles reporting on concurrent validity). Exclusion criteria included the following: duplicates, full-text unavailable, no primary data (eg, protocol paper, consensus/statement paper), and conference abstracts and case reports (<5 cases). Articles using the same cohort were selected for inclusion in the following hierarchical manner: (1) larger sample size; (2) more reliability and concurrent validity measures reported; (3) more sarcopenia definitions reported; or (4) if more SARC-F versions were reported. No restrictions were applied for sex, age, or population. The screening of titles and abstracts and, subsequently, full-text articles, to select eligible articles was independently performed by 2 reviewers (S.N.V. and N.M.). A third reviewer (E.M.R.) resolved any disagreements.

Data Extraction and Quality Assessment

The following data were extracted independently by 2 reviewers (S.N.V. and N.M.): population; country of included participants; sample size; age; proportion of females; SARC-F version, score, cutoff points; proportion at risk of sarcopenia; reliability and concurrent validity measures; definition(s) of sarcopenia; sarcopenia prevalence; muscle mass measurements, that is, method, device, measure, cutoff points; muscle strength measurements, that is, method, device, and cutoff points; physical performance measurements, that is, method, instructions to participants, and cutoff points. Where sex-stratified data were provided for the mean age and/or the SARC-F score, the weighted value was calculated to determine the overall mean population values.

Two reviewers (S.N.V. and N.M.) independently assessed the risk of bias using a modified version of the Newcastle-Ottawa Scale (NOS)[19,20] (Supplementary Material 1). Articles were assessed on (1) selection of the study population, (2) comparability, and (3) ascertainment of the outcome. Although the "selection" and "comparability" categories applied to all articles, the "outcome" category was dependent on whether the article reported on reliability measures, concurrent validity measures, or both. The total scores varied depending on the reliability and/or concurrent validity measures reported. For articles reporting only on reliability (0-7 points), concurrent validity (0-5 points), or both (0-8 points), a relative score of 5 points was adopted across all articles. Because modified versions of the NOS were used, a median score of ≥5 points was considered as a low risk of bias.[21] Discrepancies over data extraction and risk of bias were resolved by a third reviewer (E.M.R.).

Data Synthesis

Reliability

Reliability measures included inter-rater reliability, test-retest reliability, and internal consistency. The inter-rater reliability and test-retest reliability were reported using intraclass correlation coefficient values (≥0.90 excellent, ≥0.75 to <0.90 good, ≥0.50 to <0.75 moderate, and <0.50 poor)[22] and Cohen kappa (>0.80 very good, >0.60 to ≤0.80 good, >0.40 to ≤0.60 moderate, >0.20 to ≤0.40 fair, and ≤0.20 poor).[23] Internal consistency was reported using Cronbach alpha coefficients (>0.70 high, ≤0.70 low).[24] Reliability measures were analyzed qualitatively and were not pooled because of asymmetrical 95% confidence interval (CI) values for inter-rater reliability and test-retest reliability, and no 95% CI values for internal consistency.

Concurrent validity

The concurrent validity of the SARC-F against sarcopenia definitions was reported using sensitivity, specificity, Youden Index, positive likelihood ratio (LR+), negative likelihood ratio (LR−), diagnostic odds ratio (DOR), area under the curve (AUC), positive predictive value (PPV), negative predictive value (NPV), and accuracy. Sensitivity and specificity were classified as follows: high ≥80%, moderate ≥50% to <80%, and low <50%.[25] The Youden Index (*J* statistic), where $J =$ (sensitivity + specificity) − 1, was classified as high if ≥0.6, and low if <0.6.[26] Because a negative *J* value implies the same diagnostic accuracy as $J = 0$, all calculated negative *J* values were reported as $J = 0$.[27] LR+ was classified as follows: large >10, moderate 5 to 10, and small <5.[28] LR− was classified as follows: large <0.1, moderate 0.1 to 0.5, and small >0.5.[28] DOR values were classified as no discriminatory ability if 1.0, low discriminatory ability if >1.0 to ≤4.0, moderate discriminatory ability if >4.0 to ≤10.0, and high discriminatory ability if >10.0.[28,29] AUC was classified as follows: high >0.80, moderate 0.60 to 0.80, and low <0.60.[25,30] Because no cutoff points were available for

Table 1
Characteristics of Included Articles, Grouped by SARC-F Versions

First Author, Year	Population	Country	Sample Size, N	Age, y Mean ± SD or %	Female, n (%)
SARC-F					
Bahat, 2018a*,[37]	CD	TUR	207	74.6 ± 6.7	140 (67.6)
Bahat, 2018*,[38]	CD	TUR	20	71.4 ± 8.1	10 (50.0)
Barbosa-Silva, 2016[39]	CD	BRA	179	60-69 y: 57.5% 70-79 y: 31.3% ≥80 y: 11.2%	110 (61.4)
Beaudart, 2018[40]	CD	BEL	306	74.8 ± 5.9	182 (59.6)
Borges, 2019[41]	IP	BRA	77	56.1 ± 12.0‖	32 (41.6)
Drey, 2020[42]	OP	GER	117	79.1 ± 5.2	94 (80.4)
Fu, 2020[43]	IPwC	CHN	309	54.7 ± 11.3	120 (38.8)
Ha, 2020[44]	IPwHF	KOR	115	80.2 ± 7.42‖	93 (80.9)
Ida, 2017[45]	OP	JPN	207	71.9 ± 5.44‖	81 (39.2)
Ida, 2019[46]	OP	JPN	140	70.8 ± 9.68‖	42 (30.0)
Kemmler, 2017[47]	CD	DEU	74	82.2 ± 5.6	0
Kera, 2019[48]	CD	JPN	733	73 ± 66.7	445 (60.7)
Kera, 2020[49]	CD	JPN	1060	73.1 ± 6.02‖	772 (72.8)
Kim, 2018[50]	CD	KOR	1222	76.0 ± 4.0	645 (52.8)
Kim, 2019[51]	CD	KOR	2099	75.9 ± 4.0	1046 (49.8)
Kotlarczyk, 2018[52]	LCR	USA	141	83.6 ± 7.0	141 (100)
Kurita, 2019[53]	IPwMSD	JPN	959	69.4 ± 9.3	699 (72.9)
Li, 2019[54]	IP	CHN	138	70.3 ± 8.82‖	69 (50.0)
Lima, 2020[55]	OPwPD	BRA	218	67.2 ± 10.9	93 (42.7)
Lim, 2018[56]	CD	SGP	200	67.9 ± 2.2	137 (68.5)
Malmstrom, 2016[57]	CD†	USA	853	59.2 ± 4.34‖	534 (62.6)
	CD‡	USA	1053	84.6 ± 9.17‖	500 (47.5)
	CD§	USA	3288	72.0 ± 7.86‖	1702 (51.8)
Mienche, 2019[58]	OP	IDN	120	71.9 ± 6.1	74 (61.7)
Parra-Rodriguez, 2016[59]	CD	MEX	487	73.2 ± 8.0	390 (80.1)
Rolland, 2017[60]	CD	FRA	2705	NR	2705 (100)
Sanchez-Rodriguez, 2019[61]	CD	ESP	90	81.4 ± 5.9	68 (75.6)
Woo, 2014[63]	CD	HKG	3997	72.5 ± 5.15‖	1998 (50.0)
Yang, 2018[64]	CD	CHN	384	71.5 ± 5.8	224 (58.3)
Yang, 2019[65]	NH	CHN	277	81.6 ± 3.3	194 (70.0)
SARC-CalF					
Bahat, 2018[37]	CD	TUR	207	74.6 ± 6.7	140 (67.6)
Barbosa-Silva, 2016[39]	CD	BRA	179	60-69 y: 57.5% 70-79 y: 31.3% ≥80 y: 11.2%	110 (61.4)
Fu, 2020[43]	IPwC	CHN	309	54.7 ± 11.3	120 (38.8)
Mienche, 2019[58]	OP	IDN	120	71.9 ± 6.1	74 (61.7)
Urzi, 2017[62,**]	NH	SVN	80	84.3 ± 7.9	56 (70.0)
Yang, 2018[64]	CD	CHN	384	71.5 ± 5.8	224 (58.3)
Yang, 2019[65]	NH	CHN	277	81.6 ± 3.3	194 (70.0)
SARC-F+TC					
Mienche, 2019[58]	OP	IDN	120	71.9 ± 6.1	74 (61.7)
SARC-F+CC+TC					
Mienche, 2019[58]	OP	IDN	120	71.9 ± 6.1	74 (61.7)
SARC-F+E					
Kurita, 2019[53]	IPwMSD	JPN	959	69.4 ± 9.3	699 (72.9)
SARC-F+EBM					
Kurita, 2019[53]	IPwMSD	JPN	959	69.4 ± 9.3	699 (72.9)
SARC-F-slowness					
Lim, 2018[56]	CD	SGP	200	67.9 ± 2.2	137 (68.5)

BEL, Belgium; BRA, Brazil; CD, community-dwellers; CHN, China; DEU, Germany; ESP, Spain; FRA, France; HKG, Hong Kong; IDN, Indonesia; IP, inpatients; IPwC, inpatients with cancer; IPwHF, inpatients with hip fracture; IPwMSD, inpatients with musculoskeletal disease; JPN, Japan; KOR, Korea; LCR, long-term care residents; MEX, Mexico; NH, nursing home residents; NR, not reported; OP, outpatients; OPwPD, outpatients with Parkinson's disease; SD, standard deviation; SGP, Singapore; SVN, Slovenia; TUR, Turkey.

*Same cohort: Bahat (2018)[37] concurrent validity, Bahat (2018)[38] reliability.

†African American Health cohort.

‡Baltimore Longitudinal Study of Aging cohort.

§National Health and Nutrition Examination cohort.

‖Calculated based on reported data.

**Article only reported on the concurrent validity data of the SARC-CalF.

PPV, NPV, and accuracy, and since these measures are reported on the same scale as sensitivity and specificity, the same cutoff points were applied as for sensitivity and specificity. The SARC-F or its modified versions were considered good when its sensitivity and specificity meet a high AUC or Youden Index value. In the case of sarcopenia, high sensitivity was prioritized over specificity because of the disease's serious yet preventable nature.[31] If concurrent validity measures were not reported, the values were calculated where possible, based on cross-tabulations of reported values of the SARC-F and sarcopenia (except for AUC). Authors were contacted if additional information was needed for inclusion in the meta-analysis.

A bivariate random effects model was used in the meta-analyses to assess the concurrent validity of the SARC-F.[32] When articles reported multiple diagnostic criteria of the same sarcopenia definition, the

concurrent validity data for the definition closest to the original definition were included in the meta-analyses. When articles reported on multiple cutoff points of the same SARC-F version, the concurrent validity data for the cutoff point closest to the original cutoff points (SARC-F: ≥4 points, SARC-CalF: ≥11 points) were included in the meta-analyses. When articles reported on probable, confirmed, and severe sarcopenia (EWGSOP2), the concurrent validity data for confirmed sarcopenia was included in the meta-analyses. All concurrent validity measures were pooled, except for likelihood ratios because separate pooling ignores correlations between LR+ and LR−,[33] and the Youden Index, DOR, and accuracy because they were calculated (ie, not reported in articles), hence lacking 95% CIs. Meta-analyses were stratified by sarcopenia definition for definitions included in at least 2 articles. Subgroup analyses were performed for SARC-F version, population, sex, and risk of bias. Heterogeneity in point estimates between the studies was assessed using the I^2 statistic, with I^2 values ≤25%, 25% to 75%, and ≥75% reflecting low, moderate, and high heterogeneity, respectively.[34] The Cochran Q value was used to assess between-group heterogeneity in point estimates, with P values <.05 of the Q value (Qb) considered statistically significant.[35]

Publication bias was visually evaluated by funnel plots showing the standard error against the point estimate and quantified using the Egger regression test,[36] with P values <.05 considered statistically significant (2-tailed). Analyses including fewer than 3 articles were excluded from the publication bias analysis.

Pooled estimates and forest plots were produced using Comprehensive Meta-Analysis (version 2.0; Biostat Inc, Englewood, NJ). Summary forest plots of the meta-analyses were created using GraphPad Prism 8.0.

Results

Search Results

A total of 3915 articles were screened for titles and abstracts, and 82 articles were identified for full-text screening. Twenty-nine articles[37–65] were included in this systematic review (Supplementary Figure 1). Eleven authors were contacted, of which 7 responded.

Study Characteristics

A total of 21,855 individuals were included (20 to 3,997 individuals per article), with a weighted mean age of 63.3 ± 14.6 years (range: 54.7 ± 11.3 years to 84.6 ± 9.17 years) and of which 61.3% were females (Table 1). Populations included community-dwelling older adults (n = 16; 18,957 individuals),[37–40,47–51,56,57,59–61,63,64] geriatric inpatients (n = 5; 1,598 individuals),[41,43,44,53,54] geriatric outpatients (n = 5; 802 individuals),[42,45,46,58] nursing home residents (n = 2; 357 individuals),[62,65] and long-term care residents (n = 1; 141 individuals).[52] Fifteen articles included individuals from Asia, 8 articles from Europe, 3 articles from South America, and 3 articles from North America. Three articles reported on reliability, 20 on concurrent validity, and 6 on both reliability and concurrent validity. Twenty articles[37–39,42,45,47–52,54,56–58,60–62,64,65] had a low risk of bias (Supplementary Table 2).

Reliability

Four articles reported on inter-rater reliability, 6 on test-retest reliability and 6 (8 cohorts) on internal consistency (Table 2). The inter-rater reliability of the SARC-F was excellent in 2 articles[40,42] and good in 2 articles.[38,61] Test-retest reliability was good in 5 articles[38,40,42,45,59] and moderate in 1 article.[61] Internal consistency was low in 4 articles[42,48,56,59] and high in 2 articles (including 4 cohorts).[57,61] One article[56] reported low internal consistency of the SARC-F—slowness, a modified SARC-F version.

Concurrent Validity

The prevalence of sarcopenia ranged from 1.8% to 66.2% depending on the applied definition. The prevalence of individuals at risk of sarcopenia ranged from 2.8% to 84.2% depending on the SARC-F version. Concurrent validity measures are reported in Table 3 and Supplementary Table 3. The following sarcopenia definitions were

Table 2
Reliability Measures of SARC-F Versions

First Author, Year	SARC-F			Inter-rater Reliability		Test-Retest Reliability		Internal Consistency
	Cutoff	Score, Mean ± SD	Prevalence, *n* (%)	ICC (95% CI)	Cohen Kappa	ICC (95% CI)	Cohen Kappa	Cronbach Alpha
SARC-F (0-10 points)								
Bahat, 2018[38]	≥4	1.71 ± 2.30*	NR	0.78 (0.60, 0.89)	NR	0.78 (0.61, 0.90)	NR	NR
Beaudart, 2018[40]	≥4	1.69 ± 2.00*	52 (17.0)	0.90 (0.76, 0.96)	NR	0.86 (0.66, 0.94)	NR	NR
Drey, 2020[42]	≥4	NR	63 (53.8)	0.93 (0.82, 0.97)	NR	0.90 (0.78, 0.95)	NR	0.672
Ida, 2017[45]	≥4	NR	41 (19.8)	NR	NR	NR	0.660	NR
Kera, 2019[48]	≥4	4.86 ± 1.06*	24 (3.3)	NR	NR	NR	NR	0.610
Lim, 2018[56]	≥4	0.45 ± 0.74	10 (0.5)	NR	NR	NR	NR	0.093
Malmstrom, 2016[57]	≥4	NR	157 (18.4)†	NR	NR	NR	NR	0.810
	≥4	NR	66 (6.3)‡	NR	NR	NR	NR	0.780
	≥4	NR	505 (15.4)§	NR	NR	NR	NR	0.760
Parra-Rodriguez, 2016[59]	≥4	1.95 ± 1.90	95 (19.5)	NR	NR	0.80 (NR)	NR	0.641
Sanchez-Rodriguez, 2019[61]	≥4	3.80 ± 2.30	51 (56.7)	NR	0.603	NR	0.588	0.779
SARC-F-slowness (0-10 points)								
Lim, 2018[56]	≥4	0.51 ± 0.83	4 (0.2)	NR	NR	NR	NR	0.244

ICC, intraclass correlation coefficient; NR, not reported; SD, standard deviation.
*Weighted score calculated based on reported sex-stratified scores.
†African American Health cohort.
‡Baltimore Longitudinal Study of Aging cohort.
§National Health and Nutrition Examination cohort.

Table 3
Concurrent Validity Measures of SARC-F Versions to Identify Sarcopenia

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Youden Index (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
SARC-F (0-10 points)									
Bahat, 2018[37]	EWGSOP, 2010	8 (3.86)	≥4	39 (18.8)	25.0 (0.00, 55.0)	81.4 (76.0, 86.8)	0.064 (NR)*	5.13 (−1.79, 12.1)*	96.4 (93.6, 99.2)*
	FNIH**, 2014	19 (9.18)	≥4	39 (18.8)	31.6 (10.6, 52.5)	82.4 (77.0, 87.9)	0.140 (NR)*	15.4 (10.7, 52.5)*	92.3 (88.2, 96.3)*
	IWGS, 2011	4 (1.93)	≥4	39 (18.8)	50.0 (1.00, 99.0)	81.8 (76.5, 87.1)	0.318 (NR)*	5.13 (−1.79, 12.1)*	98.8 (97.2, 100)*
	SCWD, 2011	5 (2.42)	≥4	39 (18.8)	40.0 (0.00, 82.9)	81.7 (76.3, 87.0)	0.217 (NR)*	5.13 (−1.79, 12.1)*	98.2 (96.2, 100)*
Barbosa-Silva, 2016[39]	EWGSOP, 2010	15 (8.38)	≥6	31 (17.3)	33.3 (11.8, 61.6)	84.2 (77.6, 89.4)	0.175 (NR)*	16.1 (5.5, 33.7)	93.2 (87.9, 96.7)
Beaudart, 2018[40]	EWGSOP, 2010	51 (16.7)	≥4	52 (17.0)	36.0 (30.6, 41.4)	87.1 (83.3, 90.9)	0.231 (NR)*	35.3 (29.9, 40.7)	87.4 (83.7, 91.1)
	AWGS, 2014	12 (3.92)	≥4	52 (17.0)	75.0 (70.1, 79.9)	85.4 (81.4, 89.4)	0.604 (NR)*	17.3 (13.1, 21.5)	94.9 (92.4, 97.4)
	FNIH**, 2014	11 (3.59)	≥4	52 (17.0)	40.9 (35.4, 46.4)	84.9 (80.9, 88.9)	0.258 (NR)*	17.3 (13.1, 21.5)	98.8 (97.6, 100)
	IWGS, 2011	37 (12.1)	≥4	52 (17.0)	43.2 (37.6, 48.8)	86.6 (82.8, 90.4)	0.298 (NR)*	30.8 (25.6, 36.0)	91.7 (88.6, 94.8)
	SCWD, 2011	18 (5.90)	≥4	52 (17.0)	55.6 (50.0, 61.2)	85.4 (81.4, 89.4)	0.410 (NR)*	19.2 (14.8, 23.6)	96.8 (94.8, 98.8)
	Baumgartner, 1998	81 (26.5)	≥4	52 (17.0)	22.2 (17.5, 26.9)	84.9 (80.9, 88.9)	0.071 (NR)*	34.6 (29.3, 39.9)	75.2 (70.4, 80.0)
	Delmonico, 2007	104 (34.0)	≥4	52 (17.0)	22.1 (17.5, 26.7)	85.6 (81.7, 89.5)	0.077 (NR)*	44.2 (38.6, 49.8)	68.1 (62.9, 73.3)
Borges, 2019[41]	EWGSOP2‡, 2018	47 (41.9)	≥4	32 (41.6)	59.4 (NR)*	28.9 (NR)*	0.00 (NR)*	61.3 (NR)*	28.3 (NR)*
	EWGSOP2§, 2018	30 (38.5)	≥4	39 (50.6)	30.8 (NR)*	71.1 (NR)*	0.019 (NR)*	38.7 (NR)*	58.7 (NR)*
Drey, 2020[42]	EWGSOP, 2010	8 (6.84)	≥4	63 (53.8)	50 (16, 84)	54 (44, 64)	0.040 (NR)*	7 (4, 14)	94 (88, 97)
	EWGSOP2‡, 2018	49 (41.9)	≥4	63 (53.8)	75 (62, 86)	67 (53, 78)	0.420 (NR)*	68 (59, 76)	74 (64, 82)
	EWGSOP2§, 2018	8 (6.8)	≥4	63 (53.8)	63 (25, 92)	47 (37, 57)	0.100 (NR)*	8 (5, 13)	94 (87, 98)
	EWGSOP, 2010	NR	≥3	NR	69 (59, 77)	38 (9, 76)	0.070 (NR)*	8 (3, 18)	94 (90, 96)
	EWGSOP2‡, 2018	49 (41.9)	≥3	NR	86 (74, 94)	48 (35, 62)	0.340 (NR)*	61 (55, 67)	78 (64, 88)
	EWGSOP2§, 2018	8 (6.8)	≥3	NR	75 (35, 97)	32 (24, 42)	0.070 (NR)*	9 (5, 11)	95 (84, 98)
	EWGSOP, 2010	NR	≥2	NR	88 (47, 100)	23 (15, 32)	0.110 (NR)*	8 (6, 10)	96 (80, 99)
	EWGSOP2‡, 2018	49 (41.9)	≥2	NR	91 (81, 97)	35 (23, 48)	0.260 (NR)*	57 (52, 62)	81 (63, 91)
	EWGSOP2§, 2018	8 (6.8)	≥2	NR	88 (47, 100)	23 (15, 32)	0.110 (NR)*	8 (6, 10)	96 (80, 99)
Fu, 2020[43]	EWGSOP2§, 2018	156 (50.5)	≥4	77 (24.9)	22.4 (16.2, 29.8)	92.1 (86.7, 95.9)	0.145 (NR)*	67.6 (55.5, 67.5)	61.7 (55.5, 67.5)
	AWGS, 2014	81 (26.2)	≥4	77 (24.9)	32.1 (22.2, 43.4)	90.7 (86.3, 94.2)	0.228 (NR)*	55.0 (40.3, 69.8)	79.0 (73.5, 83.7)
Ha, 2020[44]	EWGSOP, 2010	36 (31.3)	≥4	72 (62.6)	88.4 (78.8, 98.0)†	43.0 (31.6, 54.5)†	0.314 (NR)*	48.1 (37.1, 59.1)†	86.1 (74.8, 97.4)†
	EWGSOP2§, 2018	43 (37.4)	≥4	72 (62.6)	95.4 (89.1, 100)†	56.9 (45.5, 68.4)†	0.523 (NR)*	56.9 (45.5, 68.4)†	95.3 (89.1, 100)†
	AWGS, 2014	35 (30.4)	≥4	72 (62.6)	82.9 (70.3, 95.3)†	46.3 (35.3, 57.2)†	0.291 (NR)*	40.3 (28.9, 51.6)†	86.1 (75.7, 96.4)†
	IWGS, 2011	60 (52.2)	≥4	72 (62.6)	71.7 (60.3, 83.1)†	47.3 (34.1, 60.5)†	0.189 (NR)*	59.7 (48.4, 71.1)†	60.5 (45.9, 75.1)†
Ida, 2017[45]	EWGSOP, 2010	T: 53 (25.6)	≥4	41 (19.8)	18.9 (8.33, 29.4)*	79.9 (73.5, 86.2)*	0.00 (NR)*	24.4 (11.2, 37.5)*	74.1 (67.4, 80.8)*
		M: 41 (32.5)		18 (14.3)	14.6 (3.82, 25.5)†	85.8 (78.5, 93.3)†	0.004 (NR)*	33.3 (11.6, 55.1)†	67.5 (58.8, 76.4)†
		F: 12 (14.8)		23 (28.4)	33.3 (6.66, 60.0)†	72.4 (61.9, 83.0)†	0.057 (NR)*	17.3 (1.90, 32.9)†	86.2 (77.3, 95.1)†
Ida, 2019[46]	JSH, 2016	T: 75 (53.6)	≥4	22 (15.7)	24.0 (14.3, 33.7)*	93.8 (88.0, 99.7)*	0.178 (NR)*	81.8 (65.7, 97.9)*	51.7 (42.7, 60.7)*
		M: 55 (56.1)		11 (11.2)	16.3 (6.59, 26.1)†	95.3 (89.1, 102)†	0.116 (NR)*	81.8 (59.0, 105)†	47.1 (36.6, 57.6)†
		F: 20 (47.6)		11 (26.2)	45.0 (23.2, 66.8)†	90.9 (78.9, 103)†	0.359 (NR)*	81.8 (59.0, 105)†	64.5 (47.7, 81.4)†
Kemmler, 2017[47]	EWGSOP, 2010	M: 49 (66.2)	≥4	35 (33.5)	38.8 (25.1, 52.4)*	76.0 (59.3, 92.7)*	0.148 (NR)*	76.0 (59.3, 92.7)*	38.8 (25.1, 52.4)*
	FNIH‡‡, 2014	M: 32 (43.2)	≥4	35 (33.5)	40.6 (23.6, 57.6)*	71.4 (57.8, 85.1)*	0.120 (NR)*	52.0 (32.4, 71.6)*	61.2 (47.6, 74.9)*
	IWGS, 2011	M: 37 (50.0)	≥4	35 (33.5)	54.1 (38.0, 70.1)*	86.5 (75.5, 97.5)*	0.406 (NR)*	80.0 (64.3, 95.7)*	65.3 (52.0, 78.6)*
Kera, 2019[48]	EWGSOP, 2010	T: 75 (10.2)	≥4	24 (3.3)	8.0 (1.86, 14.1)†	97.3 (96.0, 98.5)†	0.053 (NR)*	25.0 (7.68, 42.3)†	90.3 (88.1, 92.5)†
		M: 66 (22.9)		9 (3.1)	7.58 (1.19, 14.0)*	98.2 (96.4, 100)*	0.058 (NR)*	55.6 (23.1, 88.0)*	78.1 (73.3, 83.0)*
		F: 9 (2.0)		15 (3.4)	11.1 (−9.42, 31.6)*	96.8 (95.1, 98.4)*	0.079 (NR)*	6.67 (−5.96, 19.3)*	98.1 (96.9, 99.4)*
	AWGS, 2014	T: 75 (10.2)	≥4	24 (3.3)	5.30 (0.25, 10.4)†	97.0 (95.6, 98.2)†	0.023 (NR)*	16.7 (1.76, 31.6)†	90.0 (87.8, 92.2)†
		M: 19 (6.60)		9 (3.1)	5.26 (−4.78, 15.3)*	97.0 (95.0, 99.1)*	0.023 (NR)*	11.1 (−9.42, 31.6)*	93.5 (90.7, 96.4)*
		F: 56 (12.6)		15 (3.4)	5.36 (−0.54, 11.3)*	96.9 (95.2, 98.6)*	0.023 (NR)*	20.0 (−0.24, 40.2)*	87.7 (84.6, 90.8)*
	Japanese AWGS, 2012	T: 167 (22.8)	≥4	24 (3.3)	6.00 (2.39, 9.59)†	97.5 (96.2, 98.8)†	0.035 (NR)*	41.7 (21.9, 61.4)†	77.9 (74.8, 80.9)†
		M: 52 (18.1)		9 (3.1)	3.85 (−1.38, 9.07)*	97.0 (94.9, 99.2)*	0.009 (NR)*	22.2 (−4.94, 49.4)*	82.1 (77.6, 86.6)*
		F: 115 (25.8)		15 (3.4)	6.96 (2.31, 11.6)*	97.9 (96.3, 99.4)*	0.049 (NR)*	53.3 (28.1, 78.6)*	75.1 (71.0, 79.2)*
Kera, 2020[49]	AWGS, 2014	129 (12.2)	≥4	30 (2.8)	3.9 (0.55, 7.21)†	97.3 (96.3, 98.4)†	0.012 (NR)*	16.7 (3.33, 30.0)*	88.0 (86.0, 89.9)*

(*continued on next page*)

Table 3 (*continued*)

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Youden Index (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
Kim, 2018[50]	EWGSOP, 2010	T: 95 (7.77)	≥4	123 (10.1)	25.3 (16.5, 34.0)†	91.2 (89.6, 92.9)†	0.165 (NR)*	19.5 (12.5, 26.5)†	93.5 (92.1, 95.0)†
		M: 48 (8.3)		24 (4.2)	16.7 (6.12, 27.2)†	97.0 (95.5, 98.4)†	0.137 (NR)*	33.3 (14.5, 52.2)†	92.8 (90.6, 94.9)†
		F: 47 (7.3)		99 (15.3)	34.0 (20.5, 47.6)†	86.1 (83.3, 88.9)†	0.201 (NR)*	16.2 (8.91, 23.4)†	94.3 (92.4, 96.3)†
	AWGS, 2014	T: 125 (10.2)	≥4	123 (10.1)	24.0 (16.5, 31.5)†	91.5 (89.9, 93.2)†	0.155 (NR)*	24.4 (16.8, 32.0)†	91.4 (89.7, 93.0)†
		M: 66 (11.4)		24 (4.2)	16.7 (7.68, 25.7)†	97.5 (96.1, 98.8)†	0.142 (NR)*	45.8 (25.9, 65.8)†	90.1 (87.6, 92.5)†
		F: 59 (9.1)		99 (15.3)	32.2 (20.3, 44.1)†	86.3 (83.6, 89.1)†	0.185 (NR)*	19.2 (11.4, 26.9)†	92.7 (90.5, 94.9)†
	FNIH††, 2014	T: 95 (7.77)	≥4	123 (10.1)	17.9 (10.2, 25.6)†	90.6 (88.9, 92.3)†	0.085 (NR)*	13.8 (7.72, 19.9)†	92.9 (91.4, 94.4)†
		M: 59 (10.2)		24 (4.2)	11.9 (3.61, 20.1)†	96.7 (95.2, 98.3)†	0.086 (NR)*	29.2 (11.0, 47.4)†	90.6 (88.2, 93.0)†
		F: 36 (5.6)		99 (15.3)	27.8 (13.1, 42.4)†	85.4 (82.6, 88.2)†	0.132 (NR)*	10.1 (4.17, 16.0)†	95.2 (93.5, 97.0)†
	FNIH‡‡, 2014	T: 23 (1.88)	≥4	123 (10.1)	43.5 (23.2, 63.7)†	90.6 (88.9, 92.2)†	0.341 (NR)*	8.1 (3.30, 13.0)†	98.8 (98.2, 99.5)†
		M: 10 (1.7)		24 (4.2)	60.0 (29.6, 90.4)†	96.8 (95.4, 98.3)†	0.568 (NR)*	25.0 (7.68, 42.3)†	99.3 (98.6, 100)†
		F: 13 (2.0)		99 (15.3)	30.8 (5.68, 55.9)†	85.0 (82.2, 87.8)†	0.158 (NR)*	4.04 (0.16, 7.92)†	98.4 (97.3, 99.4)†
	IWGS, 2011	T: 167 (13.7)	≥4	123 (10.1)	26.8 (20.1, 33.5)†	92.6 (91.0, 94.2)†	0.194 (NR)*	36.6 (28.1, 45.1)†	88.8 (86.9, 90.7)†
		M: 74 (12.8)		24 (4.2)	18.9 (1.00, 27.8)†	98.0 (96.7, 99.2)†	0.169 (NR)*	58.3 (38.6, 78.1)†	89.2 (86.6, 91.7)†
		F: 94 (14.6)		99 (15.3)	33.0 (23.5, 42.5)†	87.7 (84.9, 90.4)†	0.207 (NR)*	31.3 (22.2, 40.4)†	88.5 (85.8, 91.1)†
	CFNIH, 2015	T: 78 (6.38)	≥4	123 (10.1)	19.2 (10.5, 28.0)†	90.6 (88.9, 92.3)†	0.098 (NR)*	12.2 (6.41, 18.0)†	94.3 (92.9, 95.6)†
		M: 46 (8.0)		24 (4.2)	13.0 (3.31, 22.8)†	96.6 (95.1, 98.1)†	0.096 (NR)*	25.0 (7.68, 42.3)†	92.8 (90.6, 94.9)†
		F: 32 (5.0)		99 (15.3)	28.1 (12.5, 43.7)†	85.3 (82.5, 88.1)†	0.134 (NR)*	9.09 (3.43, 14.8)†	95.8 (94.1, 97.5)†
Kim, 2019[51]	EWGSOP2§, 2018	195 (9.3)	≥4	152 (7.2)	15.4 (10.3, 20.4)*	93.6 (92.5, 94.7)*	0.090 (NR)*	19.7 (13.4, 26.1)*	91.5 (90.3, 92.8)*
	EWGSOP2‖, 2018	38 (1.8)	≥4	152 (7.2)	42.1 (26.4, 57.8)*	93.4 (92.3, 94.5)*	0.355 (NR)*	10.5 (5.65, 15.4)*	98.9 (98.4, 99.3)*
Kotlarczyk, 2018[52]	EWGSOP, 2010	F: 11 (7.80)	≥4	30 (21.3)	18.2 (−4.61, 41.0)†	78.7 (71.4, 85.5)†	0.00 (NR)*	6.67 (−2.26, 15.6)*	91.9 (86.8, 97.0)*
	FNIH††, 2014	F: 6 (4.25)	≥4	30 (21.3)	33.3 (−4.39, 71.1)†	79.3 (72.4, 86.1)†	0.126 (NR)*	6.67 (−2.26, 15.6)*	96.4 (92.9, 99.9)*
	FNIH§§, 2014	F: 46 (32.6)	≥4	30 (21.3)	26.1 (13.4, 38.8)†	81.1 (73.2, 88.9)†	0.072 (NR)*	40.0 (22.5, 57.5)*	69.4 (60.8, 77.9)*
Kurita, 2019[53]	EWGSOP2§, 2018	19 (1.98)	≥6	67 (6.99)	26.3 (9.2, 51.2)	93.4 (91.6, 94.9)	0.197 (NR)*	7.46 (1.17, 13.8)*	98.4 (97.6, 99.2)*
	EWGSOP2§, 2018	19 (1.98)	≥5	161 (16.8)	26.3 (9.2, 51.2)	83.4 (80.9, 85.7)	0.097 (NR)*	3.11 (0.426, 5.79)*	98.2 (97.3, 99.2)*
	EWGSOP2§, 2018	19 (1.98)	≥4	306 (31.9)	47.4 (24.4, 71.1)	68.4 (65.3, 71.4)	0.158 (NR)*	2.94 (1.05, 4.83)*	98.5 (97.5, 99.4)*
	EWGSOP2§, 2018	19 (1.98)	≥3	454 (47.3)	57.9 (33.5, 79.7)	52.9 (49.6, 56.1)	0.108 (NR)*	2.42 (1.01, 3.84)*	98.4 (97.3, 99.5)*
	EWGSOP2§, 2018	19 (1.98)	≥2	623 (65.0)	57.9 (33.5, 79.7)	34.9 (31.8, 38.0)	0.00 (NR)*	1.77 (0.731, 2.80)*	97.6 (96.0, 99.2)*
	EWGSOP2§, 2018	19 (1.98)	≥1	815 (85.0)	84.2 (60.4, 96.6)	15.0 (12.8, 17.4)	0.00 (NR)*	1.96 (1.01, 2.92)*	97.9 (95.6, 100)*
	AWGS, 2014	36 (3.75)	≥6	67 (6.99)	19.4 (8.2, 36.0)	93.5 (91.7, 95.0)	0.129 (NR)*	10.4 (3.12, 17.8)*	96.7 (95.6, 97.9)*
	AWGS, 2014	36 (3.75)	≥5	161 (16.8)	22.2 (10.1, 39.2)	83.4 (80.9, 85.8)	0.056 (NR)*	4.97 (1.61, 8.33)*	96.5 (95.2, 97.8)*
	AWGS, 2014	36 (3.75)	≥4	306 (31.9)	41.7 (25.5, 59.2)	68.5 (65.4, 71.5)	0.102 (NR)*	4.90 (2.48, 7.32)*	96.8 (95.4, 98.1)*
	AWGS, 2014	36 (3.75)	≥3	454 (47.3)	61.1 (43.5, 76.9)	53.2 (49.9, 56.5)	0.143 (NR)*	4.85 (2.87, 6.82)*	97.2 (95.8, 98.7)*
	AWGS, 2014	36 (3.75)	≥2	623 (65.0)	63.9 (46.2, 79.2)	35.0 (31.9, 38.2)	0.00 (NR)*	3.69 (2.21, 5.17)*	96.1 (94.1, 98.2)*
	AWGS, 2014	36 (3.75)	≥1	815 (85.0)	86.1 (70.5, 95.3)	15.1 (12.8, 17.5)	0.012 (NR)*	3.80 (2.49, 5.12)*	96.5 (93.5, 99.5)*
Li, 2019[54]	AWGS, 2014	35 (25.4)	≥4	23 (16.7)	42.9 (26.5, 59.3)†	92.2 (87.1, 97.4)†	0.351 (NR)*	65.2 (45.8, 84.7)*	82.6 (75.7, 89.5)*
Lima, 2020[55]	EWGSOP2‡, 2018	123 (56.4)	≥4	121 (55.5)	21.1 (13.9, 28.4)*	0.00 (0.00, 0.00)*	0.00 (NR)*	21.5 (14.2, 28.8)*	0.00 (0.00, 0.00)*
Mienche, 2019[58]	AWGS, 2014	19 (15.8)	≥4	32 (26.7)	31.6 (13, 57)	74.3 (65, 82)	0.058 (NR)*	18.8 (10.0, 33.0)	85.2 (81.0, 89.0)
Parra-Rodriguez, 2016[59]	EWGSOP, 2010	45 (9.30)	≥4	94 (19.4)	35.6 (23.5, 47.7)†	82.2 (79.1, 85.3)†	0.178 (NR)*	17.0 (10.5, 23.6)†	92.6 (90.3, 94.8)†
	AWGS, 2014	54 (11.2)	≥4	93 (19.3)	31.5 (19.1, 43.9)†	82.1 (78.7, 85.9)†	0.136 (NR)*	18.2 (10.4, 26.1)†	90.5 (87.6, 93.4)†
	IWGS, 2011	99 (20.8)	≥4	91 (19.1)	28.3 (19.4, 37.2)†	83.3 (79.5, 87.1)†	0.116 (NR)*	30.8 (21.3, 40.3)†	81.6 (77.7, 85.4)†
Rolland, 2017[60]	FNIH††, 2014	F: 49 (1.8)	≥4	401 (14.8)	34.0 (21.4, 48.0)†	85.0 (84.1, 86.8)†	0.190 (NR)*	4.2 (2.27, 6.21)†	98.6 (98.1, 99.1)
Sanchez-Rodriguez, 2019[61]	EWGSOP, 2010	16 (17.8)	≥4	51 (56.7)	81.3 (62.1, 100)†	48.7 (37.3, 60.0)†	0.300 (NR)*	25.5 (13.5, 37.5)†	92.3 (83.9, 101)†
	EWGSOP2§, 2018	23 (25.6)	≥4	51 (56.7)	78.3 (61.4, 95.1)†	50.8 (38.8, 62.7)†	0.291 (NR)*	35.3 (22.2, 48.4)†	87.2 (76.7, 97.7)†
Woo, 2014[63]	EWGSOP, 2010	T: 361 (9.0)	≥4	150 (3.8)	6.9 (4.3, 9.5)*	96.6 (96.0, 97.2)*	0.035 (NR)*	16.7 (10.7, 22.6)*	91.3 (90.4, 92.2)*
		M: 190 (9.5)		31 (1.6)	4.2 (1.4, 7.1)†	98.7 (98.2, 99.2)†	0.029 (NR)*	25.8 (10.4, 41.2)†	90.8 (89.5, 92.0)†
		F: 171 (8.6)		119 (6.0)	9.9 (5.5, 14.4)†	94.4 (93.4, 95.5)†	0.043 (NR)*	14.3 (8.00, 20.6)†	91.8 (90.6, 93.0)†
	AWGS, 2014	T: 293 (7.3)	≥4	150 (3.8)	6.5 (3.7, 9.3)*	96.5 (95.9, 97.1)*	0.030 (NR)*	12.7 (7.34, 18.0)*	92.9 (92.1, 93.7)*
		M: 187 (9.4)		31 (1.6)	4.8 (1.8, 7.9)†	98.8 (98.3, 99.3)†	0.036 (NR)*	29.0 (13.1, 45.0)†	91.0 (89.7, 92.2)†
		F: 106 (5.3)		119 (6.0)	9.4 (3.9, 15.0)†	94.2 (93.2, 95.3)†	0.036 (NR)*	8.4 (3.42, 13.4)†	94.9 (93.9, 95.9)†

	IWGS, 2011	T: 806 (20.2)	≥4	150 (3.8)	5.8 (4.2, 7.5)*	96.8 (96.2, 97.4)*	0.026 (NR)*	31.3 (23.9, 38.8)*	80.3 (79.0, 81.5)*
		M: 442 (22.1)		31 (1.6)	3.8 (2.1, 5.6)†	99.1 (98.6, 99.6)†	0.029 (NR)*	54.8 (37.3, 72.4)†	78.4 (76.6, 80.2)†
		F: 364 (18.2)		119 (6.0)	8.2 (5.4, 11.1)†	94.6 (93.5, 95.7)†	0.028 (NR)*	25.2 (17.4, 33.0)†	82.2 (80.5, 84.0)†
Yang, 2018[64]	EWGSOP, 2010	T: 45 (11.7)	≥4	47 (12.2)	20.0 (0.6, 34.6)	95.6 (92.8, 97.5)	0.156 (NR)*	NR	NR
		M: 17 (10.6)		15 (9.4)	5.9 (0.1, 28.7)	93.7 (88.4, 97.1)	0.00 (NR)*		
		F: 28 (12.5)		32 (14.3)	28.6 (13.2, 48.7)	94.9 (90.8, 97.5)	0.235 (NR)*		
	AWGS, 2014	T: 61 (15.9)	≥4	47 (12.2)	29.5 (18.5, 42.6)	98.1 (96.0, 99.3)	0.276 (NR)*	NR	NR
		M: 19 (11.9)		15 (9.4)	15.8 (3.4, 39.5)	97.8 (93.9, 99.6)	0.136 (NR)*		
		F: 42 (18.8)		32 (14.3)	35.7 (21.6, 52.0)	98.3 (95.3, 99.7)	0.340 (NR)*		
	FNIH**, 2014	T: 59 (15.4)	≥4	47 (12.2)	30.5 (19.2, 43.9)	98.2 (96.0, 99.3)	0.287 (NR)*	NR	NR
		M: 23 (14.4)		15 (9.4)	21.7 (7.5, 43.7)	99.2 (96.0, 100)	0.209 (NR)*		
		F: 36 (16.1)		32 (14.3)	36.1 (20.8, 53.8)	97.3 (93.9, 99.1)	0.334 (NR)*		
	IWGS, 2011	T: 96 (25.0)	≥4	47 (12.2)	19.8 (12.4, 29.2)	98.2 (96.0, 99.4)	0.180 (NR)*	NR	NR
		M: 39 (24.4)		15 (9.4)	10.3 (2.9, 14.2)	98.3 (94.2, 99.8)	0.086 (NR)*		
		F: 57 (25.4)		32 (14.3)	26.3 (15.5, 39.7)	98.2 (94.8, 99.6)	0.245 (NR)*		
Yang, 2019[65]	EWGSOP, 2010	90 (32.5)	≥4	90 (32.5)	17.8 (10.5, 27.3)	96.8 (93.1, 98.8)	0.146 (NR)*	NR	NR
	AWGS, 2014	95 (34.3)	≥4	90 (32.5)	20.0 (12.5, 29.5)	98.4 (95.3, 99.7)	0.184 (NR)*	NR	NR
	FNIH**, 2014	87 (31.4)	≥4	90 (32.5)	21.8 (13.7, 32.0)	98.4 (95.5, 99.7)	0.202 (NR)*	NR	NR
	IWGS, 2011	106 (38.3)	≥4	90 (32.5)	17.0 (10.4, 25.5)	97.6 (94.1, 99.4)	0.146 (NR)*	NR	NR
SARC-CalF (0-18 points)									
Bahat, 2018[37]	EWGSOP, 2010	8 (3.86)	≥11‖	5 (2.42)	25.0 (0.00, 55.0)	98.0 (96.0, 99.0)	0.230 (NR)*	NR	NR
	FNIH**, 2014	19 (9.18)	≥11‖	5 (2.42)	10.5 (0.00, 24.0)	98.4 (96.0, 99.9)	0.089 (NR)*	NR	NR
	IWGS, 2011	4 (1.93)	≥11‖	5 (2.42)	50.0 (1.00, 99.0)	98.5 (96.8, 99.9)	0.485 (NR)*	NR	NR
	SCWD, 2011	5 (2.42)	≥11‖	5 (2.42)	40.0 (0.00, 82.0)	98.5 (96.8, 99.9)	0.385 (NR)*	NR	NR
	EWGSOP, 2010	8 (3.86)	≥11***	21 (10.1)	25.0 (0.00, 55.0)	90.0 (86.0, 94.5)	0.150 (NR)*	NR	NR
	FNIH**, 2014	19 (9.18)	≥11***	21 (10.1)	15.7 (0.00, 32.0)	90.4 (86.0, 94.6)	0.061 (NR)*	NR	NR
	IWGS, 2011	4 (1.93)	≥11***	21 (10.1)	50.0 (1.00, 99.0)	90.6 (86.6, 94.6)	0.406 (NR)*	NR	NR
	SCWD, 2011	5 (2.42)	≥11***	21 (10.1)	40.0 (0.00, 82.0)	90.5 (86.5, 94.6)	0.305 (NR)*	NR	NR
Barbosa-Silva, 2016[39]	EWGSOP, 2010	15 (8.38)	≥11†††	38 (21.2)	66.7 (38.4, 88.2)	82.9 (76.3, 88.4)	0.496 (NR)*	26.3 (13.4, 43.1)	96.5 (91.9, 98.8)
Fu, 2020[43]	EWGSOP2§, 2018	156 (50.5)	≥11†††	166 (53.7)	55.1 (47.0, 63.1)	76.4 (68.9, 82.9)	0.315 (NR)*	63.2 (53.6, 72.0)	69.8 (62.8, 76.1)
	AWGS, 2014	81 (26.2)	≥11†††	166 (53.7)	66.6 (55.3, 76.8)	70.1 (63.8, 76.9)	0.367 (NR)*	44.1 (35.1, 53.4)	85.5 (79.6, 90.3)
Mienche, 2019[58]	AWGS, 2014	19 (15.8)	≥4‡‡‡	6 (5.00)	21.1 (6.05, 45.57)	98.0 (93.0, 99.8)	0.191 (NR)*	66.7 (28.0, 91.0)	86.8 (84.0, 89.0)
Urzi, 2017[62]	EWGSOP, 2010	T: 31 (38.7)	≥11†††	29 (36.2)	77.4 (58.9, 90.4)	89.8 (77.7, 96.6)	0.672 (NR)*	82.8 (67.2, 91.8)	86.3 (76.5, 92.4)
		M: 8 (33.3)		8 (33.3)	87.5 (47.3, 99.7)	93.7 (69.8, 99.8)	0.812 (NR)*	87.5 (50.7, 97.9)	93.8 (70.5, 98.9)
		F: 23 (41.1)		21 (37.5)	73.9 (51.6, 89.8)	87.9 (71.8, 96.6)	0.618 (NR)*	80.9 (62.2, 91.7)	82.9 (70.6, 90.7)
Yang, 2018[64]	EWGSOP, 2010	T: 45 (11.7)	≥11†††	99 (25.8)	48.9 (33.7, 64.2)	90.6 (86.9, 93.5)	0.395 (NR)*	NR	NR
		M: 17 (10.6)		29 (18.1)	29.4 (10.3, 56.0)	93.7 (88.4, 97.1)	0.231 (NR)*		
		F: 28 (12.5)		70 (31.3)	46.4 (27.5, 66.1)	91.8 (87.1, 95.3)	0.382 (NR)*		
	AWGS, 2014	T: 61 (15.9)	≥11†††	99 (25.8)	60.7 (47.3, 72.9)	94.7 (91.7, 96.9)	0.554 (NR)*	NR	NR
		M: 19 (11.9)		29 (18.1)	47.4 (24.4, 71.1)	96.5 (91.9, 98.8)	0.439 (NR)*		
		F: 42 (18.8)		70 (31.3)	57.1 (41.0, 72.3)	97.3 (93.7, 99.1)	0.544 (NR)*		
	FNIH**, 2014	T: 59 (15.4)	≥11†††	99 (25.8)	55.9 (42.4, 68.8)	93.5 (90.5, 96.0)	0.494 (NR)*	NR	NR
		M: 23 (14.4)		29 (18.1)	43.9 (23.2, 65.5)	97.1 (92.7, 99.5)	0.410 (NR)*		
		F: 36 (16.1)		70 (31.3)	50.0 (32.9, 67.1)	94.2 (89.8, 97.0)	0.442 (NR)*		
	IWGS, 2011	T: 96 (25.0)	≥11†††	99 (25.8)	42.7 (32.7, 53.2)	95.5 (92.4, 97.6)	0.382 (NR)*	NR	NR
		M: 39 (24.4)		29 (18.1)	30.8 (17.0, 47.6)	98.4 (94.2, 99.8)	0.292 (NR)*		
		F: 57 (25.4)		70 (31.3)	50.9 (37.3, 64.4)	93.4 (88.5, 96.7)	0.443 (NR)*		
Yang, 2019[65]	EWGSOP, 2010	90 (32.5)	≥11†††	131 (47.3)	58.9 (48.0, 69.2)	84.5 (78.5, 89.4)	0.434 (NR)*	NR	NR
	AWGS, 2014	95 (34.3)	≥11†††	131 (47.3)	59.0 (48.4, 68.9)	85.7 (79.8, 90.5)	0.447 (NR)*	NR	NR
	FNIH**, 2014	87 (31.4)	≥11†††	131 (47.3)	64.4 (53.4, 74.4)	86.3 (80.6, 90.9)	0.507 (NR)*	NR	NR
	IWGS, 2011	106 (38.3)	≥11†††	131 (47.3)	55.7 (45.7, 65.3)	86.5 (80.5, 91.3)	0.422 (NR)*	NR	NR

(*continued on next page*)

Table 3 (*continued*)

First Author, Year	Sarcopenia Definition, Year	Sarcopenia Prevalence, n (%)	SARC-F Cutoff	SARC-F Prevalence, n (%)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Youden Index (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)
SARC-F+TC (0-10 points and TC below cutoff point)									
Mienche, 2019[58]	AWGS, 2014	19 (15.8)	≥4§§§	5 (4.17)	21.1 (6, 46)	99.0 (95, 100)	0.201 (NR)*	80.0 (32.0, 97.0)	87.0 (84.0, 89.0)
SARC-F+CC+TC (0-10 points and CC and TC below cutoff point)									
Mienche, 2019[58]	AWGS, 2014	19 (15.8)	≥4‡‡‡,§§§	4 (3.33)	15.8 (3, 40)	99.0 (95, 100)	0.148 (NR)*	75.0 (25.0, 96.0)	86.2 (84.0, 88.0)
SARC-F+E (0-20 points)									
Kurita, 2019[53]	EWGSOP, 2010	19 (1.98)	≥9	337 (35.1)*	84.2 (60.4, 96.6)	66.2 (63.0, 69.2)	0.504 (NR)*	4.75 (2.48, 7.02)*	99.5 (99.0, 100)*
	AWGS, 2014	36 (3.75)	≥9	334 (34.8)*	63.9 (46.2, 79.2)	66.3 (63.2, 69.4)	0.302 (NR)*	6.89 (4.17, 9.60)*	97.9 (96.8, 99.0)*
SARC-F+EBM (0-30 points)									
Kurita, 2019[53]	EWGSOP, 2010	19 (1.98)	≥12	312 (32.5)	84.2 (60.4, 96.6)	68.8 (65.8, 71.8)	0.530 (NR)*	5.13 (2.68, 7.58)*	99.5 (99.0, 100)*
	AWGS, 2014	36 (3.75)	≥12	309 (32.2)	77.8 (60.8, 89.9)	69.6 (66.5, 72.5)	0.474 (NR)*	9.06 (5.86, 12.3)*	98.8 (97.9, 99.6)*

CFNIH, Chinese FNIH; cont., continued; F, female; JSH, Japan Society of Hepatology; M, male; NR, not reported; T, total.
*Point estimate and 95% CI/†95% CI calculated based on reported data. ‡Probable, §confirmed, and ‖severe sarcopenia according to EWGSOP2. Sarcopenia according to FNIH on **low lean mass, weakness and slowness; ††low lean mass and weakness; ‡‡low lean mass and slowness; and §§low lean mass and weakness. CC cutoff points: ***≤33 cm; †††M: ≤34 cm, F: ≤33 cm; ‡‡‡M: <34 cm, F: <29 cm. TC cutoff points: §§§M: <49 cm, F: <44 cm; ‖‖‖≤31 cm.

applied: EWGSOP[2] (n = 17), EWGSOP2[1] (n = 8), Asian Working Group for Sarcopenia[66] (AWGS; n = 13), Japanese AWGS[67] (n = 1), Foundation for the National Institutes of Health[68] (FNIH; n = 8), Chinese FNIH[69] (n = 1), International Working Group on Sarcopenia[70] (IWGS; n = 9), Society on Sarcopenia, Cachexia and Wasting Disorders[71] (SCWD; n = 2), Baumgartner[72] (n = 1), Delmonico[73] (n = 1), and Japan Society of Hepatology[74] (n = 1) (Supplementary Table 4). For concurrent validity measures that were not pooled, the SARC-F had low (0.00) to high (0.604) Youden Index, small (0.211) to large (18.9) LR+, small (0.988) to moderate (0.289) LR−, low (0.594) to high (27.1) DOR, and low (11.9%) to high (98.2%) accuracy. The SARC-CalF had low (0.191) to high (0.812) Youden Index, small (1.6) to large (18.6) LR+, small (0.9) to moderate (0.133) LR−, low (1.77) to high (104) DOR, and moderate (66.2%) to high (91.7%) accuracy.

Pooled concurrent validity measures of the SARC-F (cutoff point: 4 points)

Figure 1 shows the summary of the meta-analyses of the sensitivity, specificity, PPV, and NPV of the SARC-F, stratified by sarcopenia definition (exact values are reported in Supplementary Table 5). Sensitivity was low using the EWGSOP (34.0%, 95% CI 20.9%, 47.1%), AWGS (28.9%, 95% CI 19.7%, 38.1%), FNIH (30.6%, 95% CI 22.3%, 39.0%) and IWGS (33.7%, 95% CI 18.9%, 48.5%), and moderate using the EWGSOP2 (53.3%, 95% CI 18.9%, 87.8%) and SCWD (55.3%, 95% CI 49.8%, 60.9%). Heterogeneity was high (I^2 = 75.5%-98.9%), except for the SCWD (I^2 = 0%). Specificity was high using the EWGSOP (82.9%, 95% CI 79.0%, 86.8%), AWGS (86.8%, 95% CI 83.1%, 90.5%), IWGS (88.9%, 95% CI 85.2%, 92.6%), SCWD (84.0%, 95% CI 80.5%, 87.5%), and FNIH (85.8%, 95% CI 80.4%, 91.3%), and moderate using the EWGSOP2 (68.9%, 95% CI 54.4%, 83.4%). Heterogeneity was high (I^2 = 95.6%-98.0%), except for the SCWD (I^2 = 15.2%).

PPV was low using the EWGSOP (28.2%, 95% CI 18.7%, 37.7%), EWGSOP2 (23.2%, 95% CI 10.5%, 35.9%), AWGS (24.5%, 95% CI 16.6%, 32.4%), FNIH (15.1%, 95% CI 7.10%, 23.2%), IWGS (38.2%, 95% CI 24.5%, 52.0%), and SCWD (12.4%, 95% CI −1.40%, 26.2%). High heterogeneity was reported for all definitions (I^2 = 91.2%-96.8%). NPV was high using the EWGSOP (89.0%, 95% CI 86.2%, 91.8%), EWGSOP2 (88.5%, 95% CI 82.2%, 94.7%), AWGS (89.8%, 95% CI 87.5%, 92.1%), FNIH (94.7%, 95% CI 91.8%, 97.6%), IWGS (83.1%, 95% CI 75.7%, 90.4%), and SCWD (97.5%, 95% CI 96.1%, 98.9%). Heterogeneity was high (I^2 = 91.0%-98.3%), except for the SCWD (I^2 = 0%). AUC was moderate using the EWGSOP (0.674, 95% CI 0.534, 0.815), EWGSOP2 (0.644, 95% CI 0.566, 0.722), AWGS (0.694, 95% CI 0.592, 0.796), FNIH (0.772, 95% CI 0.713, 0.830), and IWGS (0.737, 95% CI 0.657, 0.817). Heterogeneity was high (I^2 = 75.2%-94.9%), except for the EWGSOP2 (I^2 = 24.1%).

There was statistically significant heterogeneity (Qb $P < .05$) for specificity and NPV between definitions. There were no differences in sensitivity, PPV, and AUC between definitions.

Subgroup analyses for populations showed statistically significant heterogeneity (Qb $P < .05$) for sensitivity and specificity between populations using the EWGSOP and AWGS, and for specificity and PPV using the IWGS. Low sensitivity was observed for community-dwelling older adults (EWGSOP, EWGSOP2, AWGS, FNIH, and IWGS), geriatric outpatients (EWGSOP and AWGS), nursing home residents (EWGSOP, AWGS, FNIH, and IWGS), and long-term care residents (FNIH); moderate sensitivity in geriatric inpatients (EWGSOP2, AWGS, and IWGS) and geriatric outpatients (EWGSOP2); and high sensitivity in geriatric inpatients (EWGSOP). Low specificity was observed in geriatric inpatients (EWGSOP and IWGS) and geriatric outpatients (EWGSOP2) and moderate to high specificity in community-dwelling older adults (EWGSOP, EWGSOP2, AWGS, FNIH, IWGS, and SCWD), geriatric inpatients (EWGSOP2 and AWGS), geriatric outpatients (EWGSOP and AWGS), and nursing home residents (EWGSOP and IWGS). Low PPV was observed in community-dwelling older adults (EWGSOP, EWGSOP2, AWGS, FNIH, IWGS, and SCWD), geriatric

inpatients (EWGSOP, EWGSOP2, and AWGS), geriatric outpatients (EWGSOP, EWGSOP2, and AWGS), nursing home residents (EWGSOP2), and long-term care residents (FNIH) and moderate PPV in geriatric inpatients (IWGS). Moderate NPV was observed in geriatric inpatients (IWGS) and high NPV in community-dwelling older adults (EWGSOP, EWGSOP2, AWGS, FNIH, IWGS, and SCWD), geriatric inpatients (EWGSOP, EWGSOP2, and AWGS), geriatric outpatients (EWGSOP, EWGSOP2, and AWGS), nursing home residents (EWGSOP2), and long-term care residents (FNIH). Moderate AUC was observed in community-dwelling older adults (EWGSOP, AWGS, FNIH, and IWGS), geriatric inpatients (EWGSOP2 and AWGS), and nursing home residents (EWGSOP, AWGS, and FNIH) and low AUC in geriatric outpatients (EWGSOP, EWGSOP2, and AWGS).

Subgroup analyses for sexes (Supplementary Table 6) showed that the SARC-F had lower specificity (EWGSOP, AWGS, and IWGS), lower PPV (EWGSOP, AWGS, FNIH, and IWGS) and higher NPV (EWGSOP and FNIH) in females compared to males (statistically significant heterogeneity Qb $P < .05$). There were no differences in sensitivity and AUC between both sexes for the SARC-F. Subgroup analyses for risk of bias showed no statistically significant heterogeneity between high− and low−risk of bias articles, except for sensitivity using the FNIH (Qb $P = .048$) and specificity using the AWGS (Qb $P = .008$) (Supplementary Table 7).

Concurrent validity measures of the SARC-F (cutoff point: 1 and 6 points)

Subgroup analyses by different cutoff points were not performed because of the low number of studies. Twenty-eight of the 29 included articles reported on the proposed cutoff points for the SARC-F (≥4 points), of which 2 articles applied multiple cutoff points (Table 3).[42,53] Lower cutoff points resulted in higher sensitivity and lower specificity values compared to the cutoff of ≥4 points for all sarcopenia definitions applied.[42,53] The lowest applied cutoff point (≥1 point[42] and ≥2 points[53]) resulted in high sensitivity and low specificity. Higher cutoff points resulted in lower sensitivity and higher specificity compared with the cutoff of ≥4 points,[53] with low sensitivity and high specificity.

Pooled concurrent validity measures of the SARC-CalF

Figure 2 shows the summary of the meta-analyses of the sensitivity, specificity, PPV, and NPV of the SARC-CalF, stratified by sarcopenia definition (exact values are reported in Supplementary Table 5). Sensitivity was moderate using the EWGSOP (57.2%, 95% CI 43.2%, 71.2%) and AWGS (53.8%, 95% CI 39.1%, 68.6%), and low using the FNIH (45.9%, 95% CI 19.0%, 72.8%) and IWGS (49.4%, 95% CI 39.2%, 59.6%). Heterogeneity was high (I^2 = 81.4%-92.1%), except for the EWGSOP and IWGS (I^2 = 38.1%-69.6%). Specificity was high using the EWGSOP (87.9%, 95% CI 84.8%, 91.1%), AWGS (87.7%, 95% CI 78.9%, 96.4%), FNIH (90.5%, 95% CI 86.5%, 94.6%), and IWGS (91.3%, 95% CI 86.1%, 96.4%). Heterogeneity was high (I^2 = 81.0%-95.8%), except for the EWGSOP and FNIH (I^2 = 46.4%-67.7%).

PPV was moderate using the EWGSOP (54.7%, 95% CI -0.70%, 110%) and AWGS (50.1%, 95% CI 30.6%, 69.7%). Heterogeneity was high for the EWGSOP (I^2 = 97.0%) and moderate for the AWGS (I^2 = 45.0%). NPV was high using the EWGSOP (92.1%, 95% CI 82.1%, 102%) and AWGS (86.6%, 95% CI 84.3%, 88.9%). Heterogeneity was high for the EWGSOP (I^2 = 81.2%) and low for the AWGS (I^2 = 0%). AUC was high using the EWGSOP (0.810, 95% CI 0.766, 0.854), AWGS (0.830, 95% CI 0.724, 0.937), FNIH (0.817, 95% CI 0.716, 0.918), and IWGS (0.827, 95% CI 0.801, 0.854). Heterogeneity was high (I^2 = 93.8%-95.3%), except for the EWGSOP and IWGS (I^2 = 0%-55.1%). There was no statistically significant heterogeneity in any concurrent validity measures between definitions.

Subgroup analyses for populations showed statistically significant heterogeneity (Qb $P < .05$) for sensitivity and specificity between populations using the AWGS. There were no differences between populations for the other definitions. Low sensitivity was observed in community-dwelling older adults (EWGSOP, FNIH, and IWGS) and geriatric outpatients (AWGS), and moderate sensitivity in community-dwelling older adults (AWGS) and nursing home residents (EWGSOP, AWGS, FNIH, and IWGS). Moderate specificity was observed in geriatric inpatients (AWGS) and high specificity was observed in community-dwelling older adults (EWGSOP, AWGS, FNIH, and IWGS), geriatric outpatients (AWGS), and nursing home residents (EWGSOP, AWGS, FNIH, and IWGS). Low to moderate PPV was observed in community-dwelling older adults (EWGSOP and AWGS), geriatric inpatients (AWGS), and geriatric outpatients (AWGS) and high PPV in nursing home residents (EWGSOP). High NPV was observed in community-dwelling older adults (EWGSOP and AWGS), geriatric inpatients (AWGS), and geriatric outpatients (AWGS). Moderate AUC was observed in community-dwelling older adults (EWGSOP and FNIH), geriatric inpatients (AWGS), and geriatric outpatients (AWGS) and high AUC in community-dwelling older adults (AWGS and IWGS) and nursing home residents (EWGSOP, AWGS, FNIH, and IWGS).

Subgroup analyses for sexes revealed no statistically significant differences using all definitions (Supplementary Table 6). Subgroup analyses for risk of bias showed no statistically significant heterogeneity between high− and low−risk of bias articles, except for specificity using the AWGS (Qb $P < .001$) (Supplementary Table 7).

Concurrent validity of other modified SARC-F versions

Other SARC-F modified versions were used in 2 articles,[53,58] in which additional measurements were added to the original SARC-F, that is, calf circumference (CC), thigh circumference (TC), age (E), and body mass index (BM): SARC-F+TC (n = 1), SARC-F+CC+TC (n = 1), SARC-F+E (n = 1), and SARC-F+EBM (n = 1). No meta-analyses were performed because of the low number of articles. SARC-F+TC (using the AWGS) had low sensitivity, high specificity, low Youden Index, high PPV, high NPV, large LR+, small LR−, high DOR, high accuracy, and high AUC. SARC-F+CC+TC (using the AWGS) had low sensitivity, high specificity, low Youden Index, moderate PPV, high NPV, large LR+, small LR−, high DOR, high accuracy, and high AUC. SARC-F+E (using the EWGSOP and AWGS) had moderate to high sensitivity, moderate specificity, low Youden Index, low PPV, high NPV, small LR+, small to moderate LR−, low to high DOR, moderate accuracy, and moderate AUC. SARC-F+EBM (using the EWGSOP and AWGS) had moderate to high sensitivity, moderate specificity, low Youden Index, low PPV, high NPV, small LR+, moderate LR−, moderate to high DOR, moderate accuracy, and high AUC (Table 3, Supplementary Table 3).

Publication Bias

Visual evaluation of funnel plots revealed asymmetry and the Egger regression test showed statistically significant bias for the sensitivity of the SARC-F with the EWGSOP, AWGS, and IWGS; specificity using the EWGSOP, AWGS, and IWGS; and PPV, NPV, and AUC using the AWGS (Supplementary Table 8). No publication bias was observed for the FNIH, and for any validity measures of the SARC-CalF.

Discussion

The SARC-F had good to excellent inter-rater reliability, moderate to good test-retest reliability, low to high internal consistency, low to moderate sensitivity, moderate to high specificity, low PPV, high NPV, and moderate to high AUC, in identifying sarcopenia, independent of the sarcopenia definition. The SARC-CalF had low to moderate sensitivity, high specificity, low to moderate PPV, high NPV, and high AUC in identifying sarcopenia, independent of the sarcopenia definition. Population differences were found in concurrent validity measures for

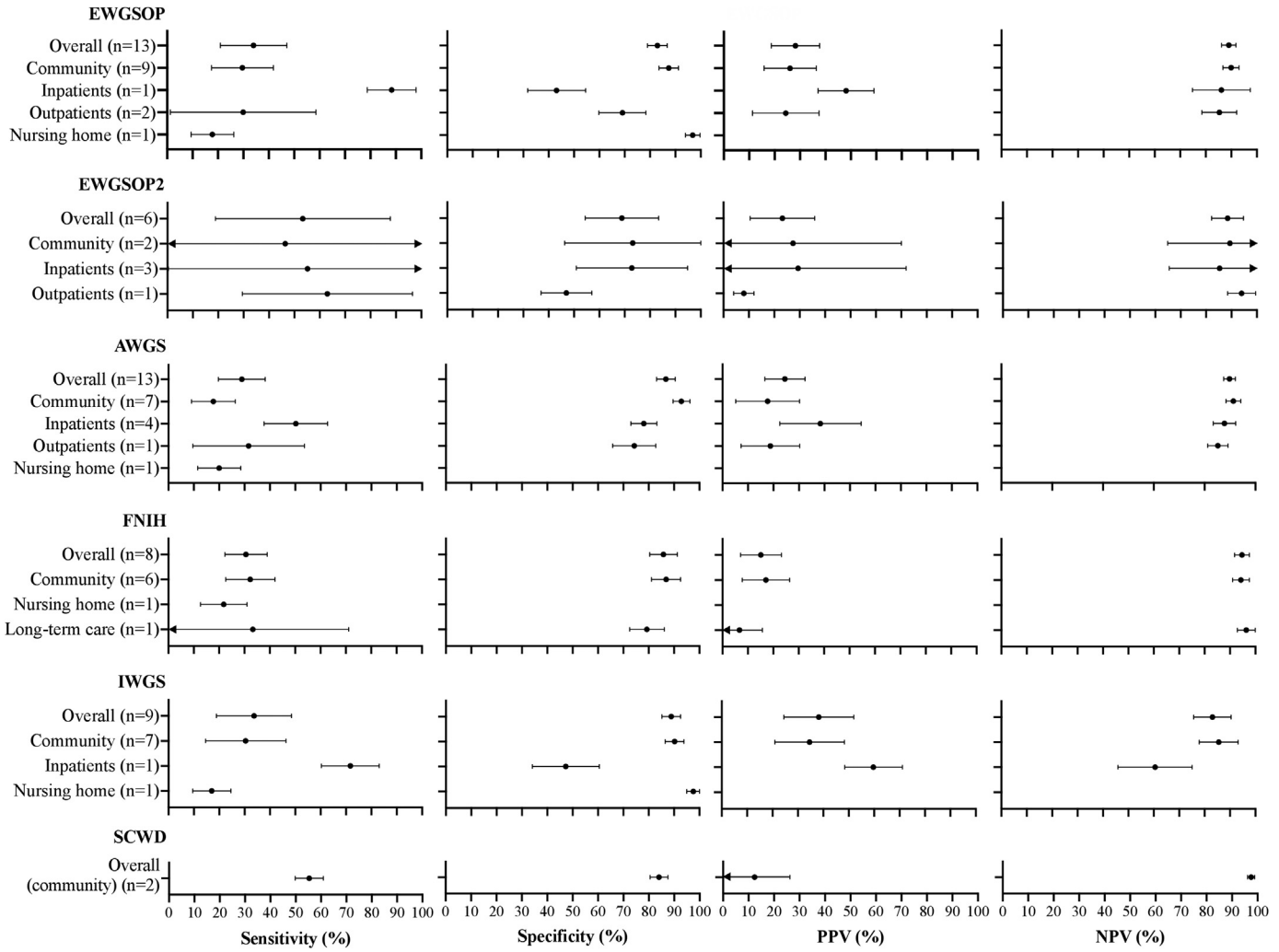


Fig. 1. Summary of meta-analyses of concurrent validity measures of the SARC-F to identify sarcopenia.

the SARC-F using the EWGSOP, AWGS, and IWGS, and for the SARC-CalF using the AWGS; however, a low number of studies were included in population subgroups. Sex differences were found in concurrent validity measures for the SARC-F using the EWGSOP, AWGS, FNIH, and IWGS, but no sex differences were found for SARC-CalF; however, a low number of studies were included in the sex subgroups too.

Sarcopenia Screening

The International Clinical Practice Guidelines for Sarcopenia (ICFSR, 2018) recommended screening for sarcopenia using gait speed or the SARC-F. However, these recommendations were conditional and had low certainty of evidence.[16] Despite this, the SARC-F was subsequently proposed as a case-finding tool in the EWGSOP2[1] and AWGS 2019[75] definitions. Although these definitions propose case-finding to identify those at risk of sarcopenia using the SARC-F and/or clinical judgement to identify symptoms or risk factors, the SARC-F was initially developed as a tool to "rapidly diagnose sarcopenia." In contrast, the SARC-F has been widely applied in the literature as a screening tool. There is a difference between case-finding and screening in that case-finding targets high-risk groups, whereas screening is not restricted to any population.[76] However, the purpose of both a case-finding tool and a screening tool is risk stratification, which was the proposed use of the SARC-F according to the authors who developed it, the EWGSOP2 and the AWGS 2019. Therefore, the labels "diagnostic tool," "case-finding tool," and "screening tool" are used interchangeably in the case of the SARC-F. The results of this review add to the body of evidence for sarcopenia screening using the SARC-F and the SARC-CalF when using multiple sarcopenia definitions across various populations.

Reliability of the SARC-F

This is the first review assessing the reliability of the SARC-F. The good to excellent inter-rater reliability and good test-retest reliability of the SARC-F signifies the high degree of agreement between assessors that ensures the results are representative over time. The SARC-F was found to have high internal consistency. However, the SARC-F also showed low internal consistency, indicating variability within and between the individual SARC-F components.

Concurrent Validity of the SARC-F

Based on the low to moderate sensitivity, moderate to high specificity, low PPV, and high NPV, the SARC-F was found to have a low ability to identify those with sarcopenia correctly and a high ability to identify those without sarcopenia correctly.[77,78] These results were

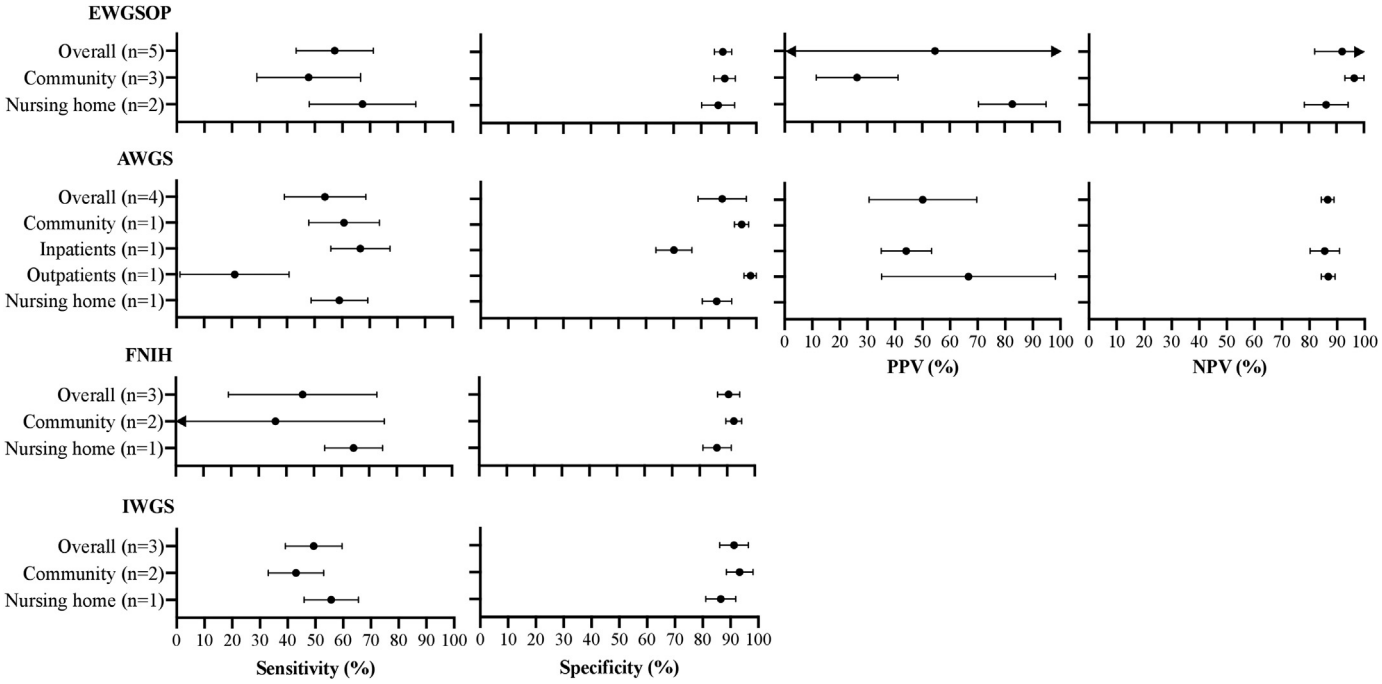


Fig. 2. Summary of meta-analyses of concurrent validity measures of the SARC-CalF to identify sarcopenia.

similar to a previous meta-analysis reporting low sensitivity and high specificity of the SARC-F.[17] Although the addition of CC measurements as a proxy measure of muscle mass in the SARC-CalF[39] resulted in higher sensitivity compared to the SARC-F, the addition of TC measurements in the SARC-F+TC and SARC-F+CC+TC[58] resulted in a lower sensitivity. Additionally, the inclusion of age (SARC-F+E) and age and low body mass index (SARC-F+EBM)[53] were reported to have higher sensitivity than the SARC-F. However, limited studies were available on using these modified SARC-F versions, thereby impairing the representation of these versions. It can be suggested that the low to moderate sensitivity of the SARC-F could be attributed to the questionnaire only assessing self-reported measures of muscle strength and physical function, and not muscle mass. It can also be proposed that the addition of determinants of sarcopenia, such as age and low body mass index, could increase the sensitivity of the SARC-F.

Concurrent validity can be influenced by the disease prevalence in a population,[79] which depends on population characteristics such as disease severity and comorbidity that could determine the risk of sarcopenia.[9] A large proportion of the included studies were of community-dwelling older adults, in which the SARC-F was found to have moderate overall concurrent validity (measured using AUC) in this population. These results echoed the overall meta-analysis across sarcopenia definitions. Although statistically significant differences between populations were found for concurrent validity measures of the SARC-F and the SARC-CalF, subgroup analyses on populations did not reveal that the SARC-F or the SARC-CalF performed better in one population over the other as the overall concurrent validity (measured using AUC) were similar across the included populations and definitions. The low representation of these populations may have affected the observed differences between populations. Therefore, further research is needed to determine if the SARC-F and its modified versions are best-performing in these populations, namely, geriatric inpatients, geriatric outpatients, and nursing home residents.

Sex differences across all concurrent validity measures of the SARC-F were only statistically significant for specificity, PPV, and NPV, whereas none were statistically significant for the SARC-CalF. Despite the lack of statistical significance across the board, higher sensitivity, lower specificity, lower PPV, and higher NPV of the SARC-F were observed for females compared with males with the SARC-F and the SARC-CalF (except for the higher specificity observed in females with the SARC-CalF using one of the four definitions). A low number of studies were included for the sex subgroups and would require further research to determine sex differences in diagnostic accuracy of the SARC-F. The sex differences in diagnostic accuracy of the present analysis suggests that just as diagnostic assessment for sarcopenia has sex-specific cutoff points across all definitions, the diagnostic accuracy of the SARC-F could be sex-specific, hence potentially requiring sex-specific cutoff points.

Clinical Implications

In a clinical setting, high sensitivity and high specificity are important so that interventions can be initiated for those at risk of or diagnosed with sarcopenia, and unnecessary diagnostic assessments can be prevented to decrease the economic burden on the health care system.[80] The pooled results from the SARC-F did not satisfy the AUC, Youden Index, or high sensitivity requirements of a valid screening tool for sarcopenia. Despite the low number of studies examining alternative cutoff points, none of these other cutoff points improved the sensitivity of the SARC-F beyond 61.1% (moderate) while maintaining at least a moderate level of specificity. The high AUC was achieved by the SARC-CalF; however, the high specificity yet low to moderate sensitivity of the SARC-CalF made it unsuitable as a screening tool for sarcopenia. As such, despite the moderate to high specificity of the SARC-F, its low to moderate sensitivity makes it nonoptimal to screen for sarcopenia using the SARC-F because of the high risk of missed sarcopenia diagnosis that could leave these at-risk individuals prone to progressing toward adverse health outcomes.[6–10]

This stresses the need for a screening tool that has a better AUC and Youden Index. Creating a new screening tool that includes determinants of sarcopenia could be considered.

Strengths and Limitations

This systematic review and meta-analysis reports on the reliability and concurrent validity of the SARC-F and modified SARC-F versions across 6 sarcopenia definitions and 5 different populations. This is the first review assessing the reliability of the SARC-F. However, not all articles were included in the meta-analysis either because of a limited number of articles, asymmetrical 95% CIs for reliability measures, or the lack of data required to calculate 95% CIs for concurrent validity measures.

Conclusions and Implications

The SARC-F had good reliability, low to moderate sensitivity, and moderate to high specificity in identifying sarcopenia, independent of the sarcopenia definition. The SARC-CalF had higher sensitivity compared with the SARC-F. Despite the good reliability of the SARC-F, its low to moderate sensitivity makes it nonoptimal to use the tool to screen for sarcopenia in older adults. Additionally, the SARC-F and the SARC-CalF were not shown to perform better in certain populations. As such, it is recommended to apply the diagnostic criteria for sarcopenia without screening.

Acknowledgments

The authors thank Patrick Condron (senior liaison librarian, Brownless Biomedical Library, Faculty of Medicine, Dentistry & Health Sciences, The University of Melbourne), who assisted with the development of the search strategy.

References

1. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: Revised European Consensus on Definition and Diagnosis. Age Ageing 2018;48:16–31.
2. Cruz-Jentoft AJ, Baeyens JP, Bauer JM, et al. Sarcopenia: European Consensus on Definition and Diagnosis: Report of the European Working Group on Sarcopenia in Older People. Age Ageing 2010;39:412–423.
3. Reijnierse EM, Trappenburg MC, Leter MJ, et al. The impact of different diagnostic criteria on the prevalence of sarcopenia in healthy elderly participants and geriatric outpatients. Gerontology 2015;61:491–496.
4. Cruz-Jentoft AJ, Landi F, Schneider SM, et al. Prevalence of and interventions for sarcopenia in ageing adults: A systematic review. Report of the International Sarcopenia Initiative (EWGSOP and IWGS). Age Ageing 2014;43:748–759.
5. Shafiee G, Keshtkar A, Soltani A, et al. Prevalence of sarcopenia in the world: A systematic review and meta-analysis of general population studies. J Diabetes Metab Disord 2017;16:21.
6. Janssen I, Heymsfield S, Ross R. Low relative skeletal muscle mass (sarcopenia) in older persons is associated with functional impairment and physical disability. J Am Geriatr Soc 2002;50:889–896.
7. Wang DXM, Yao J, Zirek Y, et al. Muscle mass, strength, and physical performance predicting activities of daily living: A meta-analysis. J Cachexia Sarcopenia Muscle 2020;11:3–25.
8. Yeung SSY, Reijnierse EM, Pham VK, et al. Sarcopenia and its association with falls and fractures in older adults: A systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2019;10:485–500.
9. Pacifico J, Geerlings MAJ, Reijnierse EM, et al. Prevalence of sarcopenia as a comorbid disease: A systematic review and meta-analysis. Exp Gerontol 2019;131:1–19.
10. Beaudart C, Zaaria M, Pasleau F, et al. Health outcomes of sarcopenia: A systematic review and meta-analysis. PLoS One 2017;12:e0169548.
11. Yeung SSY, Reijnierse EM, Trappenburg MC, et al. Current knowledge and practice of Australian and New Zealand health-care professionals in sarcopenia diagnosis and treatment: Time to move forward! Australas J Ageing 2019;15:15.
12. Reijnierse EM, de van der Schueren MAE, Trappenburg MC, et al. Lack of knowledge and availability of diagnostic equipment could hinder the diagnosis of sarcopenia and its management. PLoS One 2017;12:e0185837.
13. Van Ancum JM, Meskers CGM, Reijnierse EM, et al. Lack of knowledge contrasts the willingness to counteract sarcopenia among community-dwelling adults. J Aging Health; 2019. 898264319852840.
14. Sayer A, Syddall H, Martin H, et al. The developmental origins of sarcopenia. J Nutr Health Aging 2008;12:427–432.
15. Malmstrom TK, Morley JE. SARC-F: A simple questionnaire to rapidly diagnose sarcopenia. J Am Med Dir Assoc 2013;14:531–532.
16. Dent E, Morley JE, Cruz-Jentoft AJ, et al. International Clinical Practice Guidelines for Sarcopenia (ICFSR): Screening, diagnosis and management. J Nutr Health Aging 2018;22:1148–1161.
17. Ida S, Kaneko R, Murata K. SARC-F for screening of sarcopenia among older adults: A meta-analysis of screening test accuracy. J Am Med Dir Assoc 2018;19:685–689.
18. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009;6:e1000097.
19. Lo CK-L, Mertz D, Loeb M. Newcastle-Ottawa Scale: Comparing reviewers' to authors' assessments. BMC Med Res Methodol 2014;14:45.
20. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. Ottawa, Ontario, Canada: Ottawa Hospital Research Institute, University of Ottawa; 2001.
21. Jüni P, Witschi A, Bloch R, et al. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999;282:1054–1060.
22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–163.
23. Masson LF, McNeill G, Tomany JO, et al. Statistical approaches for assessing the relative validity of a food-frequency questionnaire: Use of correlation coefficients and the kappa statistic. Public Health Nutr 2003;6:313–321.
24. Nunnally J, Bernstein L. Psychometric Theory. 3rd ed. New York: McGraw-Hill Inc; 1994.
25. Power L, Mullally D, Gibney ER, et al. A review of the validity of malnutrition screening tools used in older adults in community and healthcare settings—A MaNuEL Study. Clin Nutr ESPEN 2018;24:1–13.
26. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J 2005;47:458–472.
27. Schisterman EF, Perkins N. Confidence intervals for the Youden Index and corresponding optimal cut-point. Commun Stat Simul Comput 2007;36:549–563.
28. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. JAMA 1994;271:703–707.
29. Glas AS, Lijmer JG, Prins MH, et al. The diagnostic odds ratio: A single indicator of test performance. J Clin Epidemiol 2003;56:1129–1135.
30. van Bokhorst-de van der Schueren MAE, Guaitoli PR, Jansma EP, de Vet HCW. A systematic review of malnutrition screening tools for the nursing home setting. J Am Med Dir Assoc 2014;15:171–184.
31. Herman C. What makes a screening exam "good"? Virtual Mentor 2006;8:34–37.
32. Reitsma JB, Glas AS, Rutjes AWS, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58:982–990.
33. Zwinderman AH, Bossuyt PM. We should not pool diagnostic likelihood ratios in systematic reviews. Stat Med 2008;27:687–697.
34. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–1558.
35. Higgins JPT, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. Br Med J 2003;327:557–560.
36. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. Br Med J 1997;315:629–634.
37. Bahat G, Oren MM, Yilmaz O, et al. Comparing SARC-F with SARC-CalF to screen sarcopenia in community living older adults. J Nutr Health Aging 2018;22:1034–1038.
38. Bahat G, Yilmaz O, Kilic C, et al. Performance of SARC-F in regard to sarcopenia definitions, muscle mass and functional measures. J Nutr Health Aging 2018;22:898–903.
39. Barbosa-Silva TG, Menezes AM, Bielemann RM, et al. Enhancing SARC-F: Improving sarcopenia screening in the clinical practice. J Am Med Dir Assoc 2016;17:1136–1141.
40. Beaudart C, Locquet M, Bornheim S, et al. French translation and validation of the sarcopenia screening tool SARC-F. Eur Geriatr Med 2018;9:29–37.
41. Borges TC, Gomes TLN, Pimentel GD. Sarcopenia as a predictor of nutritional status and comorbidities in hospitalized patients with cancer: A cross-sectional study. Nutrition 2019;73:110703.
42. Drey M, Ferrari U, Schraml M, et al. German version of SARC-F: Translation, adaption, and validation. J Am Med Dir Assoc 2020;21:747–751e741.
43. Fu X, Tian Z, Thapa S, et al. Comparing SARC-F with SARC-CalF for screening sarcopenia in advanced cancer patients. Clin Nutr 2020;22:22.
44. Ha YC, Won Won C, Kim M, et al. SARC-F as a useful tool for screening sarcopenia in elderly patients with hip fractures. J Nutr Health Aging 2020;24:78–82.
45. Ida S, Murata K, Nakadachi D, et al. Development of a Japanese version of the SARC-F for diabetic patients: An examination of reliability and validity. Aging Clin Exp Res 2017;29:935–942.
46. Ida S, Kojima Y, Hamaoka S, et al. Validity of Japanese version of SARC-F questionnaire in patients with chronic liver disease. J Gastroenterol Hepatol 2019;34:947–953.
47. Kemmler W, Sieber C, Freiberger E, et al. The SARC-F questionnaire: Diagnostic overlap with established sarcopenia definitions in older German men with sarcopenia. Gerontology 2017;63:411–416.

48. Kera T, Kawai H, Hirano H, et al. SARC-F: A validation study with community-dwelling older Japanese adults. Geriatr Gerontol Int 2019;19:1172–1178.
49. Kera T, Kawai H, Hirano H, et al. Limitations of SARC-F in the diagnosis of sarcopenia in community-dwelling older adults. Arch Gerontol Geriatr 2020;87:103959.
50. Kim S, Kim M, Won CW. Validation of the Korean version of the SARC-F questionnaire to assess sarcopenia: Korean Frailty and Aging Cohort Study. J Am Med Dir Assoc 2018;19:40–45.e41.
51. Kim M, Won CW. Prevalence of sarcopenia in community-dwelling older adults using the definition of the European Working Group on Sarcopenia in Older People 2: Findings from the Korean Frailty and Aging Cohort Study. Age Ageing 2019;48:910–916.
52. Kotlarczyk MP, Perera S, Nace DA, et al. Identifying sarcopenia in female long-term care residents: A comparison of current guidelines. J Am Geriatr Soc 2018;66:316–320.
53. Kurita N, Wakita T, Kamitani T, et al. SARC-F validation and SARC-F+EBM derivation in musculoskeletal disease: The SPSS-OK Study. J Nutr Health Aging 2019;23:732–738.
54. Li M, Kong Y, Chen H, et al. Accuracy and prognostic ability of the SARC-F questionnaire and Ishii's score in the screening of sarcopenia in geriatric inpatients. Braz J Med Biol Res 2019;52:e8204.
55. Lima DP, de Almeida SB, Bonfadini JC, et al. Clinical correlates of sarcopenia and falls in Parkinson's disease. PLoS One 2020;15:e0227238.
56. Lim WS, Tay L, Yeo A, et al. Modulating effect of contextual factors on factor structure and reliability of SARC-F. J Am Med Dir Assoc 2018;19:551–553.
57. Malmstrom TK, Miller DK, Simonsick EM, et al. SARC-F: A symptom score to predict persons with sarcopenia at risk for poor functional outcomes. J Cachexia Sarcopenia Muscle 2016;7:28–36.
58. Mienche M, Setiati S, Setyohadi B, et al. Diagnostic performance of calf circumference, thigh circumference, and SARC-F questionnaire to identify sarcopenia in elderly compared to Asian Working Group for Sarcopenia's Diagnostic Standard. Acta Med Indones 2019;51:117–127.
59. Parra-Rodriguez L, Szlejf C, Garcia-Gonzalez AI, et al. Cross-cultural adaptation and validation of the Spanish-language version of the SARC-F to assess sarcopenia in Mexican community-dwelling older adults. J Am Med Dir Assoc 2016;17:1142–1146.
60. Rolland Y, Dupuy C, Abellan Van Kan G, et al. Sarcopenia screened by the SARC-F questionnaire and physical performances of elderly women: A cross-sectional study. J Am Med Dir Assoc 2017;18:848–852.
61. Sanchez-Rodriguez D, Marco E, Davalos-Yerovi V, et al. Translation and validation of the Spanish version of the SARC-F questionnaire to assess sarcopenia in older people. J Nutr Health Aging 2019;23:518–524.
62. Urzi F, Simunic B, Buzan E. Basis for sarcopenia screening with the SARC-CalF in nursing homes. J Am Med Dir Assoc 2017;18:991.e5–991.e10.
63. Woo J, Leung J, Morley JE. Validating the SARC-F: A suitable community screening tool for sarcopenia? J Am Med Dir Assoc 2014;15:630–634.
64. Yang M, Hu X, Xie L, et al. Screening sarcopenia in community-dwelling older adults: SARC-F vs SARC-F combined with calf circumference (SARC-CalF). J Am Med Dir Assoc 2018;19:277e271–277e278.
65. Yang M, Lu J, Jiang J, et al. Comparison of four sarcopenia screening tools in nursing home residents. Aging Clin Exp Res 2019;31:1481–1489.
66. Chen LK, Liu LK, Woo J, et al. Sarcopenia in Asia: Consensus report of the Asian Working Group for Sarcopenia. J Am Med Dir Assoc 2014;15:95–101.
67. Tanimoto Y, Watanabe M, Sun W, et al. Association between sarcopenia and higher-level functional capacity in daily living in community-dwelling elderly subjects in Japan. Arch Gerontol Geriatr 2012;55:e9–e13.
68. Studenski SA, Peters KW, Alley DE, et al. The FNIH sarcopenia project: Rationale, study description, conference recommendations, and final estimates. J Gerontol A Biol Sci Med Sci 2014;69:547–558.
69. Woo J, Leung J. Anthropometric cut points for definition of sarcopenia based on incident mobility and physical limitation in older Chinese people. J Gerontol A Biol Sci Med Sci 2016;71:935–940.
70. Fielding RA, Vellas B, Evans WJ, et al. Sarcopenia: An undiagnosed condition in older adults. Current consensus definition: Prevalence, etiology, and consequences. International Working Group on Sarcopenia. J Am Med Dir Assoc 2011;12:249–256.
71. Morley JE, Abbatecola AM, Argiles JM, et al. Sarcopenia with limited mobility: An international consensus. J Am Med Dir Assoc 2011;12:403–409.
72. Baumgartner RN, Koehler KM, Gallagher D, et al. Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol 1998;147:755–763.
73. Delmonico MJ, Harris TB, Lee JS, et al. Alternative definitions of sarcopenia, lower extremity performance, and functional impairment with aging in older men and women. J Am Geriatr Soc 2007;55:769–774.
74. Nishikawa H, Shiraki M, Hiramatsu A, et al. Japan Society of Hepatology guidelines for sarcopenia in liver disease (1st edition): Recommendation from the Working Group for Creation of Sarcopenia Assessment Criteria. Hepatol Res 2016;46:951–963.
75. Chen LK, Woo J, Assantachai P, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020;21:300–307e302.
76. Diagnosis of dementia; 2012. Available at: https://www.who.int/mental_health/mhgap/evidence/resource/dementia_q6.pdf?ua=1. Accessed March 18, 2021.
77. Parikh R, Mathai A, Parikh S, et al. Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol 2009;56:45–50.
78. Parikh R, Parikh S, Arun E, et al. Likelihood ratios: Clinical application in day-to-day practice. Indian J Ophthalmol 2009;57:217–221.
79. Leeflang MMG, Rutjes AWS, Reitsma JB, et al. Variation of a test's sensitivity and specificity with disease prevalence. CMAJ 2013;185:E537–E544.
80. Bruyere O, Beaudart C, Ethgen O, et al. The health economics burden of sarcopenia: A systematic review. Maturitas 2019;119:61–69.

Supplementary Material 1. Modified Newcastle-Ottawa Scale

Note: An article can be awarded a maximum of 1 star for each numbered item within the Selection and Comparability categories. There are no limits of the number of stars an article can be awarded for each numbered item within the Outcome category. (Lowest maximum score: 5; Highest maximum score: 8)

Selection

1) *Representativeness of the exposed cohort*
   a) Truly representative of the average subjects with sarcopenia ✳
   b) Somewhat representative of the average subjects with sarcopenia ✳
   c) No description of the derivation of the cohort
2) *Selection of the non-exposed cohort: subjects without sarcopenia drawn from the same population as the exposed cohort*
   a) Yes, drawn from the same population as the exposed cohort ✳
   b) No, drawn from a different population
   c) No description of the derivation of the non-exposed cohort
3) *Ascertainment of exposure: screening for sarcopenia using the SARC-F*
   a) Screening for sarcopenia using SARC-F or its modified versions ✳
   b) Screening for sarcopenia using methods other than those listed above
   c) No description or unclear

Comparability

1) *Comparability of cohorts on the basis of the screening for sarcopenia using the SARC-F*
   a) No selection of individuals to be screened for sarcopenia using SARC-F or its modified versions ✳
   b) Selection of individuals to be screened for sarcopenia using SARC-F or its modified versions.

Outcome

For assessment of articles reporting solely reliability measures, use 1) only. For assessment of articles reporting solely concurrent validity measures, use 2) only. For assessment of articles reporting on both reliability and concurrent validity measures, use both 1) and 2).

1) *Ascertainment of outcome: how the reliability of the SARC-F was assessed*
   a) Assessment of inter-rater reliability was done *between* assessors and *within* the same individual using the same SARC-F version; ✳
   b) Assessment of test-retest/intra-rater reliability was done within a *single* assessor and *within* the same individual using the same SARC-F version ✳
   c) Assessment of internal consistency reliability was done by measuring *all* SARC-F questions (sub scores) using the same SARC-F version ✳
   d) Assessment of inter-rater reliability, test-retest/intra-rater reliability, and internal consistency reliability using methods other than those listed above
2) *Ascertainment of outcome: how the diagnosis of sarcopenia was made*
   a) Sarcopenia diagnosed using the diagnostic criteria of 1 or more of the following definitions:
   EWGSOP (2010), EWGSOP2 (2018), AWGS (2014), AWGS (2019), FNIH (2014), IWGS (2011), or SDOC (2019), AND diagnosed with muscle mass using BIA, BIS, DXA, MRI, MRS, CT scan, air displacement plethysmography, hydrometry OR muscle strength (handgrip strength, chair stand test, knee extension strength test); physical performance (gait speed, SPPB, TUG, 400-metre walk) according to the definition ✳
   b) Sarcopenia diagnosed using diagnostic criteria of a definition listed above, but with methods other than those listed above
   c) Sarcopenia diagnosed using diagnostic criteria of a definition other than those listed above
   d) No description or unclear

✳ = 1 star

Supplementary Table 1
Search Strategy

Database: Ovid MEDLINE(R) epub ahead of print, in-process and other non-indexed citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE (2013 to April 6, 2020)

No.	Searches	Results
1	muscle weakness/or muscular atrophy/or sarcopenia/	22,347
2	(sarcopenia or "muscular atrophy" or "muscle weakness" or "muscular weakness").mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	46,343
3	1 or 2	46,343
4	(SARC-F or SARC-F-5 or SARC-F-3 or SARC-CalF or SARC-CalF-31 or SARC-CalF-33).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	85
5	screening.mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	593,181
6	((diagnosis or diagnostic*) and (prognosis or survey or surveys or questionnaire* or qualitative or "self report" or schedule or interview* or tool or tools or test* or form or forms)).mp. [mp=title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	1,626,265
7	3 and (4 or 5 or 6)	6722
8	(case report* or editorial or letter or review*).pt.	5,947,988
9	7 not 8	4116
10	limit 9 to (english language and yr="2013 - 2020")	2209

Database: Embase Classic + Embase, January 1, 2013, to April 6, 2020

No.	Searches	Results
1	muscular atrophy/or sarcopenia/	30,288
2	(sarcopenia or "muscular atrophy").mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	28,686
3	1 or 2	45,688
4	(SARC-F or SARC-F-5 or SARC-F-3 or SARC-CalF or SARC-CalF-31 or SARC-CalF-33).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	157
5	screening.mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	1,060,034
6	((diagnosis or diagnostic*) and (prognosis or survey or surveys or questionnaire* or qualitative or "self report" or schedule or interview* or tool or tools or test* or form or forms)).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword, floating subheading word, candidate term word]	2,349,226
7	3 and (4 or 5 or 6)	6656
8	(case report* or editorial or letter or review*).pt.	4,358,359
9	7 not 8	5504
10	limit 9 to (english language and yr="2013 - 2020")	3038
11	limit 10 to conference abstract status	1153
12	10 not 11	1885

Database: EBM Reviews—Cochrane Central Register of Controlled Trials February 2020

No.	Searches	Results
1	muscle weakness/or muscular atrophy/or sarcopenia/	1188
2	(sarcopenia or "muscular atrophy" or "muscle weakness" or "muscular weakness").mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	4392
3	1 or 2	4392
4	(SARC-F or SARC-F-5 or SARC-F-3 or SARC-CalF or SARC-CalF-31 or SARC-CalF-33).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	13
5	screening.mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	55,091
6	((diagnosis or diagnostic*) and (prognosis or survey or surveys or questionnaire* or qualitative or "self report" or schedule or interview* or tool or tools or test* or form or forms)).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]	79,637
7	3 and (4 or 5 or 6)	426
8	(case report* or editorial or letter or review*).pt.	12,066
9	7 not 8	425
10	limit 9 to (english language and yr="2013 - 2020")	156

Database: CINAHL (January 1, 2013, to April 6, 2020)

No.	Searches	Results
S1	“muscle weakness” or “muscular weakness” or “muscular atrophy” or sarcopenia	14,476
S2	SARC-F or SARC-F-5 or SARC-F-3 or SARC-CalF or SARC-CalF-31 or SARC-CalF-33	71
S3	“screening”	175,829
S4	(diagnosis or diagnostic*) and (prognosis or survey or surveys or questionnaire* or qualitative or "self report" or schedule or interview* or tool or tools or test* or form or forms)	432,861
S5	S2 or S3 or S4	557,613
S6	S1 and S5	2297
S7	S6 (Limiters - Published Date: 2013-2020; English Language)	1407

Supplementary Table 2
Results of the Risk of Bias Assessment of Included Articles

First Author, Year	Section of NOS						Raw Score/ Max Score	Adjusted Score*	Risk of Bias
	Selection			Com.	Outcome				
	Q1	Q2	Q3	Q1	Q1	Q2			
Bahat, 2018[37]	*	*	*	*	*	N/A	5/5	5.00	Low
Bahat, 2018[38]	*	*	*	*	N/A	**	6/6	5.00	Low
Barbosa-Silva, 2016[39]	*	*	*	*	*	N/A	5/5	5.00	Low
Beaudart, 2018[40]		*	*	*	*	**	6/7	4.29	High
Borges, 2019[41]	*	*	*	*		N/A	4/5	4.00	High
Drey, 2020[42]	*	*	*	*	*	***	8/8	5.00	Low
Fu, 2020[43]	*	*	*	*		N/A	4/5	4.00	High
Ha, 2020[44]	*	*	*	*		N/A	4/5	4.00	High
Ida, 2017[45]	*	*	*	*	*	*	6/6	5.00	Low
Ida, 2019[46]	*	*	*	*		N/A	4/5	4.00	High
Kemmler, 2017[47]	*	*	*	*	*	N/A	5/5	5.00	Low
Kera, 2019[48]	*	*	*	*	*	*	6/6	5.00	Low
Kera, 2020[49]	*	*	*	*	*	N/A	5/5	5.00	Low
Kim, 2018[50]	*	*	*	*	*	N/A	5/5	5.00	Low
Kim, 2019[51]	*	*	*	*	*	N/A	5/5	5.00	Low
Kotlarczyk, 2018[52]	*	*	*	*	*	N/A	5/5	5.00	Low
Kurita, 2019[53]		*	*	*	*	N/A	4/5	4.00	High
Li, 2019[54]	*	*	*	*	*	N/A	5/5	5.00	Low
Lima, 2020[55]	*	*	*	*		N/A	4/5	4.00	High
Lim, 2018[56]	*	*	*	*	N/A	*	5/5	5.00	Low
Malmstrom, 2016[57]	*	*	*	*	N/A	*	5/5	5.00	Low
Mienche, 2019[58]	*	*	*	*	*	N/A	5/5	5.00	Low
Parra-Rodriguez, 2016[59]		*	*	*	*	**	6/7	4.29	High
Rolland, 2017[60]	*	*	*	*	*	N/A	5/5	5.00	Low
Sanchez-Rodriguez, 2019[61]	*	*	*	*	*	***	8/8	5.00	Low
Urzi, 2017[62]	*	*	*	*	*	N/A	5/5	5.00	Low
Woo, 2014[63]		*	*	*	*	N/A	4/5	4.00	High
Yang, 2018[64]	*	*	*	*	*	N/A	5/5	5.00	High
Yang, 2019[65]	*	*	*	*	*	N/A	5/5	5.00	High

Com., comparability; N/A, not applicable; NOS, Newcastle-Ottawa Scale.
Selection: Q1, representativeness; Q2, selection; Q3, ascertainment. Comparability: Q1, comparability of cohorts. Outcome: Q1, ascertainment (reliability); Q2, ascertainment (sarcopenia).
*Calculated based on raw score.

Supplementary Table 3
Additional Concurrent Validity Measures of the SARC-F Versions to Identify Sarcopenia

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	LR+ (95% CI)	LR− (95% CI)	DOR (95% CI)	Accuracy, % (95% CI)	AUC (95% CI)
SARC-F (0-10 points)									
Bahat, 2018a[36]	EWGSOP, 2010	8 (3.86)	≥4	39 (18.8)	1.3 (0.39, 4.62)†	0.9 (0.61, 1.38)†	1.45 (NR)*	NR	0.522 (0.452, 0.592)
	FNIH**, 2014	19 (9.18)	≥4	39 (18.8)	1.8 (0.87, 3.74)†	0.8 (0.61, 1.13)†	2.17 (NR)*	NR	0.701 (0.634, 0.763)
	IWGS, 2011	4 (1.93)	≥4	39 (18.8)	2.7 (0.98, 7.63)†	0.6 (0.23, 1.63)†	4.50 (NR)*	NR	0.640 (0.571, 0.706)
	SCWD, 2011	5 (2.42)	≥4	39 (18.8)	2.1 (0.72, 6.64)†	0.7 (0.36, 1.51)†	2.98 (NR)*	NR	0.551 (0.481, 0.620)
Barbosa-Silva, 2016[38]	EWGSOP, 2010	15 (8.38)	≥6	31 (17.3)	2.11 (NR)*	0.79 (NR)*	2.66 (NR)*	79.9 (NR)*	0.592 (0.445, 0.739)
Beaudart, 2018[39]	EWGSOP, 2010	51 (16.7)	≥4	52 (17.0)	2.71 (1.67, 4.40)*	0.74 (0.60, 0.91)*	3.80 (NR)*	78.4 (NR)*	NR
	AWGS, 2014	12 (3.92)	≥4	52 (17.0)	2.70 (1.52, 4.79)*	0.70 (0.49, 0.99)*	3.88 (NR)*	81.7 (NR)*	NR
	FNIH**, 2014	11 (3.59)	≥4	52 (17.0)	2.71 (NR)*	0.70 (NR)*	3.89 (NR)*	82.7 (NR)*	NR
	IWGS, 2011	37 (12.1)	≥4	52 (17.0)	3.23 (2.00, 5.21)*	0.66 (0.49, 0.87)*	4.93 (NR)*	81.4 (NR)*	NR
	SCWD, 2011	18 (5.90)	≥4	52 (17.0)	3.81 (2.31, 6.27)*	0.52 (0.31, 0.87)*	7.32 (NR)*	83.7 (NR)*	NR
	Baumgartner, 1998	81 (26.5)	≥4	52 (17.0)	1.47 (NR)*	0.92 (NR)*	1.60 (NR)*	68.3 (NR)*	NR
	Delmonico, 2007	104 (34.0)	≥4	52 (17.0)	1.53 (NR)*	0.91 (NR)*	1.69 (NR)*	64.1 (NR)*	NR
Borges, 2019[40]	EWGSOP2‡, 2018	47 (41.9)	≥4	32 (41.6)	0.84 (NR)*	1.41 (NR)*	0.594 (NR)*	65.8 (NR)*	NR
	EWGSOP2§, 2018	30 (38.5)	≥4	39 (50.6)	1.06 (NR)*	0.97 (NR)*	1.09 (NR)*	40.0 (NR)*	NR
Drey, 2020[41]	EWGSOP, 2010	8 (6.84)	≥4	63 (53.8)	1.09 (NR)*	0.93 (NR)*	1.17 (NR)*	98.2 (NR)*	0.515 (0.298, 0.733)
	EWGSOP2‡, 2018	49 (41.9)	≥4	63 (53.8)	2.27 (NR)*	0.37 (NR)*	6.09 (NR)*	NR	0.781 (0.697, 0.864)
	EWGSOP2§, 2018	8 (6.8)	≥4	63 (53.8)	1.19 (NR)*	0.79 (NR)*	1.51 (NR)*	NR	0.580 (0.367, 0.793)
	EWGSOP, 2010	NR	≥3	NR	1.11 (NR)*	0.83 (NR)*	1.36 (NR)*	NR	NR
	EWGSOP2‡, 2018	49 (41.9)	≥3	NR	1.65 (NR)*	0.29 (NR)*	5.67 (NR)*	NR	NR
	EWGSOP2§, 2018	8 (6.8)	≥3	NR	1.10 (NR)*	0.78 (NR)*	1.41 (NR)*	NR	NR
	EWGSOP, 2010	NR	≥2	NR	1.14 (NR)*	0.52 (NR)*	2.19 (NR)*	NR	NR
	EWGSOP2‡, 2018	49 (41.9)	≥2	NR	1.40 (NR)*	0.26 (NR)*	5.44 (NR)*	NR	NR
	EWGSOP2§, 2018	8 (6.8)	≥2	NR	1.14 (NR)*	0.52 (NR)*	2.19 (NR)*	NR	NR
Fu, 2020[42]	EWGSOP2§, 2018	156 (50.5)	≥4	77 (24.9)	2.8 (NR)	0.8 (NR)	3.50 (NR)*	NR	0.680 (0.620, 0.730)
	AWGS, 2014	81 (26.2)	≥4	77 (24.9)	3.4 (NR)	0.7 (NR)	4.86 (NR)*	NR	0.700 (0.640, 0.750)
Ha, 2020[43]	EWGSOP, 2010	36 (31.3)	≥4	72 (62.6)	1.66 (1.29, 2.13)*	0.29 (0.12, 0.67)*	5.75 (NR)*	60.0 (NR)	NR
	EWGSOP2§, 2018	43 (37.4)	≥4	72 (62.6)	2.21 (1.68, 2.91)*	0.08 (0.02, 0.32)*	27.1 (NR)*	71.3 (NR)	NR
	AWGS, 2014	35 (30.4)	≥4	72 (62.6)	1.54 (1.20, 1.99)*	0.37 (0.17, 0.80)*	4.16 (NR)*	57.4 (NR)	NR
	IWGS, 2011	60 (52.2)	≥4	72 (62.6)	1.36 (1.01, 1.83)*	0.60 (0.37, 0.98)*	2.27 (NR)*	60.0 (NR)	NR
Ida, 2017[44]	EWGSOP, 2010	T: 53 (25.6) M: 41 (32.5) F: 12 (14.8)	≥4	41 (19.8) 18 (14.3) 23 (28.4)	0.94 (0.49, 1.78)* 1.04 (0.42, 2.57)* 1.21 (0.50, 2.94)*	1.02 (0.87, 1.18)* 0.99 (0.85, 1.16)* 0.92 (0.60, 1.41)*	0.919 (NR)* 1.05 (NR)* 1.32 (NR)*	64.3 (NR)* 62.7 (NR)* 66.7 (NR)*	NR
Ida, 2019[45]	JSH, 2016	T: 75 (53.6) M: 55 (56.1) F: 20 (47.6)	≥4	22 (15.7) 11 (11.2) 11 (26.2)	3.90 (1.39, 10.9)* 3.52 (0.80, 15.4)* 4.95 (1.21, 20.2)*	0.81 (0.70, 0.93)* 0.88 (0.77, 1.0)* 0.61 (0.40, 0.92)*	4.81 (NR)* 4.00 (NR)* 8.11 (NR)*	56.4 (NR)* 51.0 (NR)* 69.0 (NR)*	NR

(*continued on next page*)

Supplementary Table 3 (*continued*)

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	LR+ (95% CI)	LR− (95% CI)	DOR (95% CI)	Accuracy, % (95% CI)	AUC (95% CI)
Kemmler, 2017[46]	EWGSOP, 2010	M: 49 (66.2)	≥4	35 (33.5)	1.62 (0.74, 3.53)*	0.81 (0.59, 1.10)*	2.01 (NR)*	51.4 (NR)*	NR
	FNIH‡‡, 2014	M: 32 (43.2)	≥4	35 (33.5)	1.42 (0.75, 1.68)*	0.83 (0.59, 1.17)*	1.71 (NR)*	58.1 (NR)*	NR
	IWGS, 2011	M: 37 (50.0)	≥4	35 (33.5)	4.00 (1.68, 9.52)*	0.53 (0.37, 0.77)*	7.53 (NR)*	70.3 (NR)*	NR
Kera, 2019[47]	EWGSOP, 2010	T: 75 (10.2) M: 66 (22.9) F: 9 (2.0)	≥4 	24 (3.3) 9 (3.1) 15 (3.4)	2.93 (1.20, 7.14)† 4.20 (1.16, 15.2)* 3.46 (50.8, 23.6)*	0.95 (0.88, 1.01)† 0.94 (0.88, 1.01)* 0.92 (72.9, 1.16)*	3.08 (NR)* 4.47 (NR)* 3.77 (NR)*	88.1 (NR)* 77.4 (NR)* 95.1 (NR)*	NR 0.587 (0.504, 0.670) 0.562 (0.376, 0.747)
	AWGS, 2014	T: 75 (10.2) M: 19 (6.60) F: 56 (12.6)	≥4 	24 (3.3) 9 (3.1) 15 (3.4)	1.76 (0.62, 5.00)† 1.77 (0.23, 13.4)* 1.74 (0.51, 6.00)*	0.98 (0.92, 1.03)† 0.976 (0.88, 1.09)* 0.98 (0.98, 0.92)*	1.80 (NR)* 1.81 (NR)* 1.78 (NR)*	87.6 (NR)* 91.0 (NR)* 85.4 (NR)*	NR 0.620 (0.474, 0.766) 0.628 (0.551, 0.705)
	Japanese AWGS, 2012	T: 167 (22.8) M: 52 (18.1) F: 115 (25.8)	≥4 	24 (3.3) 9 (3.1) 15 (3.4)	2.43 (1.10, 5.35)† 1.30 (0.28, 6.06)* 3.28 (1.22, 8.84)*	0.96 (0.93, 1.00)† 0.991 (0.93, 1.05)* 0.951 (0.90, 1.00)*	2.53 (NR)* 1.31 (NR)* 3.45 (NR)*	76.7 (NR)* 80.2 (NR)* 74.4 (NR)*	NR 0.571 (0.480, 0.663) 0.637 (0.578, 0.697)
Kera, 2020[48]	AWGS, 2014	129 (12.2)	≥4	30 (2.8)	1.44 (0.56, 3.70)*	0.99 (0.95, 1.02)*	1.46 (NR)*	85.9 (NR)*	0.63 (0.57, 0.68)
Kim, 2018[49]	EWGSOP, 2010	T: 95 (7.77) M: 48 (8.3) F: 47 (7.3)	≥4 	123 (10.1) 24 (4.2) 99 (15.3)	2.88 (1.94, 4.26)* 5.51 (2.49, 12.2)* 2.45 (1.57, 3.83)*	0.82 (0.73, 0.92)* 0.86 (0.76, 0.98)* 0.77 (0.62, 0.94)*	3.52 (NR)* 6.41 (NR)* 3.20 (NR)*	86.1 (NR) 90.3 (NR) 82.3 (NR)	NR
	AWGS, 2014	T: 125 (10.2) M: 66 (11.4) F: 59 (9.1)	≥4 	123 (10.1) 24 (4.2) 99 (15.3)	2.83 (1.96, 4.09)* 6.55 (3.06, 14.0)* 2.36 (1.55, 3.60)*	0.83 (0.75, 0.92)* 0.86 (0.77, 0.95)* 0.79 (0.66, 0.94)*	3.41 (NR)* 7.66 (NR)* 3.00 (NR)*	84.6 (NR) 88.2 (NR) 81.4 (NR)	NR
	FNIH††, 2014	T: 95 (7.77) M: 59 (10.2) F: 36 (5.6)	≥4 	123 (10.1) 24 (4.2) 99 (15.3)	1.90 (1.19, 3.04)* 3.62 (1.56, 8.36)* 1.90 (1.09, 3.33)*	0.91 (0.82, 1.00)* 0.91 (0.83, 1.00)* 0.85 (0.69, 1.04)*	2.10 (NR)* 3.97 (NR)* 2.25 (NR)*	84.9 (NR) 88.0 (NR) 82.2 (NR)	NR
	FNIH‡‡, 2014	T: 23 (1.88) M: 10 (1.7) F: 13 (2.0)	≥4 	123 (10.1) 24 (4.2) 99 (15.3)	4.61 (2.80, 7.59)* 18.9 (9.57, 37.3)* 2.05 (0.89, 4.72)*	0.62 (0.44, 0.89)* 0.41 (0.19, 0.88)* 0.82 (0.57, 1.17)*	7.39 (NR)* 45.8 (NR)* 2.51 (NR)*	89.7 (NR) 96.2 (NR) 83.9 (NR)	NR
	IWGS, 2011	T: 167 (13.7) M: 74 (12.8) F: 94 (14.6)	≥4 	123 (10.1) 24 (4.2) 99 (15.3)	3.62 (2.61, 5.03)* 9.52 (4.39, 20.6)* 2.67 (1.86, 3.85)*	0.79 (0.72, 0.87)* 0.83 (0.74, 0.92)* 0.77 (0.66, 0.88)*	4.58 (NR)* 11.5 (NR)* 3.50 (NR)*	83.6 (NR) 87.9 (NR) 79.7 (NR)	NR
	CFNIH, 2015	T: 78 (6.38) M: 46 (8.0) F: 32 (5.0)	≥4 	123 (10.1) 24 (4.2) 99 (15.3)	2.04 (1.25, 3.32)* 3.85 (1.61, 9.22)* 1.92 (1.07, 3.44)*	0.89 (0.80, 1.00)* 0.90 (0.80, 1.01)* 0.84 (0.68, 1.05)*	2.29 (NR)* 4.28 (NR)* 2.27 (NR)*	86.0 (NR) 89.9 (NR) 82.5 (NR)	NR
Kim, 2019[50]	EWGSOP2§, 2018	195 (9.3)	≥4	152 (7.2)	2.40 (1.66, 3.48)*	0.90 (0.85, 0.96)*	2.66 (NR)*	86.3 (NR)*	NR
	EWGSOP2‖, 2018	38 (1.8)	≥4	152 (7.2)	6.38 (4.25, 9.59)*	0.62 (0.47, 0.81)*	10.3 (NR)*	92.5 (NR)*	NR
Kotlarczyk, 2018[51]	EWGSOP, 2010	F: 11 (7.80)	≥4	30 (21.3)	0.84 (0.23, 3.08)*	1.04 (0.78, 1.40)*	0.81 (NR)*	73.8 (NR)*	NR
	FNIH††, 2014	F: 6 (4.25)	≥4	30 (21.3)	1.61 (0.50, 5.22)*	0.84 (0.48, 1.49)*	1.91 (NR)*	77.3 (NR)*	NR
	FNIH§§, 2014	F: 46 (32.6)	≥4	30 (21.3)	1.38 (0.73, 2.61)*	0.91 (0.75, 1.11)*	1.51 (NR)*	63.1 (NR)*	NR

(*continued on next page*)

Supplementary Table 3 (*continued*)

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	LR+ (95% CI)	LR− (95% CI)	DOR (95% CI)	Accuracy, % (95% CI)	AUC (95% CI)
Kurita, 2019[52]	EWGSOP2[§], 2018	19 (1.98)	≥6	67 (6.99)	3.99 (1.81, 8.79)*	0.79 (0.60, 1.03)*	5.06 (NR)*	92.1 (NR)*	NR
	EWGSOP2[§], 2018	19 (1.98)	≥5	161 (16.8)	1.59 (0.74, 3.41)*	0.88 (0.67, 1.16)*	1.80 (NR)*	82.3 (NR)*	NR
	EWGSOP2[§], 2018	19 (1.98)	≥4	306 (31.9)	1.50 (0.93, 2.43)*	0.77 (0.50, 1.18)*	1.95 (NR)*	68.0 (NR)*	0.558 (0.399, 0.716)
	EWGSOP2[§], 2018	19 (1.98)	≥3	454 (47.3)	1.23 (0.83, 1.81)*	0.80 (0.47, 1.35)*	1.54 (NR)*	53.0 (NR)*	NR
	EWGSOP2[§], 2018	19 (1.98)	≥2	623 (65.0)	0.89 (0.60, 1.31)*	1.21 (0.71, 2.06)*	0.74 (NR)*	35.3 (NR)*	NR
	EWGSOP2[§], 2018	19 (1.98)	≥1	815 (85.0)	0.99 (0.81, 1.21)*	1.05 (0.37, 3.01)*	0.94 (NR)*	16.4 (NR)*	NR
	AWGS, 2014	36 (3.75)	≥6	67 (6.99)	2.99 (1.47, 6.07)*	0.86 (0.73, 1.01)*	3.47 (NR)*	90.7 (NR)*	NR
	AWGS, 2014	36 (3.75)	≥5	161 (16.8)	1.34 (0.72, 2.51)*	0.93 (0.78, 1.11)*	1.44 (NR)*	81.1 (NR)*	NR
	AWGS, 2014	36 (3.75)	≥4	306 (31.9)	1.32 (0.89, 1.97)*	0.85 (0.64, 1.13)*	1.55 (NR)*	67.5 (NR)*	0.557 (0.452, 0.662)
	AWGS, 2014	36 (3.75)	≥3	454 (47.3)	1.31 (99.7, 1.71)*	0.73 (0.48, 1.11)*	1.79 (NR)*	53.5 (NR)*	NR
	AWGS, 2014	36 (3.75)	≥2	623 (65.0)	0.98 (0.77, 1.26)*	1.03 (0.66, 1.61)*	0.95 (NR)*	36.1 (NR)*	NR
	AWGS, 2014	36 (3.75)	≥1	815 (85.0)	1.01 (88.7, 1.16)*	0.92 (0.40, 2.11)*	1.10 (NR)*	17.7 (NR)*	NR
Li, 2019[53]	AWGS, 2014	35 (25.4)	≥4	23 (16.7)	5.52 (2.56, 11.9)*	0.62 (0.46, 0.83)*	8.90 (NR)*	79.7 (NR)*	0.720 (0.610, 0.830)
Lima, 2020[54]	EWGSOP2[‡], 2018	123 (56.4)	≥4	121 (55.5)	0.211 (0.15, 0.30)*	N/A‖‖‖	N/A‖‖‖	11.9 (NR)*	NR
Mienche, 2019[57]	AWGS, 2014	19 (15.8)	≥4	32 (26.7)	1.23 (0.59, 2.57)	0.92 (0.66, 1.28)	1.34 (NR)*	67.5 (NR)*	0.535 (0.390, 0.680)
Parra-Rodriguez, 2016[58]	EWGSOP, 2010	45 (9.30)	≥4	94 (19.4)	2.00 (1.37, 2.93)†	0.78 (0.65, 0.95)†	2.56 (NR)*	77.9 (NR)*	NR
	AWGS, 2014	54 (11.2)	≥4	93 (19.3)	1.76 (1.14, 2.77)†	0.83 (0.69, 1.0)†	2.12 (NR)*	76.6 (NR)*	NR
	IWGS, 2011	99 (20.8)	≥4	91 (19.1)	1.69 (1.15, 2.49)†	0.86 (1.76, 0.98)†	1.97 (NR)*	71.8 (NR)*	NR
Rolland, 2017[59]	FNIH[††], 2014	F: 49 (1.8)	≥4	401 (14.8)	2.40 (1.62, 3.56)*	0.76 (0.62, 0.94)*	0.27 (NR)*	84.6 (NR)*	NR
Sanchez-Rodriguez, 2019[60]	EWGSOP, 2010	16 (17.8)	≥4	51 (56.7)	1.58 (1.15, 2.19)†	0.38 (0.14, 1.10)†	5.57 (NR)	54.0 (NR)	NR
	EWGSOP2[§], 2018	23 (25.6)	≥4	51 (56.7)	1.59 (1.15, 2.20)†	0.43 (0.19, 0.96)†	5.62 (NR)	58.0 (NR)	NR
Woo, 2014[62]	EWGSOP, 2010	T: 361 (9.0) M: 190 (9.5) F: 171 (8.6)	≥4	150 (3.8) 31 (1.6) 119 (6.0)	2.01 (1.33, 3.05)* 3.20 (1.50, 7.30)† 1.80 (1.09, 2.90)†	0.96 (0.94, 0.99)* 0.97 (0.94, 1.00)† 0.95 (0.91, 1.00)†	2.09 (NR)* 3.30 (NR)* 1.89 (NR)*	88.5 (NR)* 89.7 (NR) 87.2 (NR)	NR
	AWGS, 2014	T: 293 (7.3) M: 187 (9.4) F: 106 (5.3)	≥4	150 (3.8) 31 (1.6) 119 (6.0)	1.83 (1.15, 2.92)* 4.00 (1.85, 8.48)† 1.60 (0.88, 3.04)†	0.97 (0.94, 1.00)* 0.96 (0.93, 1.00)† 0.96 (0.90, 1.02)†	1.89 (NR)* 4.17 (NR)* 1.67 (NR)*	89.9 (NR)* 90.0 (NR) 89.7 (NR)	NR
	IWGS, 2011	T: 806 (20.2) M: 442 (22.1) F: 364 (18.2)	≥4	150 (3.8) 31 (1.6) 119 (6.0)	1.80 (1.29, 2.53)* 4.20 (2.13, 8.61)† 1.50 (1.02, 2.25)†	0.97 (0.96, 0.99)* 0.97 (0.95, 0.99)† 0.97 (0.94, 1.00)†	1.85 (NR)* 4.33 (NR)* 1.55 (NR)*	78.4 (NR)* 78.0 NR) 78.8 (NR)	NR
Yang, 2018[63]	EWGSOP, 2010	T: 45 (11.7) M: 17 (10.6) F: 28 (12.5)	≥4	47 (12.2) 15 (9.4) 32 (14.3)	4.5 (2.1, 9.7) 1.70 (0.2, 13.6) 5.6 (90.8, 97.5)	0.8 (0.7, 1.0) 0.9 (0.9, 1.0) 0.8 (0.6, 1.0)	5.63 (NR)* 1.89 (NR)* 7.00 (NR)*	NR	0.81 (0.77, 0.85) 0.83 (0.77, 0.88) 0.81 (0.75, 0.86)

(*continued on next page*)

Supplementary Table 3 (*continued*)

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	LR+ (95% CI)	LR− (95% CI)	DOR (95% CI)	Accuracy, % (95% CI)	AUC (95% CI)
	AWGS, 2014	T: 61 (15.9) M: 19 (11.9) F: 42 (18.8)	≥4	47 (12.2) 15 (9.4) 32 (14.3)	15.9 (6.6, 38.4) 7.4 (1.6, 34.2) 21.7 (6.6, 71.5)	0.7 (0.6, 0.8) 0.8 (0.7, 1.0) 0.7 (0.5, 0.8)	22.7 (NR)* 9.25 (NR)* 31.0 (NR)*	NR	0.89 (0.86, 0.92) 0.86 (0.82, 0.93) 0.90 (0.86, 0.94)
	FNIH**, 2014	T: 59 (15.4) M: 23 (14.4) F: 36 (16.1)	≥4	47 (12.2) 15 (9.4) 32 (14.3)	16.5 (6.8, 39.9) 29.8 (3.6, 47.1) 13.6 (5.2, 35.7)	0.7 (0.6, 0.8) 0.8 (0.6, 1.0) 0.7 (0.5, 0.8)	23.6 (NR)* 37.3 (NR)* 19.4 (NR)*	NR	0.81 (0.77, 0.85) 0.82 (0.75, 0.88) 0.81 (0.76, 0.86)
	IWGS, 2011	T: 96 (25.0) M: 39 (24.4) F: 57 (25.4)	≥4	47 (12.2) 15 (9.4) 32 (14.3)	11.4 (4.4, 29.7) 6.2 (1.2, 32.6) 14.7 (4.4, 48.8)	0.8 (0.7, 0.9) 0.9 (0.8, 1.0) 0.8 (0.6, 0.9)	14.3 (NR)* 6.89 (NR)* 18.4 (NR)*	NR	0.79 (0.75, 0.83) 0.77 (0.69, 0.83) 0.82 (0.77, 0.87)
Yang, 2019[64]	EWGSOP, 2010	90 (32.5)	≥4	90 (32.5)	5.5 (2.2, 13.7)	0.85 (0.8, 0.9)	6.47 (NR)*	NR	0.782 (0.728, 0.829)
	AWGS, 2014	95 (34.3)	≥4	90 (32.5)	12.1 (3.7, 40.0)	0.8 (0.7, 0.9)	15.1 (NR)*	NR	0.775 (0.721, 0.823)
	FNIH**, 2014	87 (31.4)	≥4	90 (32.5)	13.8 (4.2, 45.5)	0.8 (0.7, 0.9)	17.3 (NR)*	NR	0.791 (0.738, 0.837)
	IWGS, 2011	106 (38.3)	≥4	90 (32.5)	7.3 (2.5, 20.9)	0.8 (0.8, 0.9)	9.13 (NR)*	NR	0.769 (0.715, 0.817)
SARC-CalF (0-18 points)									
Bahat, 2018a[36]	EWGSOP, 2010	8 (3.86)	≥11‖‖	5 (2.42)	16.6 (NR)	0.8 (NR)	16.3 (NR)*	NR	0.59 (0.519, 0.657)
	FNIH**, 2014	19 (9.18)	≥11‖‖	5 (2.42)	6.7 (NR)	0.9 (NR)	7.25 (NR)*	NR	0.712 (0.645, 0.763)
	IWGS, 2011	4 (1.93)	≥11‖‖	5 (2.42)	33.8 (NR)	0.5 (NR)	66.5 (NR)*	NR	0.671 (0.603, 0.735)
	SCWD, 2011	5 (2.42)	≥11‖‖	5 (2.42)	26.9 (NR)	0.6 (NR)	44.2 (NR)*	NR	0.575 (0.504, 0.643)
	EWGSOP, 2010	8 (3.86)	≥11***	21 (10.1)	2.6 (NR)	0.8 (NR)	2.99 (NR)*	NR	0.746 (0.681, 0.804)
	FNIH**, 2014	19 (9.18)	≥11***	21 (10.1)	1.6 (NR)	0.9 (NR)	1.77 (NR)*	NR	0.682 (0.614, 0.745)
	IWGS, 2011	4 (1.93)	≥11***	21 (10.1)	5.3 (NR)	0.6 (NR)	9.67 (NR)*	NR	0.836 (0.778, 0.883)
	SCWD, 2011	5 (2.42)	≥11***	21 (10.1)	4.3 (NR)	0.70 (NR)	6.41 (NR)*	NR	0.705 (0.638, 0.767)
Barbosa-Silva, 2016[38]	EWGSOP, 2010	15 (8.38)	≥11†††	38 (21.2)	3.90 (2.39, 6.39)*	0.40 (0.20, 0.83)*	9.70 (NR)*	81.6 (NR)*	0.736 (0.575, 0.897)
Fu, 2020[42]	EWGSOP2[6], 2018	156 (50.5)	≥11†††	166 (53.7)	2.3 (NR)	0.5 (NR)	4.60 (NR)*	NR	0.700 (0.650, 0.750)
	AWGS, 2014	81 (26.2)	≥11†††	166 (53.7)	2.20 (NR)	0.4 (NR)	5.50 (NR)*	NR	0.750 (0.700, 0.800)
Mienche, 2019[57]	AWGS, 2014	19 (15.8)	≥4‡‡‡	6 (5.0)	10.6 (2.1, 54.0)	0.81 (0.64, 1.02)	13.1 (NR)*	85.8 (NR)*	0.797 (0.67, 0.92)
Urzi, 2017[61]	EWGSOP, 2010	T: 31 (38.7) M: 8 (33.3) F: 23 (41.1)	≥11†††	29 (36.2) 8 (33.3) 21 (37.5)	7.59 (3.24, 17.8)* 14.0 (2.06, 95.1)* 6.10 (2.36, 15.8)*	0.25 (0.13, 0.49)* 0.13 (0.02, 0.84)* 0.30 (0.15, 0.60)*	30.2 (NR)* 104 (NR)* 20.5 (NR)*	85.0 (NR)* 91.7 (NR)* 82.1 (NR)*	0.84 (0.74, 0.94) 0.91 (0.75, 1.00) 0.81 (0.680, 0.93)
Yang, 2018[63]	EWGSOP, 2010	T: 45 (11.7) M: 17 (10.6) F: 28 (12.5)	≥11†††	99 (25.8) 29 (18.1) 70 (31.3)	5.2 (3.3, 8.1) 4.7 (1.8, 12.3) 5.7 (3.1, 10.5)	0.6 (0.4, 0.8) 0.8 (0.6, 1.0) 0.6 (0.4, 0.8)	8.67 (NR)* 5.88 (NR)* 7.13 (NR)*	NR	0.85 (0.81, 0.89) 0.87 (0.91, 0.92) 0.85 (0.79, 0.89)
	AWGS, 2014	T: 61 (15.9) M: 19 (11.9) F: 42 (18.8)	≥11†††	99 (25.8) 29 (18.1) 70 (31.3)	11.5 (7.0, 19.1) 13.3 (5.0, 35.7) 20.8 (8.4, 51.3)	0.4 (0.3, 0.6) 0.6 (0.5, 0.8) 0.4 (0.3, 0.6)	28.8 (NR)* 22.2 (NR)* 52.0 (NR)*	NR	0.92 (0.89, 0.94) 0.88 (0.80, 0.91) 0.93 (0.89, 0.96)

(*continued on next page*)

Supplementary Table 3 (*continued*)

First Author, Year	Sarcopenia		SARC-F						
	Definition, Year	Prevalence, *n* (%)	Cutoff	Prevalence, *n* (%)	LR+ (95% CI)	LR− (95% CI)	DOR (95% CI)	Accuracy, % (95% CI)	AUC (95% CI)
	FNIH**, 2014	T: 59 (15.4) M: 23 (14.4) F: 36 (16.1)	≥11[†††]	99 (25.8) 29 (18.1) 70 (31.3)	8.6 (5.4, 13.9) 14.9 (5.1, 43.5) 8.5 (4.4, 16.5)	0.5 (0.4, 0.6) 0.6 (0.4, 0.8) 0.5 (0.4, 0.7)	17.2 (NR)* 24.8 (NR)* 17.0 (NR)*	NR	0.89 (0.86, 0.92) 0.89 (0.83, 0.93) 0.90 (0.85, 0.94)
	IWGS, 2011	T: 96 (25.0) M: 39 (24.4) F: 57 (25.4)	≥11[†††]	99 (25.8) 29 (18.1) 70 (31.3)	9.5 (5.3, 16.9) 18.6 (4.4, 79.6) 7.7 (4.1, 14.4)	0.6 (0.5, 0.7) 0.7 (0.6, 0.9) 0.5 (0.4, 0.7)	15.8 (NR)* 26.6 (NR)* 15.4 (NR)*	NR	0.83 (0.79, 0.87) 0.80 (0.73, 0.86) 0.86 (0.81, 0.91)
Yang, 2019[64]	EWGSOP, 2010	90 (32.5)	≥11[†††]	131 (47.3)	3.8 (2.6, 5.5)	0.5 (0.4, 0.6)	7.60 (NR)*	NR	0.821 (0.771, 0.865)
	AWGS, 2014	95 (34.3)	≥11[†††]	131 (47.3)	4.1 (2.8, 6.1)	0.5 (0.4, 0.6)	8.20 (NR)*	NR	0.816 (0.765, 0.860)
	FNIH**, 2014	87 (31.4)	≥11[†††]	131 (47.3)	4.7 (3.2, 6.9)	0.6 (0.5, 0.7)	7.83 (NR)*	NR	0.867 (0.821, 0.905)
	IWGS, 2011	106 (38.3)	≥11[†††]	131 (47.3)	4.1 (2.7, 6.3)	0.5 (0.4, 0.6)	8.20 (NR)*	NR	0.817 (0.766, 0.861)
Mienche, 2019[57]	AWGS, 2014	19 (15.8)	≥4[§§§]	5 (4.17)	21.3 (2.51, 179)	0.80 (0.63, 1.01)	26.6 (NR)*	86.7 (NR)*	0.804 (0.68, 0.93)
Mienche, 2019[57]	AWGS, 2014	19 (15.8)	≥4[‡‡‡,§§§]	4 (3.33)	16.0 (1.75, 145)	0.85 (0.70, 1.03)	18.8 (NR)*	85.8 (NR)*	0.862 (0.76, 0.96)
SARC-F+E (0-20 points)									
Kurita, 2019[52]	EWGSOP, 2010	19 (1.98)	≥9	337 (35.1)*	2.49 (2.01, 3.08)*	0.24 (0.08, 0.68)*	10.4 (NR)*	66.5 (NR)*	0.760 (0.640, 0.880)
	AWGS, 2014	36 (3.75)	≥9	334 (34.8)*	1.90 (1.46, 2.46)*	0.55 (0.35, 0.84)*	3.49 (NR)*	66.2 (NR)*	0.663 (0.561, 0.765)
SARC-F+EBM (0-30 points)									
Kurita, 2019[52]	EWGSOP, 2010	19 (1.98)	≥12	312 (32.5)	2.70 (2.17, 3.35)*	0.23 (0.08, 0.65)*	11.8 (NR)*	69.1 (NR)*	0.876 (0.815, 0.936)
	AWGS, 2014	36 (3.75)	≥12	309 (32.2)	2.55 (2.09, 3.12)*	0.32 (0.17, 0.59)*	7.99 (NR)*	69.9 (NR)*	0.824 (0.762, 0.886)

AUC, area under the curve; AWGS, Asian Working Group for Sarcopenia; CC, calf circumference; CFNIH, Chinese FNIH; CI, confidence interval; cont., continued; DOR, diagnostic odds ratio; E, elderly/old age; EBM, elderly/old age and body mass index; EWGSOP, European Working Group on Sarcopenia in Older People; EWGSOP2, revised EWGSOP definition; F, female; FNIH, Foundation for the National Institutes of Health; IWGS, International Working Group on Sarcopenia; JSH, Japan Society of Hepatology; LR−, negative likelihood ratio; LR+, positive likelihood ratio; M, male; NR, not reported; PPV, positive predictive value; SCWD, Society on Sarcopenia, Cachexia and Wasting Disorders; T, total; TC, thigh circumference.

Sarcopenia according to FNIH on **low lean mass, weakness and slowness; [††]low lean mass and weakness; [‡‡]low lean mass and slowness; and [§§]low lean mass and weakness. CC cutoff points: [‖‖]≤31 cm; ***≤31 cm; [†††]M: ≤34 cm, F: ≤33 cm; [‡‡‡]M: <34 cm, F: <29 cm. TC cutoff points: [§§§]M: <49 cm, F: <44 cm. [‖‖‖]Corresponding specificity equals to zero.

*Point estimate and 95% CI.

[†]95% CI calculated based on reported data.

[‡]Probable, [§]confirmed, and [‖]severe sarcopenia according to EWGSOP2.

Supplementary Table 4
Diagnostic Criteria of Sarcopenia by Articles That Reported on the Concurrent Validity of the SARC-F

First Author, Year	Sarcopenia Definition, Year	Diagnostic Criteria								
		Muscle Mass				Muscle Strength			Physical Performance	
		Method	Device	Measure	Cutoff	Method	Device[‡‡]	Cutoff	Method	Cutoff
Bahat, 2018a[36]	EWGSOP, 2010	BIA	Tanita BC532	SMI	NR	HGS	Jamar hydraulic	M: <32 kg F: <22 kg	GS (4 m)[§§]	NR
	FNIH, 2014	BIA	Tanita BC532	ASM/BMI	NR	HGS	Jamar hydraulic	M: <26 kg F: <16 kg	GS (4 m)[§§]	NR
	IWGS, 2011	BIA	Tanita BC532	SMI	NR	HGS	Jamar hydraulic	M: <32 kg F: <22 kg	GS (4 m)[§§]	NR
	SCWD, 2011	BIA	Tanita BC532	SMI	NR	HGS	Jamar hydraulic	M: <32 kg F: <22 kg	GS (4 m)[§§]	NR
Barbosa-Silva, 2016[38]	EWGSOP, 2010	DXA	GE Lunar Prodigy	ALMI	M: <7.76 kg/m^2 F: <5.62 kg/m^2	HGS	Jamar Plus digital	M: <30 kg F: <20 kg	GS (4 m)[‖‖]	<0.8 m/s
Beaudart, 2018[39]	EWGSOP, 2010	DXA	Hologic Discovery A	ALMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	HGS	Saehan hydraulic	M: <30 kg F: <20 kg	GS (4 m)[‖‖]	≤0.8 m/s
	AWGS, 2014	DXA	Hologic Discovery A	ALMI	M: <7.00 kg/m^2 F: <5.40 kg/m^2	HGS	Saehan hydraulic	M: <26 kg F: <18 kg	GS (4 m)[‖‖]	≤0.8 m/s
	FNIH, 2014	DXA	Hologic Discovery A	ALM/BMI	M: <0.789 F: <0.512	HGS	Saehan hydraulic	M: <26 kg F: <16 kg	GS (4 m)[‖‖]	≤0.8 m/s
	IWGS, 2011	DXA	Hologic Discovery A	ALMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	GS (4 m)[‖‖]	<1.0 m/s
	SCWD, 2011	DXA	Hologic Discovery A	ALMI	>2 SD below mean of healthy persons 20-30 y of same ethnicity	N/A	N/A	N/A	GS (4 m)[‖‖]	≤1.0 m/s
	Baumgartner, 1998	DXA	Hologic Discovery A	ALMI	>2 SD below mean of healthy young persons	N/A	N/A	N/A	N/A	N/A
	Delmonico, 2007	DXA	Hologic Discovery A	ALMI	M: ≤7.25 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	N/A	N/A
Borges, 2019[40]	EWGSOP2[†], 2018	BIA	Seca mBCA 525	SMI	NR	N/A	N/A	N/A	N/A	N/A
Drey, 2020[41]	EWGSOP, 2010	DXA	GE Lunar Prodigy	ALMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	HGS	Jamar hydraulic	M: <30 kg F: <20 kg	GS (4 m)[‖‖]	≤0.8 m/s
	EWGSOP2[*], 2018	N/A	N/A	N/A	N/A	HGS	Jamar hydraulic	M: <27 kg F: <16 kg	GS (4 m)[‖‖] CST SPPB	≤0.8 m/s >15 s NR
	EWGSOP2[†], 2018	DXA	GE Lunar Prodigy	ALMI	M: <7.0 kg/m^2 F: <5.5 kg/m^2	HGS	Jamar hydraulic	M: <27 kg F: <16 kg	GS (4 m)[‖‖] CST SPPB	≤0.8 m/s >15 s NR
Fu, 2020[42]	EWGSOP2[†], 2018	CT	ImageJ software	L3-SMI	M: <55 cm^2/m^2 F: <39 cm^2/m^2	HGS	Camry EH101	M: <30 kg F: <20 kg	N/A	N/A
	AWGS, 2014	CT	ImageJ software	L3-SMI	M: <40.8 cm^2/m^2 F: <34.9 cm^2/m^2	HGS	Camry EH101	M: <26 kg F: <18 kg	N/A	N/A
Ha, 2020[43]	EWGSOP, 2010	DXA	GE Lunar DPX-NT	ALMI	M: <7.26 kg/m^2 F: <5.45 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <30 kg F: <20 kg	N/A	N/A
	EWGSOP2[†], 2018	DXA	GE Lunar DPX-NT	ALMI	M: <7.0 kg/m^2 F: <6.0 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <27 kg F: <16 kg	N/A	N/A
	AWGS, 2014	DXA	GE Lunar DPX-NT	ALMI	M: <7.0 kg/m^2 F: <5.4 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <26 kg F: <18 kg	N/A	N/A
	IWGS, 2011	DXA	GE Lunar DPX-NT	ALMI	M: <7.23 kg/m^2 F: <5.67 kg/m^2	N/A	N/A	N/A	N/A	N/A
Ida, 2017[44]	EWGSOP, 2010	BIA	InBody 230	ASMI	M: ≤8.8 kg/m^2 F: ≤6.4 kg/m^2	HGS	Smedley spring-type	M: ≤30 kg F: ≤20 kg	GS (11 m)[§§]	≤0.8 m/s
Ida, 2019[45]	JSH, 2016	CT	Dynamic multidetector	L3-SMI	M: ≤6.0 cm^2/m^2 F: ≤3.4 cm^2/m^2	HGS	Smedley spring-type	M: ≤29 kg F: ≤18 kg	N/A	N/A
Kemmler, 2017[46]	EWGSOP, 2010	BIA	InBody 770	ASMI	M: <7.18 kg/m^2 F: N/A	HGS	Jamar hydraulic	M: <30 kg F: N/A	GS (10 m)[§§]	<0.8 m/s

	FNIH[**], 2014	BIA	InBody 770	ASM/BMI	M: <0.789 F: N/A	HGS	Jamar hydraulic	M: <26 kg F: N/A	N/A	N/A
	IWGS, 2011	BIA	InBody 770	ASMI	M: ≤7.23 kg/m^2 F: N/A	N/A	N/A	N/A	GS (10 m)[§§]	<1.0 m/s
Kera, 2019[47]	EWGSOP, 2010	BIA	InBody 770	SMI	M: <10.76 kg/m^2 F: <6.76 kg/m^2	HGS	Smedley spring-type	M: <30 kg F: <20 kg	GS (5 m)[§§]	<0.8 m/s
	AWGS, 2014	BIA	InBody 770	SMI	M: <7.0 kg/m^2 F: <5.7 kg/m^2	HGS	Smedley spring-type	M: <26 kg F: <18 kg	GS (5 m)[§§]	<0.8 m/s
	J-AWGS, 2012	BIA	InBody 770	SMI	M: <7.0 kg/m^2 F: <5.8 kg/m^2	HGS	Smedley spring-type	M: <30.3 kg F: <19.3 kg	GS (5 m)[§§]	M: <1.27 m/s F: <1.19 m/s
Kera, 2020[48]	AWGS, 2014	BIA	InBody 770	ALMI	M: <7.0 kg/m^2 F: <5.7 kg/m^2	HGS	Smedley spring-type	M: <26 kg F: <18 kg	GS (5 m/ 10 m)[‖‖,***]	<0.8 m/s
Kim, 2018[49]	EWGSOP, 2010	DXA	GE Lunar & Hologic	ALMI	M: <6.43 kg/m^2 F: <5.34 kg/m^2	HGS	GRIP-D T.K.K.5401	M: ≤26.5 kg F: ≤16.9 kg	GS (7 m)[§§] TUG[§§]	≤0.8 m/s NR
	AWGS, 2014	DXA	GE Lunar & Hologic	ALMI	M: <7.0 kg/m^2 F: <5.4 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <26 kg F: <18 kg	GS (7 m)[§§]	<0.8 m/s
	FNIH[‖], 2014	DXA	GE Lunar & Hologic	ALM/BMI	M: <0.789 F: <0.512	HGS	GRIP-D T.K.K.5401	M: <26 kg F: <16 kg	N/A	N/A
	FNIH[**], 2014	DXA	GE Lunar & Hologic	ALM/BMI	M: <0.789 F: <0.512	N/A	N/A	N/A	GS (7 m)[§§]	≤0.8 m/s
	IWGS, 2011	DXA	GE Lunar & Hologic	ALMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	GS (7 m)[§§]	<1.0 m/s
	CFNIH, 2015	DXA	GE Lunar & Hologic	ALM/BMI	M: <0.722 F: <0.472	HGS	GRIP-D T.K.K.5401	M: <27 kg F: <17 kg	N/A	N/A
Kim, 2019[50]	EWGSOP2[†], 2018	DXA	GE Lunar & Hologic	ALMI	M: <7.0 kg/m^2 F: <6.0 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <27 kg F: <16 kg	N/A	N/A
	EWGSOP2[‡], 2018	DXA	GE Lunar & Hologic	ALMI	M: <7.0 kg/m^2 F: <6.0 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <27 kg F: <16 kg	GS (4 m)[§§] CST TUG (3 m) SPPB	≤0.8 m/s <15 s ≥20 s ≤8 points
Kotlarczyk, 2018[51]	EWGSOP, 2010	DXA	Hologic Discovery A	ALMI	M: N/A F: ≤5.5 kg/m^2	HGS	Jamar Plus digital	M: N/A F: <20 kg	GS (4 m)[‖‖]	<0.8 m/s
	FNIH[‖], 2014	DXA	Hologic Discovery A	ALM/BMI	M: N/A F: <0.512	HGS	Jamar Plus digital	M: N/A F: <16 kg	N/A	N/A
	FNIH[††], 2014	DXA	Hologic Discovery A	ALM/BMI	M: N/A F: <0.591	HGS	Jamar Plus digital	M: N/A F: <19.9 kg	N/A	N/A
Kurita, 2019[52]	EWGSOP2[†], 2018	BIA	Tanita MC-780A	ALMI	M: <7.0 kg/m^2 F: <5.5 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <27 kg F: <16 kg	GS (10 m)[‖‖]	≤0.8 m/s
	AWGS, 2014	BIA	Tanita MC-780A	ALMI	M: <7.0 kg/m^2 F: <5.7 kg/m^2	HGS	GRIP-D T.K.K.5401	M: <26 kg F: <18 kg	GS (10 m)[‖‖]	<0.8 m/s
Li, 2019[53]	AWGS, 2014	DXA	GE Lunar Prodigy	ALMI	M: <7.0 kg/m^2 F: <5.4 kg/m^2	HGS	Camry EH101	M: <26 kg F: <18 kg	GS (6 m) [§§]	<0.8 m/s
Lima, 2020[54]	EWGSOP2[*], 2018	N/A	N/A	N/A	N/A	HGS	Jamar hydraulic	M: <27 kg F: <16 kg	CST	NR
Mienche, 2019[57]	AWGS, 2014	DXA	NR	ALMI	M: <7.0 kg/m^2 F: <5.4 kg/m^2	HGS	Jamar hydraulic	M: <26 kg F: <18 kg	GS (NR distance)[‖‖]	<0.8 m/s
Parra-Rodriguez, 2016[58]	EWGSOP, 2010	DXA	Hologic Discovery Wi	ALMI	M: ≤6.54 kg/m^2 F: ≤5.37 kg/m^2	HGS KES	Jamar hydraulic	M: ≤20 kg F: ≤12 kg NR	GS (6 m)[‖‖] SPPB	M: ≤0.85 m/s F: ≤0.7 m/s NR
	AWGS, 2014	DXA	Hologic Discovery Wi	ALMI	M: ≤6.54 kg/m^2 F: ≤5.37 kg/m^2	HGS KES	Jamar hydraulic	M: ≤20 kg F: ≤12 kg NR	GS (6 m)[‖‖] SPPB	<0.8 m/s
	IWGS, 2011	DXA	Hologic Discovery Wi	ALMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	GS (6 m)[‖‖] SPPB	<1.0 m/s NR
Rolland, 2017[59]	FNIH[††], 2014	DXA	Hologic QDR 4500W	ALM/BMI	M: N/A F: <0.512	HGS	Martin Vigorimeter	M: N/A F: <16 kg	N/A	N/A

(*continued on next page*)

Supplementary Table 4 (*continued*)

First Author, Year	Sarcopenia Definition, Year	Diagnostic Criteria								
		Muscle Mass				Muscle Strength			Physical Performance	
		Method	Device	Measure	Cutoff	Method	Device[‡‡]	Cutoff	Method	Cutoff
Sanchez-Rodriguez, 2019[60]	EWGSOP, 2010	BIA	Bodystat 1500	FFMI	NR	HGS	Jamar hydraulic	NR	GS (4 m)[‖‖]	<0.8 m/s
	EWGSOP2[*], 2018	BIA	Bodystat 1500	FFMI	NR	HGS	Jamar hydraulic	NR	GS (4 m)[‖‖]	<0.8 m/s
Urzi, 2017[61]	EWGSOP, 2010	BIA	Maltron Bioscan 920	SMI	M: <8.87 kg/m^2 F: <6.42 kg/m^2	HGS	Jamar hydraulic	M: <30 kg F: <20 kg	GS (NR distance)[§§]	<0.8 m/s
Woo, 2014[62]	EWGSOP, 2010	DXA	Hologic Delphi W4500	ALMI	M: <6.52 kg/m^2 F: <5.44 kg/m^2	HGS	Jamar hydraulic	M: ≤28 kg F: ≤18 kg	GS (6 m)[‖‖]	<0.8 m/s
	AWGS, 2014	DXA	Hologic Delphi W4500	ALMI	M: <7.0 kg/m^2 F: <5.4 kg/m^2	HGS	Jamar hydraulic	M: <26 kg F: <18 kg	N/A	N/A
	IWGS, 2011	DXA	Hologic Delphi W4500	ALMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	GS (6 m)[‖‖]	<1.0 m/s
Yang, 2018[63]	EWGSOP, 2010	BIA	InBody 230	ASMI	M: <6.28 kg/m^2 F: <5.08 kg/m^2	HGS	Camry EH101	M: <30 kg F: <20 kg	GS (4 m)[§§]	<0.8 m/s
	AWGS, 2014	BIA	InBody 230	ASMI	M: <7.0 kg/m^2 F: <5.7 kg/m^2	HGS	Camry EH101	M: <26 kg F: <18 kg	GS (4 m)[§§]	<0.8 m/s
	FNIH[§], 2014	BIA	InBody 230	ASM/BMI	M: <0.789 F: <0.512	HGS	Camry EH101	M: <26 kg F: <16 kg	GS (4 m)[§§]	<0.8 m/s
	IWGS, 2011	BIA	InBody 230	ASMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	GS (4 m)[§§]	<1.0 m/s
Yang, 2019[64]	EWGSOP, 2010	BIA	InBody 230	ASMI	M: <6.12 kg/m^2 F: <4.97 kg/m^2	HGS	Camry EH101	M: <30 kg F: <20 kg	GS (4 m)[§§]	<0.8 m/s
	AWGS, 2014	BIA	InBody 230	ASMI	M: <7.0 kg/m^2 F: <5.7 kg/m^2	HGS	Camry EH101	M: <26 kg F: <18 kg	GS (4 m)[§§]	<0.8 m/s
	FNIH[§], 2014	BIA	InBody 230	ASM/BMI	M: <0.789 F: <0.512	HGS	Camry EH101	M: <26 kg F: <16 kg	GS (4 m)[§§]	<0.8 m/s
	IWGS, 2011	BIA	InBody 230	ASMI	M: ≤7.23 kg/m^2 F: ≤5.67 kg/m^2	N/A	N/A	N/A	GS (4 m)[§§]	<1.0 m/s

ALM, appendicular lean mass; ALMI, appendicular lean mass index (appendicular lean mass/height^2); ASM, appendicular skeletal muscle; ASMI, appendicular skeletal muscle mass index (appendicular skeletal muscle/height^2); AWGS, Asian Working Group for Sarcopenia; BIA, bioelectrical impedance analysis; BMI, body mass index; CFNIH, Chinese FNIH; cont., continued; CST, chair stand test; CT, computed tomography; DXA, dual energy X-ray absorptiometry; EWGSOP, European Working Group on Sarcopenia in Older People; EWGSOP2, revised EWGSOP definition; F, female; FFMI, fat-free mass index (fat-free mass/height^2); FNIH, Foundation for the National Institutes of Health; GS, gait speed; HGS, handgrip strength; IWGS, International Working Group on Sarcopenia; J-AWGS, Japanese AWGS; JSH, Japan Society of Hepatology; KES, knee extension strength; L3-SMI, third lumbar vertebra skeletal muscle index; M, male; N/A, not applicable; NR, not reported; SCWD, Society on Sarcopenia, Cachexia and Wasting Disorders; SMI, skeletal muscle mass index (skeletal muscle mass/height^2); SPPB, short physical performance battery; TUG, Timed-Up-and-Go test.

*Probable, [†]confirmed, and [‡]severe sarcopenia according to EWGSOP2. Sarcopenia according to FNIH on [§]low lean mass, weakness and slowness; [‖]low lean mass and weakness; **low lean mass and slowness; and [††]low lean mass and weakness. [‡‡]Handheld dynamometers, unless otherwise stated (Martin Vigorimeter). Walking instruction: [§§]usual/comfortable pace; [‖‖]pace not reported. ***5 m (2011 cohort), 10 m (2017 cohort).

Supplementary Table 5
Pooled Concurrent Validity Measures of the SARC-F Versions to Identify Sarcopenia, Grouped by Definition

Population	Sarcopenia	SARC-F											
	Prevalence, *n* (%)	Prevalence, *n* (%)	*N**	Sensitivity		Specificity		PPV		NPV		*N**	AUC	
				Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %		Pt. est., (95% CI)	I^2, %
SARC-F														
EWGSOP														
Overall	932 (11.3)	881 (10.7)	13	34.0 (20.9, 47.1)	97.0	82.9 (79.0, 86.8)	97.0	28.2 (18.7, 37.7)	92.0	89.0 (86.2, 91.8)	91.0	4	0.674 (0.534, 0.815)	94.6
CD	745 (9.9)	615 (8.2)	9	29.6 (17.5, 41.8)	95.4	87.4 (83.5, 91.2)	96.6	26.1 (15.9, 36.4)	92.4	89.9 (86.7, 93.0)	91.5	2	0.667 (0.385, 0.950)	98.0
IP	36 (31.3)	72 (62.6)	1	88.4 (78.8, 98.0)	0.00	43.1 (31.7, 54.6)	0.00	48.1 (37.1, 59.1)	0.00	86.1 (74.8, 97.4)	0.00	0	N/A	N/A
OP	61 (18.8)	104 (32.1)	2	29.9 (1.3, 58.6)	65.9	69.1 (59.8, 78.4)	94.6	24.4 (11.3, 37.5)	0.00	85.3 (78.5, 92.1)	95.7	1	0.515 (0.298, 0.733)	0.00
NH	90 (32.5)	90 (32.5)	1	17.8 (9.40, 26.2)	0.00	96.8 (94.0, 99.7)	0.00	N/A	N/A	N/A	N/A	0	0.782 (0.732, 0.833)	0.00
Between-group heterogeneity (*P* value)				.015		<.001		.37		.45			.69	
EWGSOP2[†]														
Overall	444 (12.0)	721 (19.5)	6	53.3 (18.9, 87.8)	98.9	68.9 (54.4, 83.4)	98.7	23.2 (10.5, 35.9)	96.8	88.5 (82.2, 94.7)	97.5	3	0.644 (0.566, 0.722)	24.1
CD	218 (10.0)	203 (9.3)	2	46.4 (−20.6, 113)	98.0	73.2 (46.3, 100)	97.9	27.4 (−15.3, 70.0)	77.3	89.4 (64.9, 114)	0.00	0	N/A	N/A
IP	218 (15.8)	455 (32.9)	3	55.2 (0.4, 110)	99.2	72.8 (50.9, 94.8)	97.5	29.4 (−13.0, 71.9)	98.8	85.3 (65.5, 105)	98.6	2	0.643 (0.532, 0.753)	50.8
OP	8 (6.8)	63 (53.9)	1	63.0 (29.5, 96.5)	0.00	47.0 (37.0, 57.0)	0.00	8.00 (4.0, 12.0)	0.00	94.0 (88.5, 99.5)	0.00	1	0.580 (0.367, 0.793)	0.00
Between-group heterogeneity (*P* value)				.96		.48		.83		.90			.65	
AWGS														
Overall	1050 (10.4)	1119 (11.1)	13	28.9 (19.7, 38.1)	95.2	86.8 (83.1, 90.5)	98.0	24.5 (16.6, 32.4)	92.8	89.8 (87.5, 92.1)	91.7	7	0.694 (0.592, 0.796)	94.9
CD	749 (9.2)	519 (6.3)	7	17.7 (9.10, 26.3)	90.2	92.9 (89.6, 96.3)	95.4	17.7 (5.2, 30.2)	20.3	91.3 (88.4, 94.2)	83.4	2	0.761 (0.546, 0.977)	98.5
IP	187 (12.3)	478 (31.4)	4	50.2 (37.7, 62.7)	92.3	78.1 (73.0, 83.2)	97.8	38.4 (22.5, 54.4)	97.3	87.8 (83.4, 92.2)	95.0	3	0.660 (0.478, 0.842)	68.1
OP	19 (15.8)	32 (26.7)	1	31.6 (9.6, 53.6)	0.00	74.3 (65.8, 82.8)	0.00	18.8 (7.30, 30.3)	0.00	85.2 (81.2, 89.2)	0.00	1	0.535 (0.390, 0.680)	0.00
NH	95 (34.3)	90 (32.5)	1	20.0 (11.5, 28.5)	0.00	N/A	N/A	N/A	N/A	N/A	N/A	1	0.775 (0.724, 0.826)	0.00
Between-group heterogeneity (*P* value)				<.001		<.001		.12		.21			.65	
FNIH														
Overall	388 (7.3)	817 (15.4)	8	30.6 (22.3, 39.0)	75.7	85.8 (80.4, 91.3)	96.6	15.1 (7.1, 23.2)	91.3	94.7 (91.8, 97.6)	94.4	2	0.772 (0.713, 0.830)	75.2
CD	295 (6.0)	697 (14.2)	6	32.3 (22.6, 42.0)	78.9	86.9 (81.1, 92.6)	97.0	17.1 (7.8, 26.4)	93.1	94.3 (91.0, 97.6)	95.4	1	0.759 (0.652, 0.865)	87.4
NH	87 (31.4)	90 (32.5)	1	21.8 (12.7, 31.0)	0.00	N/A	N/A	N/A	N/A	N/A	N/A	1	0.791 (0.742, 0.841)	0.00
LCR	6 (4.3)	31 (21.3)	1	33.3 (−4.5, 71.1)	0.00	79.3 (72.5, 86.2)	0.00	6.70 (−2.3, 15.7)	0.00	96.4 (92.9, 99.9)	0.00	0	N/A	N/A
Between-group heterogeneity (*P* value)				.69		.35		.37		.60			.59	
IWGS														
Overall	1412 (20.0)	699 (9.9)	9	33.7 (18.9, 48.5)	97.7	88.9 (85.2, 92.6)	95.6	38.2 (24.5, 52.0)	94.9	83.1 (75.7, 90.4)	98.3	2	0.737 (0.657, 0.817)	86.1
CD	1246 (18.7)	537 (8.0)	7	30.4 (14.6, 46.3)	97.5	90.2 (86.5, 93.9)	95.3	34.6 (20.9, 48.3)	94.5	85.6 (77.9, 93.2)	98.5	1	0.718 (0.571, 0.864)	92.9

(*continued on next page*)

Supplementary Table 5 (*continued*)

Population	Sarcopenia	SARC-F											
	Prevalence, *n* (%)	Prevalence, *n* (%)	*N**	Sensitivity		Specificity		PPV		NPV		*N**	AUC	
				Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %		Pt. est., (95% CI)	I^2, %
IP	60 (52.2)	72 (62.6)	1	71.7 (60.3, 83.1)	0.00	47.3 (34.1, 60.5)	0.00	59.7 (48.4, 71.1)	0.00	60.5 (45.9, 75.1)	0.00	0	N/A	N/A
NH	106 (38.3)	90 (32.5)	1	17.0 (9.5, 24.6)	0.00	97.6 (95.0, 100)	0.00	N/A	N/A	N/A	N/A	1	0.769 (0.718, 0.820)	0.00
Between-group heterogeneity (*P* value)				.12		<.001		.18		.044			.69	
SCWD														
Overall	23 (4.5)	91 (17.7)	2	55.3 (49.8, 60.9)	0.00	84.0 (80.5, 87.5)	15.2	12.4 (−1.4, 26.2)	91.2	97.5 (96.1, 98.9)	0.00	0	N/A	N/A
CD	23 (4.5)	91 (17.7)	2	55.3 (49.8, 60.9)	0.00	84.0 (80.5, 87.5)	15.2	12.4 (−1.4, 26.2)	91.2	97.5 (96.1, 98.9)	0.00	0	N/A	N/A
SARC-CalF														
EWGSOP														
Overall	189 (16.8)	302 (26.8)	5	57.2 (43.2, 71.2)	69.6	87.9 (84.8, 91.1)	46.4	54.7 (−0.7, 110)	97.0	92.1 (82.1, 102)	81.2	5	0.810 (0.766, 0.854)	55.1
CD	68 (8.8)	142 (18.4)	3	47.8 (29.0, 66.6)	58.8	88.6 (84.7, 92.4)	59.9	26.3 (11.5, 41.2)	0.00	96.5 (93.1, 100)	0.00	3	0.793 (0.720, 0.865)	77.0
NH	121 (33.9)	160 (44.8)	2	67.2 (49.2, 85.2)	72.6	86.2 (80.2, 92.2)	0.00	82.8 (70.5, 95.1)	0.00	86.3 (78.4, 94.3)	0.00	2	0.828 (0.742, 0.915)	0.00
Between-group heterogeneity (*P* value)				.16		.52		.95		.92			.54	
AWGS														
Overall	256 (23.5)	402 (36.9)	4	53.8 (39.1, 68.6)	81.4	87.7 (78.9, 96.4)	95.8	50.1 (30.6, 69.7)	45.0	86.6 (84.3, 88.9)	0.00	4	0.830 (0.724, 0.937)	95.3
CD	81 (26.2)	166 (53.7)	1	60.7 (47.9, 73.5)	0.00	94.7 (92.1, 97.3)	0.00	N/A	N/A	N/A	N/A	1	0.920 (0.895, 0.945)	0.00
IP	19 (15.8)	6 (5.0)	1	66.6 (55.9, 77.4)	0.00	70.1 (63.6, 76.7)	0.00	44.1 (35.0, 53.3)	0.00	85.5 (80.2, 90.9)	0.00	1	0.750 (0.700, 0.800)	0.00
OP	95 (34.3)	131 (47.3)	1	21.1 (1.3, 40.8)	0.00	98.0 (95.6, 100)	0.00	66.7 (35.2, 98.2)	0.00	86.8 (84.3, 89.3)	0.00	1	0.797 (0.672, 0.922)	N/A
NH	61 (15.9)	99 (25.8)	1	59.0 (48.8, 69.3)	0.00	85.7 (80.4, 91.1)	0.00	N/A	N/A	N/A	N/A	1	0.816 (0.769, 0.864)	0.00
Between-group heterogeneity (*P* value)				.001		<.001		.18		.99			>.99	
FNIH														
Overall	165 (19.0)	251 (28.9)	3	45.9 (19.0, 72.8)	92.1	90.5 (86.5, 94.6)	67.7	N/A	N/A	N/A	N/A	3	0.817 (0.716, 0.918)	93.8
CD	78 (13.2)	120 (20.3)	2	36.1 (−3.3, 75.5)	93.1	92.4 (89.5, 95.3)	29.4	N/A	N/A	N/A	N/A	1	0.788 (0.584, 0.992)	96.9
NH	87 (31.4)	131 (47.3)	1	64.4 (53.9, 74.9)	0.00	86.3 (81.2, 91.5)	0.00	N/A	N/A	N/A	N/A	1	0.867 (0.825, 0.909)	0.00
Between-group heterogeneity (*P* value)				.41		.06		N/A		N/A			.46	
IWGS														
Overall	206 (23.7)	251 (28.9)	3	49.4 (39.2, 59.6)	38.1	91.3 (86.1, 96.4)	81.0	N/A	N/A	N/A	N/A	3	0.827 (0.801, 0.854)	0.00
CD	100 (16.9)	120 (20.3)	2	43.0 (33.0, 53.0)	0.00	93.3 (88.5, 98.1)	75.3	N/A	N/A	N/A	N/A	1	0.832 (0.800, 0.864)	0.00
NH	106 (38.3)	131 (47.3)	1	55.7 (45.9, 65.5)	0.00	86.5 (81.1, 91.9)	0.00	N/A	N/A	N/A	N/A	1	0.817 (0.770, 0.865)	0.00
Between-group heterogeneity (*P* value)				.08		.15		N/A		N/A			.60	

AUC, area under the curve; AWGS, Asian Working Group for Sarcopenia; CD, community-dwellers; CI, confidence interval; EWGSOP, European Working Group on Sarcopenia in Older People; EWGSOP2, revised EWGSOP definition; FNIH, Foundation for the National Institutes of Health; IP, inpatients; IWGS, International Working Group on Sarcopenia; LCR, long-term care residents; N/A, not applicable due to lack of data; NH, nursing home residents; NPV, negative predictive value; OP, outpatients; PPV, positive predictive value; Pt. est., point estimate; SCWD, Society on Sarcopenia, Cachexia and Wasting Disorders.

*Number of articles.

†Confirmed sarcopenia according to EWGSOP2.

Supplementary Table 6
Pooled Concurrent Validity Measures of the SARC-F Versions to Identify Sarcopenia, Grouped by Definition and Sex

	Sarcopenia	SARC-F											
	Prevalence, *n* (%)	Prevalence, *n* (%)	*N**	Sensitivity		Specificity		PPV		NPV		*N**	AUC	
				Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est, % (95% CI)	I^2, %		Pt. est. (95% CI)	I^2, %
EWGSOP														
Male	411 (12.8)	132 (4.1)	6	13.6 (5.1, 22.1)	83.2	95.3 (92.2, 98.5)	81.8	44.0 (32.1, 55.9)	81.9	79.8 (74.4, 85.1)	96.4	2	0.711 (0.475, 0.947)	95.6
Female	278 (7.9)	318 (9.0)	6	21.3 (10.8, 31.9)	69.3	90.4 (87.3, 93.4)	93.5	12.2 (3.2, 21.3)	1.51	93.1 (88.1, 98.0)	92.8	2	0.702 (0.453, 0.950)	84.2
Between-group heterogeneity (*P* value)				.26		.026		<.001		<.001			.96	
AWGS														
Male	291 (9.6)	79 (2.6)	4	9.80 (−0.4, 20.1)	58.0	98.2 (95.6, 101)	40.3	28.9 (15.0, 42.7)	64.6	87.2 (82.8, 91.7)	92.2	2	0.748 (0.483, 1.01)	89.0
Female	263 (7.9)	265 (8.0)	4	18.9 (8.5, 29.2)	86.6	92.9 (90.2, 95.6)	95.2	14.6 (4.0, 25.2)	66.2	92.0 (87.7, 96.2)	90.2	2	0.766 (0.509, 1.02)	97.4
Between-group heterogeneity (*P* value)				.22		.006		.11		.13			.92	
FNIH														
Male	114 (14.1)	74 (9.1)	3	22.2 (10.3, 34.1)	77.9	88.3 (77.2, 99.3)	92.3	39.9 (25.7, 54.0)	64.2	82.3 (73.2, 91.5)	94.2	1	0.820 (0.755, 0.885)	0.00
Female	97 (2.6)	562 (15.1)	4	32.6 (20.9, 44.3)	0.00	87.1 (80.1, 94.1)	96.1	6.50 (1.6, 11.4)	44.9	95.8 (88.1, 104)	0.00	1	0.810 (0.760, 0.860)	0.00
Between-group heterogeneity (*P* value)				.22		.86		<.001		.027			.81	
IWGS														
Male	592 (21.1)	105 (3.7)	4	19.5 (5.7, 33.4)	93.4	97.8 (94.7, 101)	73.2	65.5 (53.0, 77.9)	61.6	80.2 (72.1, 88.2)	96.2	1	0.770 (0.700, 0.840)	0.00
Female	515 (18.0)	250 (8.7)	3	21.8 (6.1, 37.6)	93.4	91.6 (88.2, 95.0)	95.2	28.1 (17.4, 38.8)	0.00	85.3 (76.4, 94.3)	93.4	1	0.820 (0.770, 0.870)	0.00
Between-group heterogeneity (*P* value)				.83		.008		<.001		.40			.98	
SARC-CalF														
EWGSOP														
Male	25 (13.6)	37 (20.1)	2	57.8 (15.9, 99.6)	90.7	93.7 (89.5, 97.9)	0.00	87.5 (63.9, 111)	0.00	93.8 (79.6, 108)	0.00	1	0.910 (0.785, 1.04)	0.00
Female	51 (18.2)	91 (32.5)	2	60.2 (19.7, 101)	74.6	91.4 (87.5, 95.3)	0.00	80.9 (66.2, 95.7)	0.00	82.9 (72.9, 93.0)	0.00	2	0.844 (0.798, 0.891)	0.00
Between-group heterogeneity (*P* value)				.94		.43		.64		.22			.34	
AWGS														
Male	19 (11.9)	29 (18.1)	1	47.4 (24.1, 70.8)	0.00	96.5 (93.1, 100)	0.00	N/A	N/A	N/A	N/A	0	N/A	N/A
Female	42 (18.8)	70 (31.3)	1	57.1 (41.5, 72.8)	0.00	97.3 (94.6, 100)	0.00	N/A	N/A	N/A	N/A	1	0.930 (0.895, 0.965)	0.00
Between-group heterogeneity (*P* value)				.65		>.99		N/A		N/A			N/A	
FNIH														
Male	23 (14.4)	29 (18.1)	1	43.9 (22.8, 65.1)	0.00	97.1 (93.7, 101)	0.00	N/A	N/A	N/A	N/A	1	0.890 (0.840, 0.940)	0.00
Female	36 (16.1)	70 (31.3)	1	50.0 (32.9, 67.1)	0.00	94.2 (90.6, 97.8)	0.00	N/A	N/A	N/A	N/A	1	0.900 (0.855, 0.945)	0.00
Between-group heterogeneity (*P* value)				.66		.99		N/A		N/A			>.99	
IWGS														
Male	39 (24.4)	29 (18.1)	1	30.8 (15.5, 46.1)	0.00	N/A	N/A	N/A	N/A	N/A	N/A	1	0.800 (0.735, 0.865)	0.00
Female	57 (25.5)	70 (31.3)	1	50.9 (37.4, 64.5)	0.00	93.4 (89.3, 97.5)	0.00	N/A	N/A	N/A	N/A	1	0.860 (0.810, 0.910)	0.00
Between-group heterogeneity (*P* value)				.05		N/A		N/A		N/A			.90	

AUC, area under the curve; AWGS, Asian Working Group for Sarcopenia; CI, confidence interval; EWGSOP, European Working Group on Sarcopenia in Older People; FNIH, Foundation for the National Institutes of Health; IWGS, International Working Group on Sarcopenia; N/A, not applicable due to lack of data; NPV, negative predictive value; PPV, positive predictive value; Pt. est., point estimate.

*Number of articles.

Supplementary Table 7
Pooled Concurrent Validity Measures of the SARC-F Versions to Identify Sarcopenia, Grouped by Definition and Risk of Bias

Risk of Bias	Sarcopenia	SARC-F											
	Prevalence	Prevalence	N*	Sensitivity		Specificity		PPV		NPV		N*	AUC	I^2, %
	n (%)	*n* (%)		Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est., % (95% CI)	I^2, %	Pt. est, % (95% CI)	I^2, %		Pt. est., (95% CI)	
SARC-F														
EWGSOP														
High	493 (10.1)	368 (7.5)	4	41.4 (15.5, 67.3)	99.1	80.0 (71.0, 89.0)	98.4	28.9 (13.6, 44.2)	93.1	89.9 (84.6, 95.3)	52.8	0	N/A	N/A
Low	439 (13.3)	513 (15.5)	9	30.6 (12.7, 48.6)	87.9	82.4 (76.2, 88.5)	96.2	27.9 (14.7, 41.0)	92.0	87.4 (83.2, 91.6)	94.1	4	0.674 (0.534, 0.815)	94.6
Between-group heterogeneity (*P* value)				.50		.67		.92		.46			N/A	
EWGSOP2†														
High	218 (15.8)	455 (32.9)	3	55.2 (1.2, 109)	99.2	72.8 (47.5, 98.0)	97.5	29.0 (−0.9, 58.8)	98.8	85.4 (69.2, 102)	98.6	2	0.643 (0.532, 0.753)	50.8
Low	226 (9.8)	266 (11.5)	3	51.5 (−3.2, 106)	96.4	64.4 (38.9, 89.9)	98.5	20.6 (−3.8, 45.0)	90.7	91.0 (74.6, 107)	0.00	1	0.580 (0.367, 0.793)	0.00
Between-group heterogeneity (*P* value)				.93		.65		.67		.63			.65	
AWGS														
High	511 (8.3)	750 (12.2)	6	38.6 (22.2, 54.9)	97.2	80.0 (73.5, 86.6)	98.9	22.8 (12.8, 32.9)	94.8	91.5 (88.6, 94.3)	92.1	2	0.632 (0.442, 0.823)	82.1
Low	539 (13.7)	369 (9.4)	7	21.6 (6.6, 36.5)	90.6	92.4 (86.0, 98.9)	92.8	26.6 (14.6, 38.5)	80.1	88.2 (85.1, 91.3)	74.4	5	0.719 (0.598, 0.839)	95.4
Between-group heterogeneity (*P* value)				.13		.008		.64		.13			.46	
FNIH														
High	11 (3.6)	52 (17.0)	1	40.9 (35.4, 46.4)	0.00	84.9 (80.9, 88.9)	0.00	17.3 (13.1, 21.5)	0.00	98.8 (97.6, 100)	0.00	0	N/A	N/A
Low	377 (7.5)	765 (15.3)	7	27.2 (20.6, 33.7)	40.3	86.0 (79.9, 92.0)	97.1	14.7 (5.5, 24.0)	88.0	92.6 (88.2, 97.1)	95.4	3	0.772 (0.713, 0.830)	75.2
Between-group heterogeneity (*P* value)				.048		.89		.81		.22			N/A	
IWGS														
High	1002 (20.4)	365 (7.4)	4	36.8 (13.0, 60.7)	99.0	82.6 (75.1, 90.1)	97.6	37.9 (18.4, 57.5)	86.5	81.0 (73.1, 88.8)	94.4	0	N/A	N/A
Low	410 (19.0)	334 (15.4)	5	31.3 (8.6, 54.0)	79.5	92.0 (85.4, 98.6)	92.4	39.0 (16.1, 61.9)	97.7	87.7 (78.6, 96.8)	97.5	2	0.737 (0.657, 0.817)	86.1
Between-group heterogeneity (*P* value)				.74		.07		.94		.27			N/A	
SCWD														
High	18 (5.9)	52 (17.0)	1	55.6 (50.0, 61.2)	0.00	85.4 (81.4, 89.4)	0.00	19.2 (14.8, 23.6)	0.00	96.8 (94.8, 98.8)	0.00	0	N/A	N/A
Low	5 (2.4)	39 (18.8)	1	40.0 (−1.5, 81.5)	0.00	81.7 (76.4, 87.1)	0.00	5.10 (−1.8, 12.1)	0.00	98.2 (96.2, 100)	0.00	0	N/A	N/A
Between-group heterogeneity (*P* value)				.47		.99		.73		.33			N/A	
SARC-CalF														
EWGSOP														
Low	189 (16.8)	302 (26.8)	5	57.2 (43.2, 71.2)	69.6	87.9 (84.8, 91.1)	46.4	54.7 (−0.7, 110)	97.0	92.1 (82.1, 102)	81.2	5	0.810 (0.766, 0.854)	55.1
AWGS														
High	81 (26.2)	166 (53.7)	1	66.6 (55.9, 77.4)	0.00	70.1 (63.6, 76.7)	0.00	44.1 (35.0, 53.3)	0.00	85.5 (80.2, 90.9)	0.00	1	0.750 (0.700, 0.800)	0.00
Low	175 (22.4)	236 (30.2)	3	48.6 (28.7, 68.6)	84.2	93.4 (88.0, 98.8)	88.6	66.7 (35.2, 98.2)	N/A	86.8 (84.3, 89.3)	0.00	3	0.854 (0.766, 0.942)	88.1
Between-group heterogeneity (*P* value)				.36		<.001		.18		.99			.23	
FNIH														
Low	165 (19.0)	251 (28.9)	3	45.9 (19.0, 72.8)	92.1	90.5 (86.5, 94.6)	67.7	N/A	N/A	N/A	N/A	3	0.817 (0.716, 0.918)	93.8
IWGS														
Low	206 (23.7)	251 (28.9)	3	49.4 (39.2, 59.6)	38.1	91.3 (86.1, 96.4)	81.0	N/A	N/A	N/A	N/A	3	0.827 (0.801, 0.854)	0.00

AUC, area under the curve; AWGS, Asian Working Group for Sarcopenia; CI, confidence interval; EWGSOP, European Working Group on Sarcopenia in Older People; EWGSOP2, revised EWGSOP definition; FNIH, Foundation for the National Institutes of Health; IWGS, International Working Group on Sarcopenia; N/A, not applicable due to lack of data; NPV, negative predictive value; PPV, positive predictive value; Pt. est., point estimate; SCWD, Society on Sarcopenia, Cachexia and Wasting Disorders.

*Number of articles.

†Confirmed sarcopenia according to EWGSOP2.

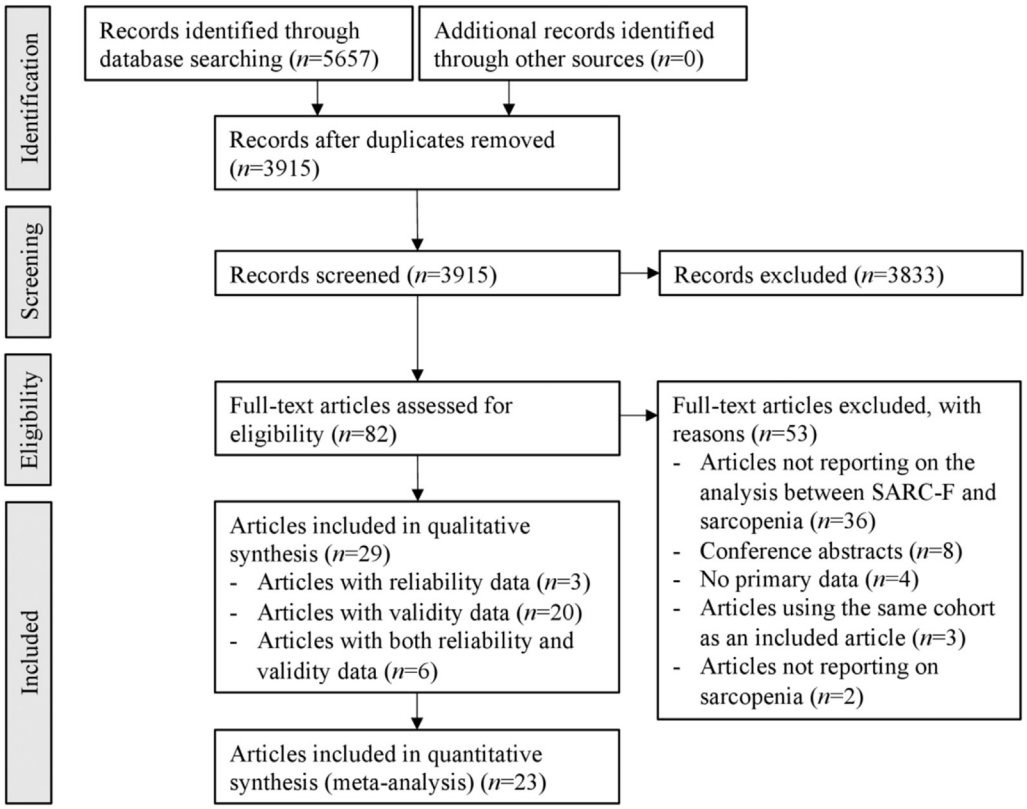
Supplementary Table 8
Publication Bias Assessment of Concurrent Validity Measures, Grouped by SARC-F Version and Definition

Sarcopenia Definition	N*	Egger Test *P* Values (2-Tailed)				
		Sensitivity	Specificity	PPV	NPV	AUC
SARC-F						
EWGSOP	13	.033	<.001	.26	.16	.35
EWGSOP2†	6	.67	.07	.015	.28	.21
AWGS	13	<.001	.009	.001	.049	.03
FNIH	7	.76	.49	.08	.08	.17
IWGS	9	.011	.014	.11	.84	.21
SCWD	2	N/A	N/A	N/A	N/A	N/A
SARC-CalF						
EWGSOP	5	.64	.34	N/A	N/A	.41
AWGS	4	.09	.043	N/A	N/A	.15
FNIH	3	.24	.14	N/A	N/A	.21
IWGS	3	.98	.03	N/A	N/A	.96

AUC, area under the curve; AWGS, Asian Working Group for Sarcopenia; EWGSOP, European Working Group on Sarcopenia in Older People; EWGSOP2, revised EWGSOP definition; FNIH, Foundation for the National Institutes of Health; IWGS, International Working Group on Sarcopenia; N/A, not applicable due to lack of data; NPV, negative predictive value; PPV, positive predictive value; SCWD, Society on Sarcopenia, Cachexia and Wasting Disorders.

*Number of articles.

†Confirmed sarcopenia according to EWGSOP2.



Supplementary Fig. 1. PRISMA flow diagram of the search.