ORIGINAL ARTICLE

Journal of Cachexia, Sarcopenia and Muscle 2022; **13**: 264–275
Published online 12 December 2021 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/jcsm.12871

Validation of a proxy-reported SARC-F questionnaire for current and retrospective screening of sarcopenia-related functional impairments

Johannes Maurus[1,2], Tobias Terzer[3], Axel Benner[3], Sabine Goisser[4], Annette Eidam[4], Anja Roth[4], Maike Janssen[1], Sonia Jaramillo[1], Hannes Martin Lorenz[1], William Micol[4], Klaus Hauer[4], Carsten Müller-Tidow[1], Jürgen M. Bauer[4,5], Karin Jordan[1] & Nina Rosa Neuendorff[1,6]*

[1]*Department of Medicine V – Haematology, Oncology and Rheumatology, University Hospital Heidelberg, Heidelberg, Germany,* [2]*Clinic for Urology, University Hospital Augsburg, Augsburg, Germany;* [3]*Division of Biostatistics, German Cancer Research Center (DKFZ), Heidelberg, Germany;* [4]*Center of Geriatric Medicine, Heidelberg University, AGAPLESION Bethanien Hospital Heidelberg, Heidelberg, Germany;* [5]*Network Aging Research (NAR), Heidelberg University, Heidelberg, Germany;* [6]*Clinic for Haematology and Stem-Cell Transplantation, University Hospital Essen, Essen, Germany*

Abstract

Background The strength, assistance walking, rise from a chair, climb stairs, and falls (SARC-F) questionnaire is a well-established instrument for screening of sarcopenia and sarcopenia-related functional impairments. As it is based on self-reporting, its use precludes patients who are unable to answer the questionnaire as a consequence of severe acute diseases or cognitive impairment. Therefore, we aimed to validate a proxy-reported version of the SARC-F for both ad-hoc as well as retrospective screening for severe sarcopenia-related functional impairments.

Methods Patients aged ≥60 years completed the SARC-F and performed the short physical performance battery (SPPB) at baseline (T1). Proxies in Cohort A gave a simultaneous assessment of the patients' functional status with the proxy-reported SARC-F at T1 and again, retrospectively, after 3 months (T2). Proxies in Cohort B only completed the SARC-F retrospectively at T2.

The questionnaires' performances were assessed through sensitivity/specificity analyses and receiver operating characteristic (ROC) curves. For non-inferiority analyses, results of both the patient-reported and proxy-reported SARC-F were correlated with the SPPB total score as well as the results of the chair-rise test subcategory; the respective correlation coefficients were tested against each other.

Results One hundred and four patients and 135 proxies participated. Using a SPPB score < 9 points as the reference standard, the proxy-reported SARC-F identified patients at high risk for sarcopenia-related functional impairment with a sensitivity of 0.81 (ad-hoc), 0.88 (retrospective Cohort A), and 0.87 (retrospective Cohort B) as well as a specificity of 0.89 (ad-hoc), 0.78 (retrospective Cohort A), and 0.64 (retrospective Cohort B). Areas under the ROC curves were ≥ 0.9 for the ad-hoc proxy-reported SARC-F and the retrospective proxy-reported SARC-F in both cohorts. The proxy-reported SARC-F showed a non-inferior correlation with the SPPB compared with the patient-reported SARC-F for ad-hoc (P = <0.001) as well as retrospective screening for severe sarcopenia-related functional impairment in both Cohorts A (P = 0.007) and B (P = 0.026).

Conclusions Proxy-reported SARC-F is a valid instrument for both ad-hoc as well as retrospective screening for sarcopenia-related functional impairment and could become the standard tool for evaluating this risk in older adults with severe acute disease, for example, in patients with quickly evolving haematological conditions.

Keywords Sarcopenia; Sarcopenia-related functional impairments; Patient-reported outcome; Proxy-reported outcome; Premorbid condition; SARC-F

Received: 25 January 2021; Revised: 20 October 2021; Accepted: 29 October 2021
*Correspondence to: Nina Rosa Neuendorff, Clinic for Haematology and Stem-Cell Transplantation, University Hospital Essen, Hufelandstrasse 55, D-45147 Essen, Germany. Phone: +49-201-72383783, Fax: +49-201-723-5961, Email: nina.neuendorff@uk-essen.de

Introduction

Sarcopenia is a progressive loss of muscle mass and function that is related to ageing.[1] It leads to functional impairment,[2,3] increased risk of falls and fractures,[4–7] loss of independence,[8,9] and an increased overall mortality.[10–13] Approximately 5–13% of adults aged 60–70 years are affected, increasing to 11–50% for those aged ≥80 years.[14] Thus, it constitutes a major public health issue.[15–18] Several adverse outcomes are associated with sarcopenia, such as an increased rate of post-operative complications after gastrointestinal surgery,[19] and a worse recovery during rehabilitation[20] and strokes.[20,21] Therefore, the diagnosis and implementation of sarcopenia into treatment plans gain more interest in a broad variety of medical specialties.

Several different definitions of sarcopenia have been propagated. We refer to the revised definition of the European Working Group on Sarcopenia in Older People,[1] where a diagnosis of probable sarcopenia is given in case of reduced muscle strength and further confirmed by the finding of reduced muscle quantity or quality.[1] To assess the severity of sarcopenia, testing physical performance is advised, for example, with the 'short physical performance battery' (SPPB).[1] SPPB test scores of <9 indicate a high risk for severe sarcopenia-related functional impairments.[1] For sarcopenia screening, the revised definition of the European Working Group on Sarcopenia in Older People recommends the use of the strength, assistance walking, rise from a chair, climb stairs, and falls (SARC-F) questionnaire.[1]

The SARC-F questionnaire contains five self-reported items evaluating the hallmarks of sarcopenia, that is, functional deficits and falls.[22] Each item has a possible score of 0 to 2 points, with higher scores suggesting a higher risk of sarcopenia.[22] A total score of >3 is regarded as a positive screening result.[22] Higher SARC-F scores have been shown to correlate with a slower chair rise, lower gait speed, overall lower SPPB scores, as well as adverse outcomes related to sarcopenia.[22] The SARC-F has a low-to-moderate sensitivity but high specificity,[23,24] thus providing a well-suited screening test to identify individuals who are not at high risk of sarcopenia-related negative outcomes. Its validation in many different languages eases its use in clinical practice.[23,25–30]

So far, the SARC-F questionnaire has only been validated for self-reporting. A person's self-reporting can be inaccurate due to cognitive impairment,[31] a non-objective perception of one's own functional capacities,[32] or even negation of functional decline. Furthermore, self-reporting is impossible in patients unable to answer questions due to an acute medical condition or altered consciousness.[33] As treatment decisions and modifications might benefit from knowledge about the presence of sarcopenia and sarcopenia-related functional impairments as part of a geriatric assessment,[34–37] proxy-reported information could be valuable to allow for patient-centred treatment strategies. If the illness has rapidly evolved within a few weeks, information on the patient's premorbid functional impairments might be crucial to differentiate between disease-related and premorbid conditions. The latter could reflect the patient's limited intrinsic physical resilience and guide the physicians towards adapted treatment intensities, such as an adjusted chemotherapy or radiotherapy regimen for the treatment of cancer. In contrast, the misperception of a disease-related deterioration as an intrinsic condition could potentially preclude the patient from a curative approach.

To our knowledge, no screening tool for sarcopenia-related functional impairments has been validated for the use of proxy-reported information so far. To close this gap, we aimed to validate a proxy-reported version of the SARC-F questionnaire for the evaluation of both a patient's current as well as retrospective, premorbid status.

Methods

A proxy-reported version of the original SARC-F questionnaire was developed substituting 'you' with the respective patient's name within the questionnaire. For retrospective evaluation, verbs were transformed from present into past tense. No further modifications were made.

Study population

The study population included patients aged ≥60 years and their proxies. Patients undergoing an in-house geriatric rehabilitation for diverse medical conditions at Agaplesion Bethanien Hospital (Heidelberg, Germany; Centre 1) as well as patients under surveillance for a rheumatological or haematological disease at the Heidelberg University Hospital outpatient clinic (Centre 2) were recruited. We selected these diverse centres with the clear intention of recruiting a representative sample of older adults with different levels of functional impairments in the hospital setting. Although patient characteristics within Cohorts A and B were further subdivided into Centres 1 and 2 to provide the detailed composition of these cohorts (e.g. *Table* 1), patients from Centres 1 and 2 did not represent independent groups and were not statistically separately analysed. Patients were excluded if they were unable to designate at least one close contact person (proxy), suffered from an acute medical condition that precluded them

DOI: 10.1002/jcsm.12871

Table 1 General patient and proxy characteristics

Characteristics	Cohort A			Cohort B		
	C1	C2	Total	C1	C2	Total
Patients						
Patients (*n*)	15	49	64	9	31	40
Age [years] (mean ± SD)	81.9 ± 8 (range 63–95)	76.6 ± 5.4 (range 64–90)	77.9 ± 6.5 (range 63–95)	83.6 ± 7.1 (range 72–93)	79.3 ± 5.1 (range 70–91)	80.2 ± 5.8 (range 70–93)
Sex (%)						
Female	73.3	40.8	48.4	55.6	38.7	42.5
SARC-F > 3 points (%)	73.3	26.5	37.5	77.8	16.1	30.0
SPPB < 9 points (%)	93.3	24.5	40.6	88.9	22.6	37.5
Gait speed < 0.8 m/s (%)	80.0	28.6	40.6	88.9	16.1	32.5
Number of participating proxies						
1	80%	65.3%	68.8%	88.9%	80.7%	82.5%
2	20%	26.5%	25%	11.1%	19.4%	17.5%
3	0%	8.2%	6.2%	0%	0%	0%
Proxies						
Age [years] (mean ± SD)	66.1 ± 13.4 (range 46–89)	67.3 ± 12.8 (range 34–87)	67 ± 12.8 (range 34–89)	59 ± 16.8 (range 25–82)	65.3 ± 13.4 (range 38–83)	63.9 ± 14.2 (range 25–83)
Sex (%)						
Female	66.7	69.4	68.8	66.7	83.9	80
Relation to patient						
Partner	26.7	59.2	51.6	22.2	61.3	52.5
Daughter	40.0	10.2	17.2	33.3	25.8	27.5
Son	13.3	18.4	17.2	22.2	9.7	12.5
Siblings	6.7	0	1.6	0	0	0
Brother-/sister-in-law	0	2.0	1.6	0	0	0
Niece/nephew	0	4.1	3.1	0	0	0
Friend	13.3	6.1	7.8	0	3.2	2.5
Grandchild	0	0	0	11.1	0	2.5
Professional caregiver	0	0	0	11.1	0	2.5
Geographical distance to patients						
Same address	33.3	69.4	60.9	11.1	67.7	55
Different town	66.7	30.6	39.1	55.6	22.6	30

C1, Centre 1 (Agaplesion Bethanien Hospital Heidelberg, geriatric hospital); C2, Centre 2 (University Hospital Heidelberg, rheumatology and haematology outpatient services).

from performing the SPPB or were unable to give informed consent (due to an at least moderate cognitive impairment or an acute alteration of consciousness).

The patient's designated proxies (caregivers, partners, children, siblings, grandchildren, neighbours, and friends) were also invited to participate. They were eligible if they were aged ≥18 years and had been in weekly contact with the patient during the past 6 months. Alternatively, contact on a weekly basis at least via telephone and with also meeting the patient in person at least twice during the past 6 months was accepted. Up to three proxies per patient were included. We classified proxies as main proxy (partner; if no partner was available, the proxy who was claimed as closest contact person by the patient) and additional proxies. We did not stratify patient and proxy recruitment according to gender to avoid any bias that could modify the choice of proxies as the person who knows the patient best. The trial flowchart is depicted in *Figure* 1.

Data collection and trial measurements

Patients and their respective proxies were recruited into two different cohorts in chronological order (Cohorts A and B). In both cohorts, the participating patients completed the self-reported SARC-F questionnaire and the SPPB once at baseline (T1). Proxies in Cohort A completed the proxy-reported SARC-F questionnaire twice, first at T1 (simultaneously to the patients' self-report) and then again after 3 months (T2). At T2, the proxies were asked to retrospectively evaluate the patients' functional status at T1. Proxies in Cohort B were not questioned at T1, but called for the first time at T2 to complete the proxy-reported SARC-F only retrospectively (i.e. evaluating the respective patient's functional status at T1). The Cohort B was established to examine a potential recall bias as proxies could potentially recall their first assessment of SARC-F questionnaire at T1 rather than assessing the premorbid condition of the patient when asked to assess the SARC-F retrospectively at T2. The presence of a recall bias would potentially limit the validity of a retrospective assessment. Study procedures are outlined in *Figure* 1.

All patients and proxies were fluent in German; nonetheless, they were offered to answer the SARC-F questionnaire in their respective mother tongue, which was preferred by three patients (once in Spanish and twice in Turkish).

The SPPB was performed as outlined elsewhere.[38] In brief, all patients carried out a balance test, performing a stand in side-by-side-position, semi-tandem, and full-tandem for 10 s

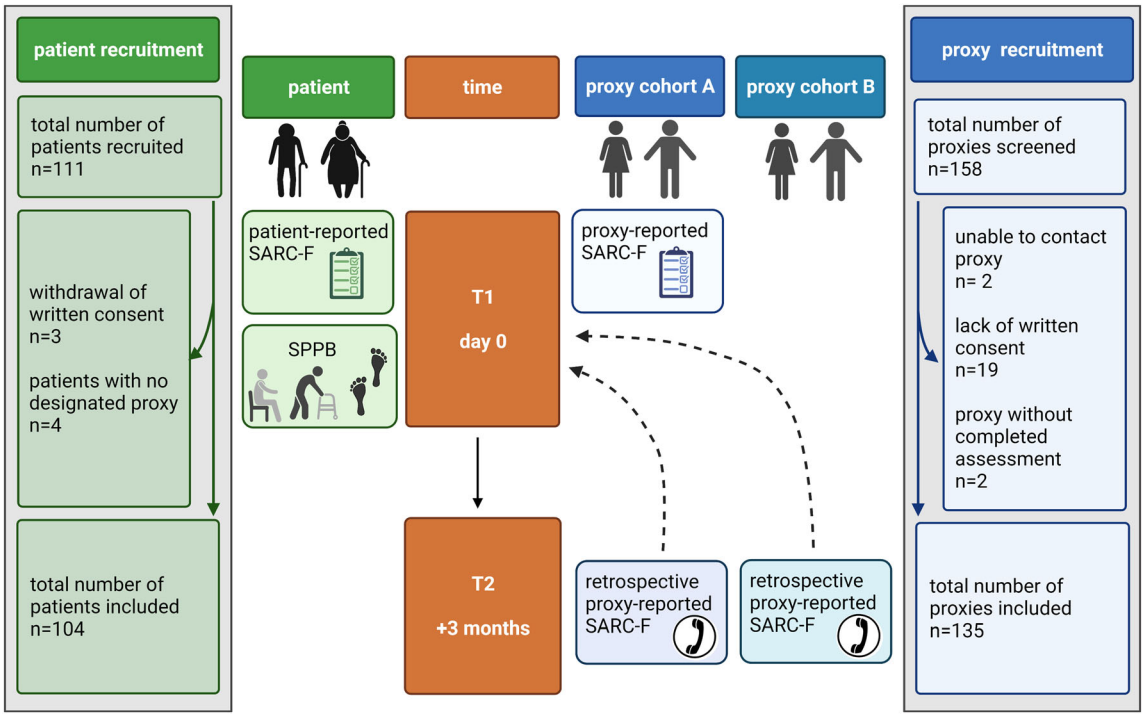

Figure 1 Trial procedures and recruitment flow chart. Patients performed SPPB and answered SARC-F questionnaire on Day 0 (=T1). Proxies in Cohort A answered SARC-F questionnaire on Day 0 and again retrospectively after 3 months (=T2). Proxies in Cohort B answered the questionnaire only once after 3 months (=T2) to test for a possible recall effect.

each. Thereafter, they were asked to walk a distance of 4 m in their usual gait speed twice; the faster walk was selected for analysis. Walking aids were allowed if used in daily routine. Finally, the patients performed the chair-rise test, getting up from a chair five times as quickly as possible without using the upper extremities or other assistance. The respective results of the three tests were each transformed into a score ranking from 0 to 4 points and summarized to yield an overall SPPB score ranking from 0 to 12 points, with higher scores indicating better performance.

Statistical analyses

Descriptive statistics were reported in absolute numbers and percentages for categorical variables. Possible differences in patient and proxy characteristics between the two study groups were assessed by Fisher's exact and Kruskal–Wallis test. All SARC-F scores were dichotomized into ≤ 3 (probably not at risk for sarcopenia-related functional impairments) and >3 points (positive screening for severe sarcopenia-related functional impairments). The SPPB score was dichotomized into ≥ 9 and <9 points.[1] The results of the chair-rise test were converted into chair rises per minute (CRPM) from the absolute time needed for five chair rises to allow for a natural numerical representation of a chair-rise test in cases where patients were not able to stand up five times in a row. If a patient was unable to perform this test, the result was set as infinite. Thereby, the introduction of artificial numbers for analysis was avoided.

The SARC-F questionnaire was primarily established to screen for persons who are at risk for sarcopenia-related poor functional outcomes.[22] Thus, defining a SPPB score <9 points as the cut-off value for patients at high risk for severe sarcopenia-related functional impairments, we assessed the diagnostic values [sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), and positive and negative likelihood ratios (PLR, NLR)] of the patient-reported and the different proxy-reported SARC-F screening modalities. We used receiver operating characteristic (ROC) curves to compare the overall accuracy of the patient-reported and proxy-reported SARC-F variants. The areas under the ROC curve (AUCs) were calculated. Comparisons between the AUC of the different SARC-F screening modalities were performed using the DeLong test.[39]

To assess the agreement between patient-reported and proxy-reported SARC-F results, Cohen's kappa was calculated and rated according to the strength of agreement.[40,41]

The results of the patient-reported and the different proxy-reported SARC-F versions were each correlated with the patient's SPPB scores and with their CRPM results. The latter were specifically analysed because SARC-F items mainly

focus on lower extremity strength, and may therefore be well reflected by the chair-rise test component of the SPPB.[1]

The resulting correlation coefficients were then tested against each other in pairwise comparisons for non-inferiority of proxy-reported in comparison with patient-reported SARC-F results to an alpha level of 0.05 with a non-inferiority margin of $\delta_0 = 0.2$.[42] Additionally, the 95% confidence intervals of the differences between these pairs of correlation coefficients were calculated. In an exploratory analysis, the obtained correlation coefficients were also tested against each other for superiority to an alpha level of 0.05.

Unless stated otherwise, the presented analyses were all performed for the group of the main proxies only. No composite score of different proxies was generated.

To compare the proxy-reported SARC-F results obtained from the different proxy subgroups (main proxies, additional proxies, and proxies with a residential distance), correlations of the retrospective proxy-reported SARC-F results with the CRPM were calculated. The 95% confidence intervals of the differences between the respective correlation coefficients were provided. No formal non-inferiority/superiority testing was performed due to the small sample sizes of the different proxy subgroups.

The sample size for the Cohort A was based on the feasibility to complete recruitment within less than 6 months; however, to ensure proper powering of the study, comprehensive sample size calculations based on different scenarios were conducted. The sample size of the Cohort B was calculated based on the correlation coefficients of the patient-reported SARC-F results with the CRPM of the first 47 patients in Cohort A, aiming at a significance level of $\alpha = 0.2$ to achieve an empirical power of 88% for non-inferiority.

All analyses were performed using the programming language 'R', 'cocor' for calculation of correlation coefficients, and 'pROC' for generation of ROC curves and AUC calculations.

Results

Study population

One hundred and eleven patients and 137 proxies were enrolled in this study; the final study population comprised 104 patients and 135 proxies (*Figure* 1). Of these, 64 patients and their corresponding proxies (64 main and 24 additional proxies) were included in Cohort A between September 2018 and January 2019. Fifteen of the patients were undergoing geriatric in-patient rehabilitation (Centre 1), and 49 were community-dwelling older adults having a scheduled outpatient visit at the rheumatology or haematology outpatient service (Centre 2). A further 40 patients and their proxies (40 main and 7 additional proxies) were assigned to Cohort B (9 at Centre 1 and 31 at Centre 2) between November 2018 and May 2019. In general, compliance of participants was excellent, with only two drop-outs due to incomplete proxy responses to the proxy-reported SARC-F.

Characteristics of the patients and proxies from the two cohorts are shown in *Table* 1 and did not reveal any significant major differences (data not shown). Patients and proxies were almost exclusively Caucasian, with only one Afro-American, one Hispanic, and a limited number of Turkish dyads. Of note, patients from Centres 1 and 2 represented different groups with regard to their functional status, as 93.3% of the patients from Centre 1 compared with only 24.5% of the patients from Centre 2 achieved a SPPB score <9 points and were therefore regarded as having a high risk for sarcopenia-related severe functional impairments. Moreover, the median age at Centre 1 was 81.9 years (range: 63–95 years) in comparison with a median age of 76.6 years (range: 64–90 years) at Centre 2.

In Cohort A, 37.5% of patients had a positive screening result for severe sarcopenia-related functional impairments according to their patient-reported SARC-F score (>3 points), and 40.6% scored <9 points on the SPPB. At T1, 40.6% of all proxies in Cohort A answered the SARC-F questionnaire with the result of a positive screening, and at T2, 50% of proxies evaluated the respective patient as scoring positive (>3 points).

In Cohort B, 30% of patients scored positive for high risk of severe sarcopenia-related functional impairments on the patient-reported SARC-F (>3 points), and 37.5% had a SPPB score <9 points. Fifty-five per cent of all proxies in Cohort B evaluated the respective patient as scoring positive (>3 points) for high risk of severe sarcopenia-related functional impairment at T1 when answering the SARC-F questionnaire at T2.

Performance of the patient-reported and proxy-reported SARC-F for ad-hoc sarcopenia screening

The sensitivity, specificity, PPV, NPV, PLR, and NLR of the patient-reported SARC-F in both cohorts and of the ad-hoc proxy-reported SARC-F in Cohort A are summarized in *Table* 2. The sensitivity of the different proxy-reported SARC-F modalities was constantly above 0.81, the NPV above 0.86. Overall, sensitivity, specificity, PPV, and NPV were in the range of a suitable screening test.[43]

In addition, ROC curves for the different SARC-F screening modalities and a SPPB score <9 points were generated (*Figure* 2). The AUCs were consistently high, in the range >0.8. The comparison of the ROC curves revealed no significant differences between the patient-reported and proxy-reported SARC-F versions (*Figure* 2F).

Table 2 Descriptive statistic measures of main proxy-reported and patient-reported SARC-F screening

	Sensitivity[a] (CI)	Specificity[a] (CI)	PPV (CI)	NPV (CI)	PLR (CI)	NLR (CI)
Ad-hoc patient-reported SARC-$F_{cohort\ A}$	0.73 (0.52, 0.88)	0.89 (0.74, 0.97)	0.83 (0.61, 0.95)	0.82 (0.66, 0.92)	6.58 (2.54, 17.06)	0.30 (0.16, 0.58)
Ad-hoc proxy-reported SARC-$F_{cohort\ A}$	0.81 (0.61, 0.93)	0.89 (0.74, 0.97)	0.84 (0.64, 0.95)	0.86 (0.71, 0.95)	7.27 (2.83, 18.66)	0.22 (0.10, 0.48)
Retrospective proxy-reported SARC-$F_{cohort\ A}$	0.88 (0.70, 0.98)	0.78 (0.61, 0.90)	0.74 (0.55, 0.88)	0.90 (0.74, 0.98)	3.98 (2.13, 7.45)	0.15 (0.05, 0.44)
Ad-hoc patient-reported SARC-$F_{cohort\ B}$	0.53 (0.27, 0.79)	0.84 (0.64, 0.95)	0.67 (0.35, 0.90)	0.75 (0.55, 0.89)	3.33 (1.21, 9.20)	0.56 (0.31, 0.98)
Retrospective proxy-reported SARC-$F_{cohort\ B}$	0.87 (0.60, 0.98)	0.64 (0.43, 0.82)	0.59 (0.36, 0.79)	0.89 (0.65, 0.99)	2.41 (1.38, 4.21)	0.21 (0.06, 0.78)

CI, 95% confidence interval; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value.
[a]SARC-F was regarded as positive with a score >3 points, and the reference to define a high risk for severe sarcopenia-related functional impairments was set at SPPB score <9 points.

Of note, the sensitivity and the PPV were markedly decreased for the ad-hoc patient-reported SARC-F in Cohort B when compared with those in Cohort A (*Table* 2). Moreover, the AUC for the ad-hoc patient-reported SARC-F in Cohort B was reduced to 0.82 in comparison with 0.92 in Cohort A (*Figure* 2), although no significant difference was demonstrated (P = 0.195). No obvious reasons for these differences could be identified.

In contrast to the originally proposed SARC-F cut-off of >3 points to define a positive screening result,[22] we observed the highest sum of sensitivity and specificity at a higher threshold of >4 points for ROC curves of all patient and proxy subgroups with the exception of the ad-hoc proxy-reported SARC-F in Cohort A (*Figure* 2). That indicates that the most suitable cut-off for a positive result on the SARC-F for severe impairment would be >4 points for our data.

To assess the interrater reliability between patient and proxy ratings, Cohen's kappa was calculated for agreement between patient-reported and proxy-reported ad-hoc SARC-F > 3 points. Agreement was demonstrated to be substantial [κ = 0.79; CI (0.64, 0.95)].

Correlations of the ad-hoc proxy-reported SARC-F results with the SPPB as well as those of the patient-reported SARC-F with the SPPB in both cohorts were moderate (*Table* 3). When tested against each other, non-inferiority of the ad-hoc proxy-reported SARC-F/dichotomized SPPB correlation coefficient compared with the patient-reported SARC-F/dichotomized SPPB correlation coefficient could be demonstrated (*Table* 3). The same was shown for correlations of the ad-hoc proxy-reported and patient-reported SARC-F results with the CRPM and revealed non-inferiority of the ad-hoc proxy-reported SARC-F (*Table* 3).

Additional testing for a potential superiority of the ad-hoc proxy-reported SARC-F/CRPM correlation coefficient compared with the patient-reported SARC-F/CRPM correlation coefficient was negative [P = 0.05; 95% CI for the difference of the correlation coefficients $r_1 - r_2$: (−0.033, 0.376)].

In summary, ad-hoc proxy-reported SARC-F was shown to be non-inferior to the patient-reported evaluation.

Performance of the retrospective proxy-reported SARC-F and evaluation of a recall effect

The sensitivity, specificity, PPV, NPV, PLR, and NLR of retrospective proxy-reported SARC-F in Cohort A are summarized in *Table* 2. The sensitivity was 0.88 [95% CI (0.70, 0.98)], specificity 0.78 [95% CI (0.61, 0.90)], PPV 0.74 [95% CI (0.55, 0.88)], and NPV 0.9 [95% CI (0.74, 0.98)]; thus, performance was comparable with that of the ad-hoc proxy-reported SARC-F.

Again, ROC curves for the different SARC-F screening modalities and a SPPB score < 9 points were generated (*Figure*

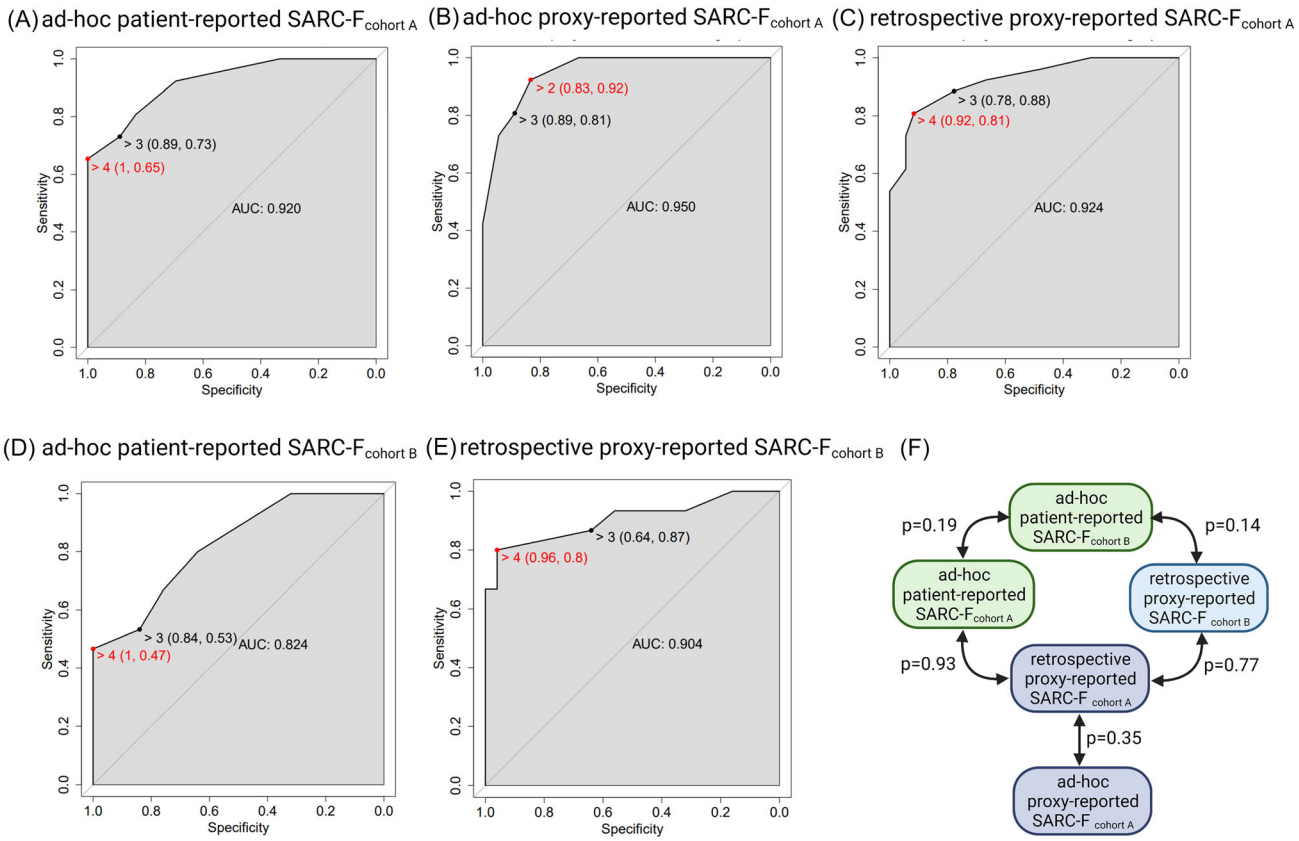

Figure 2 Receiver operating characteristic (ROC) curves. ROC curves were calculated defining a 'short physical performance battery' (SPPB) score <9 points as having a high risk of severe sarcopenia-related functional impairment for (A) ad-hoc patient-reported SARC-F (Cohort A), (B) ad-hoc proxy-reported SARC-F (Cohort A), (C) retrospective proxy-reported SARC-F (Cohort A), (D) ad-hoc patient-reported SARC-F (Cohort B), and (E) retrospective proxy-reported SARC-F (Cohort B). (F) *P* values for DeLong test comparing areas under the curves (AUC) for the different ROC curves are shown, revealing no significant differences.

Table 3 Summary of non-inferiority analyses

	CC SARC-F > 3/SPPB < 9	*P* value	95% CI ($r_1 - r_2$)	CC SARC-F > 3/chair rise	*P* value	95% CI ($r_1 - r_2$)
Ad-hoc patient-reported SARC-F (Cohort A) (r_1)	0.63	0	[−0.3311, 0.0865]	−0.59	0	[−0.0325, 0.3534]
Ad-hoc proxy-reported SARC-F (Cohort A) (r_2)	0.70			−0.66		
Ad-hoc patient-reported SARC-F (Cohort A) (r_1)	0.63	0.0065	[−0.3014, 0.2306]	−0.59	0.0011	[−0.1385, 0.3894]
Retrospective proxy-reported SARC-F (Cohort A) (r_2)	0.65			−0.67		
Ad-hoc patient-reported SARC-F (Cohort B) (r_1)	0.39	0.026	[−0.4774, 0.2313]	−0.6	0.268	[−0.5274, 0.2117]
Retrospective proxy-reported SARC-F (Cohort B) (r_2)	0.49			−0.49		
Ad-hoc patient-reported SARC-F (Cohort A) (r_1)	0.63	0.0001	[−0.0555, 0.5651]	−0.59	0.0925	[−0.2756, 0.2603]
Ad-hoc patient-reported SARC-F (Cohort B) (r_2)	0.39			−0.6		
Retrospective proxy-reported SARC-F (Cohort A) (r_1)	0.65	0.0003	[−0.1058, 0.4652]	−0.67	0.4476	[−0.482, 0.0851]
Retrospective proxy-reported SARC-F (Cohort B) (r_2)	0.49			−0.49		

CC, correlation coefficient; r_1, correlation coefficient 1; r_2, correlation coefficient 2.

2) and revealed an AUC of 0.92 (*Figure* 2F) for retrospective proxy-reported SARC-F in Cohort A.

In Cohort B, in which proxies were only asked once after 3 months to answer the proxy-reported SARC-F questionnaire to exclude a recall bias, measures for sensitivity and NPV were within the same range as those of Cohort A (*Table* 2); the specificity and PPV were reduced (*Table* 2) due to a higher rate of false positive measurements in Cohort B. Regarding ROC curves, the AUC of retrospective proxy-reported SARC-F in Cohort B was as high as 0.90 (refer to *Figure* 2).

In addition, Cohen's kappa was calculated for agreement between ad-hoc patient-reported SARC-F $>$ 3 points and retrospective proxy-reported SARC-F $>$ 3 points. Agreement was demonstrated to be substantial [κ = 0.61; CI (0.42, 0.81)] in Cohort A and moderate in Cohort B [κ = 0.42; CI (0.15, 0.69)].

In Cohort A, the retrospective proxy-reported SARC-F/dichotomized SPPB correlation was non-inferior to the patient-reported SARC-F/dichotomized SPPB correlation (*Table* 3). Similar results were obtained for the correlations of SARC-F with CRPM (*Table* 3).

In line, non-inferiority could be demonstrated for the retrospective proxy-reported SARC-F/dichotomized SPPB correlation in comparison with the patient-reported SARC-F/dichotomized SPPB in Cohort B (*Table* 3). In contrast, non-inferiority testing of the retrospective proxy-reported SARC-F/CRPM correlation versus the patient-reported SARC-F/CRPM correlation failed in Cohort B (*Table* 3). Further analyses of Cohort B revealed that the correlations differed mostly in the subgroup of proxies whose patients were treated at the outpatient rheumatology service (subgroup at Centre 2; Supporting Information, *Table* S1). No obvious reasons for this difference could be found.

Proxy-reported SARC-F: main proxies in comparison with additional proxies

As stated above, the previous analyses were all performed within the group of main proxies. For 16 and 4 patients, one or two additional proxies were enclosed, respectively. In total, 51.9% of all proxies were partners, and the majority of the remaining proxies were the patients' children (*Table* 1). Only 41.4% of all proxies lived in a different town. The 95% confidence intervals for the differences between the correlation coefficients of the retrospective proxy-reported SARC-F and the CRPM of the respective subgroups are shown in *Table* 4. The correlation coefficients for the groups with a less close relationship (in comparison with partners) or having a different address/living in a different town (compared with sharing the same address with the patient) showed a trend towards higher correlation coefficients.

Because the subgroups were too small to formally analyse interrater reliability, we assessed the subgroup of 16 patients with two available proxies. In these, only 3/16 proxies [18.75%; 95% CI (0.04, 0.46)] answered discordantly to an extent that resulted in a positive versus a negative screening result (score of $\leq$3 vs. $>$3 points).

Discussion

In the health care of older adults, complementing or substituting the patient's self-report with the appraisal of a close contact person or caregiver is an occasionally used practice. This applies particularly to cases in which the use of patient-reported outcomes or functional tests is not feasible due to acute medical conditions or dementia. Especially with regard to the stratification of treatment intensity, assessment of severe sarcopenia-related functional impairments gives valuable information on the patient's fitness and therefore on resilience factors for treatment tolerability and rehabilitation potential. In our study, we evaluated the validity of the SARC-F questionnaire for proxy-reported ad-hoc and retrospective screening for severe sarcopenia-related functional impairments. To our knowledge, this is the first study examining the validity of a screening questionnaire for sarcopenia-related functional impairments as a proxy-assisted tool.

Evaluation of a patient's current functional status with the proxy-reported SARC-F was demonstrated to be non-inferior to the patients' self-reporting and showed promising validity when compared with functional tests indicating a high risk of

Table 4 Evaluation of different proxy characteristics

Proxy subgroup	Correlation coefficient[a]	95% Confidence interval ($r_1 - r_2$)
Relationship		
Partner (n = 31; r_1) versus	−0.6	[−0.1822, 0.435]
Remaining relationship groups (n = 31; r_2)	−0.715	
Postcode		
Same town (n = 47; r_1) versus	−0.657	[−0.3431, 0.3361]
Different town (n = 15; r_2)	−0.724	
Address		
Identical address (n = 37; r_1)	−0.63	[−0.1877, 0.3935]
Different address (n = 25; r_2)	−0.741	

[a]Correlation coefficients between retrospective proxy-reported SARC-F and chair rises per minute (CRPM).

severe sarcopenia-related functional deficits (i.e. SPPB). Thus, we suggest its use as a screening tool for severe sarcopenia-related functional deficits in cases where self-reporting is not feasible.

The retrospective evaluation with the proxy-reported SARC-F also showed promising results with robust identification rates of patients at high risk for severe sarcopenia-related functional deficits. Although non-inferiority could not be demonstrated for the proxy-reported SARC-F/CRPM correlation compared with the patient-reported SARC-F/CRPM correlation in Cohort B, the respective correlation with the overall SPPB score proved to be non-inferior in both cohorts. Considering the latter findings as well as the analysis of the ROC curves and the moderate interrater reliability between patient and proxy ratings, the results of the retrospective proxy-reported SARC-F screening in Cohort A are unlikely to be severely compromised by a recall effect. Moreover, the high NPVs suggest that the retrospective screening is particularly suitable to exclude individuals who have not been at high risk for severe sarcopenia-related functional deficits.

Although evaluation of premorbid function is not always useful, it can be beneficial if an acute medical condition is thought to cause a severe temporary impairment rather than representing the patient's intrinsic condition. In this case, an assessment of the premorbid function can provide such information. Under this assumption, retrospective proxy-reported SARC-F screening can be used to guide therapy decisions if discrimination between disease-related functional decline and premorbid impairment is crucial. To further depict such a case, the rapid onset of acute myeloid leukaemia can serve as an example. Even in a previously fit and healthy older adult, the onset of acute myeloid leukaemia can lead to a rapid deterioration of the functional abilities that is mainly disease-related and does not represent the underlying physical resilience of the patient. Treatment options range between intensive chemotherapies eventually followed by an allogeneic stem-cell transplantation aiming at cure of the disease and better-tolerated palliative treatments. The choice of therapy path is crucial to avoid the high probability of treatment-related morbidity and mortality in case of a frail patient receiving an intensive treatment and undertreatment in a patient with disease-related functional impairments and high potential for physical resilience receiving a palliative treatment option.

The patient's self-perception of functional capacities can differ significantly from functional testing results,[44,45] with underestimation[46] and overestimation [32,46,47] of abilities. This perception of one's own, but also of others' (functional) abilities is a multifaceted process, influenced by many aspects, such as socio-economic factors and gender.[48–51] Functional decline might not be perceived as such if deficits are well compensated by the social environment, assistive devices, and adaptation of lifestyle to deficits. Thus, the self-reported functional status might be overrated. On the other hand, psychological factors, such as social isolation, or depression, can lead to 'feeling' physically impaired although being objectively functional. In these cases, proxy-reported functional assessments can provide important hints. On the other hand, proxy perception can also be affected by experiencing a high psychological and socio-economic burden as a caregiver.[52] Accordingly, proxy ratings might variably be superior or inferior to an objective test result or the patient's own perspective.[53,54] Our study was not sufficiently powered to assess in detail the superiority of the main proxy-reported SARC-F compared with patient-reported SARC-F results. Still, we observed a trend towards higher correlation coefficients within the proxy groups with a less close relationship to the patients (in comparison with partners). This probably indicates that the actual residential distance and a personal relationship outside the partnership are not disadvantageous for the validity of the obtained results, although this conclusion is limited by the small sample sizes, broad confidence intervals, and the exploratory nature of the analyses. Based on these findings, to contact also more distant proxies in cases where the patient cannot be assessed personally seems legitimate. Whether these proxies might respond even more objectively than the patients themselves or close proxies needs to be confirmed in a larger data set and could have important implications for day-to-day clinical decision-making.

In general, the specificity of the patient-reported SARC-F in this study was within the reported range from several other published SARC-F validation studies.[23,24,26,27,30] Compared with other pivotal trials, in which sensitivity values of $<$60% were described,[23,26,27,29,30] the sensitivity in our study was unexpectedly high. However, most of these other trials used different functional parameters for comparison, such as gait speed, grip strength, or composite scores of muscle strength and functional measures. The chair-rise test included in the SPPB employed in our trial is a very sensitive measure of functional deficits, possibly explaining the high sensitivity found in the present study. In addition, the prevalence of (sarcopenia-related) functional impairments was higher in our study than in the others, further explaining the higher sensitivity. In line, detection of severe impairments could be more obvious and therefore easier to rate for proxies than less severe impairments, which could additionally explain the higher sensitivity.

Limitations

The present study has several limitations. First, the sample size was only moderate in comparison with the pivotal SARC-F validation trial.[22] In contrast to the initial large population-based study[22] that aimed mainly at assessing the questionnaire's construct validity, we intended to validate

a proxy-reported version for the hospital setting. These different purposes explain the different sample sizes. Moreover, validity of the presented results can be claimed because statistical significance was achieved. In addition, stronger effects are required to reach statistical significance in smaller cohorts in comparison with larger trial populations. The latter further underlines the validity of our findings. Second, the mean patient age was nearly 80 years. Whether the proxy ratings in younger patient populations are equally valid was not determined. Third, patients with acute severe diseases were excluded in this trial because neither a functional reference nor a patient-reported SARC-F score would have been available as comparator for the proxy-rated scores. It cannot be excluded that the acute distress of such a situation could potentially bias the validity of the proxy-rating. Fourth, the patient cohorts were very heterogeneous as we recruited participants from two distinct inpatient and outpatient settings with the intention to include patients with and without sarcopenia-related severe functional deficits. Fifth, the group of proxies showed a clear female dominance, mostly reflecting the higher life expectancy of women and complex social patterns. A formal comparison of proxy validity with regard to gender was not performed. Finally, our participants were almost exclusively Caucasian. Because the SARC-F questionnaire is already validated in many ethnicities and languages,[22,25–27] we did not further focus on these aspects. Even more important than different ethnicities might be a culturally dominated difference of family structures with the lack of large families living with several generations in close proximity in our study. We cannot speculate on the impact of these different social structures on proxy validity as our trial was not powered to compare the validity of different proxy subgroups and these large families are scarce in Germany. These limitations should be taken into account when the proxy-reported SARC-F is used, and more research is required to assess the impact of different proxy subgroups and characteristics.

Conclusions

Under consideration of the above discussed limitations, we suggest using the proxy-reported SARC-F for the screening of a patient's current functional status in case an older patient is incapable of self-assessment. Thus, the proxy-reported SARC-F is the first instrument available for screening of severe sarcopenia-related functional impairment during acute medical emergencies and for patients with severe cognitive impairment. Adequate treatment of severe subacute diseases requires information on the patient's premorbid functional status to estimate resilience. Retrospective evaluation with the proxy-reported SARC-F could be used in this context to identify older patients without previous severe sarcopenia-related functional impairments.

Acknowledgements

We would like to thank the patients and their caregivers for participation in our study and Michael Drey (Munich) for providing the German translation of the SARC-F questionnaire. We are grateful to our colleagues Sabine Kayser, Nora Liebers, Lioba Pitierse, Johanna Schwaan, Lukas John, and Evgenija De Luca for their support in patient recruitment. The figures were created with BioRender.com. The authors of this manuscript certify that they comply with the guidelines of ethical authorship and publishing in the *Journal of Cachexia, Sarcopenia and Muscle*.[55]

Funding

N.R.N. was supported by a scholarship of Heidelberg University. No further funding was applied to this study.

Online supplementary material

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

Table S1. Subgroup analysis of center 1.

Conflicts of interest

None.

References

1. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyere O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. *Age Ageing* 2019;**48**:16–31.
2. Janssen I, Heymsfield SB, 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.
3. Visser M, Goodpaster BH, Kritchevsky SB, Newman AB, Nevitt M, Rubin SM, et al.

Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. *J Gerontol A Biol Sci Med Sci* 2005;**60**:324–333.
4. Szulc P, Feyt C, Chapurlat R. High risk of fall, poor physical function, and low grip strength in men with fracture—the STRAMBO study. *J Cachexia Sarcopenia Muscle* 2016;**7**:299–311.
5. Gadelha AB, Vainshelboim B, Ferreira AP, Neri SGR, Bottaro M, Lima RM. Stages of sarcopenia and the incidence of falls in older women: a prospective study. *Arch Gerontol Geriatr* 2018;**79**:151–157.
6. Schaap LA, van Schoor NM, Lips P, Visser M. Associations of sarcopenia definitions, and their components, with the incidence of recurrent falling and fractures: the longitudinal aging study Amsterdam. *J Gerontol A Biol Sci Med Sci* 2018;**73**:1199–1204.
7. Bischoff-Ferrari HA, Orav JE, Kanis JA, Rizzoli R, Schlogl M, Staehelin HB, et al. Comparative performance of current definitions of sarcopenia against the prospective incidence of falls among community-dwelling seniors age 65 and older. *Osteoporos Int* 2015;**26**:2793–2802.
8. Landi F, Calvani R, Ortolani E, Salini S, Martone AM, Santoro L, et al. The association between sarcopenia and functional outcomes among older patients with hip fracture undergoing in-hospital rehabilitation. *Osteoporos Int* 2017;**28**:1569–1576.
9. Dos Santos L, Cyrino ES, Antunes M, Santos DA, Sardinha LB. Sarcopenia and physical independence in older adults: the independent and synergic role of muscle mass and muscle function. *J Cachexia Sarcopenia Muscle* 2017;**8**:245–250.
10. Yang M, Jiang J, Zeng Y, Tang H. Sarcopenia for predicting mortality among elderly nursing home residents: SARC-F versus SARC-CalF. *Medicine (Baltimore)* 2019;**98**:e14546, https://doi.org/10.1097/MD.0000000000014546
11. Costanzo L, De Vincentis A, Di Iorio A, Bandinelli S, Ferrucci L, Antonelli Incalzi R, et al. Impact of low muscle mass and low muscle strength according to EWGSOP2 and EWGSOP1 in community-dwelling older people. *J Gerontol A Biol Sci Med Sci* 2020;**75**:1324–1330.
12. De Buyser SL, Petrovic M, Taes YE, Toye KR, Kaufman JM, Lapauw B, et al. Validation of the FNIH sarcopenia criteria and SOF frailty index as predictors of long-term mortality in ambulatory older men. *Age Ageing* 2016;**45**:602–608.
13. Bianchi L, Maietti E, Abete P, Bellelli G, Bo M, Cherubini A, et al. Comparing EWGSOP2 and FNIH sarcopenia definitions: agreement and 3-year survival prognostic value in older hospitalized adults: the GLISTEN study. *J Gerontol A Biol Sci Med Sci* 2020;**75**:1331–1337.
14. von Haehling S, Morley JE, Anker SD. An overview of sarcopenia: facts and numbers on prevalence and clinical impact. *J Cachexia Sarcopenia Muscle* 2010;**1**:129–133.
15. Bruyere O, Beaudart C, Ethgen O, Reginster JY, Locquet M. The health economics burden of sarcopenia: a systematic review. *Maturitas* 2019;**119**:61–69.
16. Mijnarends DM, Luiking YC, Halfens RJG, Evers S, Lenaerts ELA, Verlaan S, et al. Muscle, health and costs: a glance at their relationship. *J Nutr Health Aging* 2018;**22**:766–773.
17. Beaudart C, Rizzoli R, Bruyere O, Reginster JY, Biver E. Sarcopenia: burden and challenges for public health. *Arch Public Health* 2014;**72**:45.
18. Sousa AS, Guerra RS, Fonseca I, Pichel F, Ferreira S, Amaral TF. Financial impact of sarcopenia on hospitalization costs. *Eur J Clin Nutr* 2016;**70**:1046–1051.
19. Pipek LZ, Baptista CG, Nascimento RFV, Taba JV, Suzuki MO, do Nascimento FS, et al. The impact of properly diagnosed sarcopenia on postoperative outcomes after gastrointestinal surgery: a systematic review and meta-analysis. *PLoS ONE* 2020;**15**:e0237740. https://doi.org/10.1371/journal.pone.0237740
20. Yoshimura Y, Wakabayashi H, Bise T, Nagano F, Shimazu S, Shiraishi A, et al. Sarcopenia is associated with worse recovery of physical function and dysphagia and a lower rate of home discharge in Japanese hospitalized adults undergoing convalescent rehabilitation. *Nutrition* 2019;**61**:111–118.
21. Matsushita T, Nishioka S, Taguchi S, Yamanouchi A. Sarcopenia as a predictor of activities of daily living capability in stroke patients undergoing rehabilitation. *Geriatr Gerontol Int* 2019;**19**:1124–1128.
22. Malmstrom TK, Miller DK, Simonsick EM, Ferrucci L, Morley JE. SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes. *J Cachexia Sarcopenia Muscle* 2016;**7**:28–36.
23. Bahat G, Yilmaz O, Kilic C, Oren MM, Karan MA. Performance of SARC-F in regard to sarcopenia definitions, muscle mass and functional measures. *J Nutr Health Aging* 2018;**22**:898–903.
24. 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.
25. Bahat G, Yilmaz O, Oren MM, Karan MA, Reginster JY, Bruyère O, et al. Cross-cultural adaptation and validation of the SARC-F to assess sarcopenia: methodological report from European Union Geriatric Medicine Society Sarcopenia Special Interest Group. *Eur Geriatr Med* 2018;**9**:23–28.
26. Parra-Rodriguez L, Szlejf C, Garcia-Gonzalez AI, Malmstrom TK, Cruz-Arenas E, Rosas-Carrasco O. 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.
27. 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**:40–5.e1.
28. Kemmler W, Sieber C, Freiberger E, von Stengel S. The SARC-F questionnaire: diagnostic overlap with established sarcopenia definitions in older German men with sarcopenia. *Gerontology* 2017;**63**:411–416.
29. Ida S, Murata K, Nakadachi D, Ishihara Y, Imataka K, Uchida A, 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.
30. Barbosa-Silva TG, Menezes AM, Bielemann RM, Malmstrom TK, Gonzalez MC, Grupo de Estudos em Composicao Corporal e N. Enhancing SARC-F. Improving sarcopenia screening in the clinical practice. *J Am Med Dir Assoc* 2016;**17**:1136–1141.
31. Onor ML, Trevisiol M, Negro C, Aguglia E. Different perception of cognitive impairment, behavioral disturbances, and functional disabilities between persons with mild cognitive impairment and mild Alzheimer's disease and their caregivers. *Am J Alzheimers Dis Other Demen* 2006;**21**:333–338.
32. Roedersheimer KM, Pereira GF, Jones CW, Braz VA, Mangipudi SA, Platts-Mills TF. Self-reported versus performance-based assessments of a simple mobility task among older adults in the emergency department. *Ann Emerg Med* 2016;**67**:151–156.
33. Raveau T, Annweiler C, Chudeau N, Gergaud S, Thiery S, Gautier J, et al. Comprehensive geriatric assessment in intensive care unit: a pilot study (preSeniorea). *Geriatr Psychol Neuropsychiatr Vieil* 2013;**11**:389–395.
34. Schulkes KJ, Souwer ET, Hamaker ME, Codrington H, van der Sar-van der Brugge S, Lammers JJ, et al. The effect of a geriatric assessment on treatment decisions for patients with lung cancer. *Lung* 2017;**195**:225–231.
35. Xiao DY, Luo S, O'Brian K, Ganti A, Riedell P, Sanfilippo KM, et al. Impact of sarcopenia on treatment tolerance in United States veterans with diffuse large B-cell lymphoma treated with CHOP-based chemotherapy. *Am J Hematol* 2016;**91**:1002–1007.
36. Mohile SG, Mohamed MR, Xu H, Culakova E, Loh KP, Magnuson A, et al. Evaluation of geriatric assessment and management on the toxic effects of cancer treatment (GAP70+): a cluster-randomised study. *Lancet* 2021;**398**(10314):1894–1904. https://doi.org/10.1016/S0140-6736(21)01789-X
37. Soo W-K, King M, Pope A, Parente P, Darzins P, Davis ID. Integrated geriatric assessment and treatment (INTEGERATE) in older people with cancer planned for systemic anticancer therapy. *J Clin Oncol* 2020;**38**:12011. https://doi.org/10.1200/jco.2020.38.15_suppl.12011
38. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone com-

pared with the short physical performance battery. *J Gerontol A Biol Sci Med Sci* 2000; **55**:M221–M231.
39. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**: 837–845.
40. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall; 1991.
41. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–174.
42. Meng X-L, Rosenthal RRD. Comparing correlated correlation coefficients. *Psychol Bull* 1992;**111**:172–175.
43. Trevethan R. Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice. *Front Public Health* 2017;**5**:307, https://doi.org/10.3389/fpubh.2017.00307
44. Nielsen LM, Kirkegaard H, Ostergaard LG, Bovbjerg K, Breinholt K, Maribo T. Comparison of self-reported and performance-based measures of functional ability in elderly patients in an emergency department: implications for selection of clinical outcome measures. *BMC Geriatr* 2016;**16**:199.
45. Bean JF, Olveczky DD, Kiely DK, LaRose SI, Jette AM. Performance-based versus patient-reported physical function: what are the underlying predictors? *Phys Ther* 2011;**91**:1804–1811.
46. Kok AAL, Henstra MJ, van der Velde N, Rhebergen D, van Schoor NM. Psychosocial and health-related factors associated with discordance between 13-year trajectories of self-reported functional limitations and performance-based physical functioning in old age. *J Aging Health* 2020;**32**: 1084–1097.
47. Sager MA, Dunham NC, Schwantes A, Mecum L, Halverson K, Harlowe D. Measurement of activities of daily living in hospitalized elderly: a comparison of self-report and performance-based methods. *J Am Geriatr Soc* 1992;**40**: 457–462.
48. Tetteh J, Kogi R, Yawson AO, Mensah G, Biritwum R, Yawson AE. Effect of self-rated health status on functioning difficulties among older adults in Ghana: coarsened exact matching method of analysis of the World Health Organization's study on global AGEing and adult health, Wave 2. *PLoS ONE* 2019;**14**:e0224327, https://doi.org/10.1371/journal.pone.0224327
49. Gyasi RM, Phillips DR. Gender, self-rated health and functional decline among community-dwelling older adults. *Arch Gerontol Geriatr* 2018;**77**:174–183.
50. Jung MS, Lee KS, Kim M, Yun H. Gender-specific relationship between executive function and self-rated health. *Osong Public Health Res Perspect* 2019;**10**:93–101.
51. Arnadottir SA, Gunnarsdottir ED, Stenlund H, Lundin-Olsson L. Determinants of self-rated health in old age: a population-based, cross-sectional study using the International Classification of Functioning. *BMC Public Health* 2011;**11**:670.
52. Hsu T, Loscalzo M, Ramani R, Forman S, Popplewell L, Clark K, et al. Are disagreements in caregiver and patient assessment of patient health associated with increased caregiver burden in caregivers of older adults with cancer? *Oncologist* 2017;**22**: 1383–1391.
53. Elam JT, Graney MJ, Beaver T, el Derwi D, Applegate WB, Miller ST. Comparison of subjective ratings of function with observed functional ability of frail older persons. *Am J Public Health* 1991;**81**: 1127–1130.
54. Doble SE, Fisk JD, Rockwood K. Assessing the ADL functioning of persons with Alzheimer's disease: comparison of family informants' ratings and performance-based assessment findings. *Int Psychogeriatr* 1999;**11**:399–409.
55. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2019. *J Cachexia Sarcopenia Muscle* 2019;**10**:1143–1145.