

SARC-F VALIDATION AND SARC-F+EBM DERIVATION IN MUSCULOSKELETAL DISEASE: THE SPSS-OK STUDY

N. KURITA^{1,2,3}, T. WAKITA⁴, T. KAMITANI⁵, O. WADA⁶, K. MIZUNO⁶

1. Department of Clinical Epidemiology, Graduate School of Medicine, Fukushima Medical University, Fukushima City, Fukushima, Japan; 2. Department of Innovative Research and Education for Clinicians and Trainees (DiRECT), Fukushima Medical University Hospital, Fukushima City, Fukushima, Japan; 3. Center for Innovative Research for Communities and Clinical Excellence (CiRC2LE), Fukushima Medical University, Fukushima City, Fukushima, Japan; 4. Department of Sociology, Kansai University, Suita City, Osaka, Japan; 5. Department of Healthcare Epidemiology, School of Public Health in the Graduate School of Medicine, Kyoto University, Kyoto City, Kyoto, Japan; 6. Anshin Hospital, Kobe City, Hyogo, Japan. Corresponding author: Noriaki Kurita, MD, PhD, FACP, Department of Clinical Epidemiology, Graduate School of Medicine, Fukushima Medical University, Fukushima, Japan, 1 Hikarigaoka, Fukushima City, Fukushima 960-1295, Japan, Phone: +81-24-547-1471; Fax: +81-24-547-1468, E-mail: kuritanoriaki@gmail.com

Abstract: *Objectives:* To validate the SARC-F questionnaire for sarcopenia screening in musculoskeletal disease setting, and to assess improvements in diagnostic accuracy by adding “EBM” (elderly and body mass index information) to the SARC-F. *Design:* Diagnostic accuracy study. *Setting and Participants:* The center involved in this study was located in an urban area of Kobe City, Japan. People with musculoskeletal disease in the knee, hip, or spine who were scheduled for surgical treatment were included. *Measurements:* Sarcopenia was evaluated using the Asian Working Group for Sarcopenia (AWGS) and the European Working Group on Sarcopenia in Older People (EWGSOP2), which included bioimpedance, handgrip strength, and gait speed. Patients answered the SARC-F questionnaire and their body mass index was measured. SARC-F and “EBM” information were combined into an original score. The sensitivities, specificities, and areas under the receiver operating characteristic curve (AUC) were estimated and compared to identify sarcopenia. *Results:* A total of 959 patients were included. Sarcopenia by AWGS criteria was identified in 36 (3.8%) patients. SARC-F had a sensitivity of 41.7% and specificity of 68.5%. SARC-F+EBM had a sensitivity of 77.8% and specificity of 69.6%, with substantial improvement in sensitivity ($P<0.001$). The AUCs for SARC-F and SARC-F+EBM were 0.557 (95% confidence interval [CI] 0.452-0.662) and 0.824 (95% CI 0.762-0.886), respectively ($P<0.001$). Similar results were obtained when EWGSOP2 criteria were used as the reference standard. *Conclusion:* The SARC-F alone is not adequate for finding cases in musculoskeletal disease settings. SARC-F+EBM significantly improved the sensitivity and overall diagnostic accuracy of the SARC-F for screening sarcopenia. SARC-F+EBM is potentially useful for screening sarcopenia in different ethnic and disease settings.

Key words: Sarcopenia, muscle mass, diagnostic accuracy.

Introduction

Musculoskeletal disease is one of the most common disorders among elderly individuals, accounting for 7.5% of the total burden of disease in people aged 60 years or over. (1). Among the musculoskeletal diseases, low back pain and osteoarthritis are the main contributors that cause disability-adjusted life years (1, 2). Sarcopenia, characterized by low muscle strength and low muscle quantity or quality (3), is gaining attention in patients with musculoskeletal disease (4-6), as it is associated with increased risks of falling and mortality among the general geriatric population (7, 8). One epidemiologic study showed that the prevalence of sarcopenia is 9.1% among patients with osteoarthritis (9). However, in clinical practice settings, sarcopenia among patients with musculoskeletal disease remains underdiagnosed due to scant research on diagnostic testing (6).

For case findings in community settings, the SARC-F questionnaire is recommended by the European Working Group on Sarcopenia in Older People (EWGSOP) before performing reference standard (i.e., confirmatory testing) for sarcopenia (3). Although the SARC-F is simple (i.e., only consisting of 5 items and giving a score of 0 to 10 points), easy to administer, inexpensive, and has been validated in many ethnic populations

(3), its diagnostic accuracy has not yet been assessed among musculoskeletal diseases. In addition, the sensitivity of its cut-off (i.e., ≥ 4 points) to classify sarcopenia is admittedly low (3, 6). To overcome this challenge, the combined use of the SARC-F and routine measurements predicting sarcopenia would foster effective case finding. We conceived the “SARC-F+EBM,” which is composed of the SARC-F, elderly (“E”), and body mass index (“BM”).

In this study, we aimed to (1) validate the diagnostic accuracy of the SARC-F and (2) to derive the SARC-F+EBM for screening sarcopenia, using the “Screening for People Suffering Sarcopenia in Orthopedic cohort of Kobe study” (SPSS-OK), a large single-center cross-sectional study.

Methods

Design, setting, and participants

SPSS-OK was a cross-sectional study approved by a local institutional review board and an institutional research ethics committee. Informed consent was obtained in writing according to Japan’s ethical guidelines for medical and health research involving human subjects. The center involved in this study was located in the central part of Kobe City. Our target population comprised adult patients with musculoskeletal

Table 1
Operationalized definitions of diagnostic criteria for sarcopenia

	(1) Low muscle quantity	(2) Low muscle strength	(3) Low physical performance	Diagnostic criteria
AWGS (10)	ASMI <7.0 kg/m ² for men; ASMI <5.7 kg/m ² for women, by BIA	Grip strength <26 kg for men; Grip strength <18 kg for women	Gait speed < 0.8 for both sexes	(1) + (2) or (1) + (3) for sarcopenia
EWGSOP2 (3) ^a	ASMI <7.0 kg/m ² for men; ASMI <5.5 kg/m ² for women, by BIA	Grip strength <27 kg for men; Grip strength <16 kg for women	Gait speed ≤ 0.8 for both sexes	(1) + (2) for sarcopenia; (1) + (2) + (3) for severe sarcopenia

AWGS: Asian Working Group for Sarcopenia; ASMI: appendicular skeletal mass index; BIA: bioimpedance analysis; EWGSOP2: European Working Group on Sarcopenia in Older People; a. Cut-off value of ASMI for women was stated as 6.0 kg/m² on 13 Oct 2018. However, the value was corrected as 5.5 kg/m² on 13 May 2019.

disease who were scheduled for spinal surgery or knee or hip replacement therapy. Their diagnoses were discerned by orthopedic surgeons in charge of the patients via physical examination and imaging. Only patients waiting for their first surgery were analyzed to avoid the potential interference of the implanted artificial materials with bioimpedance analysis (BIA). Patients with neuromuscular disease were excluded as they were not likely to undergo BIA measurement. This diagnostic accuracy study was planned before data collection and questionnaire administration. The patients were recruited consecutively from August 2016 to February 2019.

Reference standards

The reference standard was sarcopenia determined according to the definition of the Asian Working Group for Sarcopenia (AWGS) (10) or the European Working Group on Sarcopenia in Older People 2 (EWGSOP2) (3) (see Table 1 for detailed diagnostic criteria). AWGS criteria use a combination of low muscle quantity (determined by appendicular skeletal mass index [ASMI], namely, appendicular skeletal muscle mass adjusted by height) and low muscle strength (determined by handgrip strength) or a combination of low muscle quantity and low physical performance (determined by gait speed). EWGSOP2 criteria use a combination of low muscle quantity and low muscle strength. All measurements for this study were obtained during the pre-surgery visit.

ASMI was calculated by dividing appendicular skeletal muscle mass by height squared (m²). Appendicular skeletal muscle mass was measured using BIA with the MC-780A body composition analyzer (Tanita, Tokyo, Japan); the analysis was performed in patients with stable condition. The monitor measures bioimpedance at three frequencies between 5 kHz and 250 kHz, as it is a multi-frequency device. The validation of BIA method using dual-energy X-ray absorptiometry (DPX-L, GE Healthcare) showed that the correlations between the two methods are high for appendicular skeletal muscle mass (12). Handgrip strength was measured twice for both left and right hands using a grip strength dynamometer (GRIP-D T.K.K. 5401; Takei Scientific Instruments Co., Ltd, Japan), and the value was calculated from the average of measured values. Gait speed was calculated from measures of walking time obtained on a 10-m length walkway. Patients were instructed to walk along a 15-m smooth and horizontal walkway, in which the segment consists of a 10-m length walkway for

measurement and 2.5-m-long sections from either end for acceleration and deceleration. Walking time in the 10-m section was measured twice using a stopwatch and was converted to m/s. The average value among the two was used for analysis. A Well-trained board-certified physiotherapists conducted all the measurements, and performers were unaware of the index test, as the index test was performed later.

Index test: SARC-F, SARC-F+E, and SARC-F+EBM

SARC-F was originally developed in English (12). It has already been translated into Japanese (Supplementary Table 1), and the conceptual equivalence of that Japanese version has been examined via back-translation and consultation with the authors of the original version (John E. Morley). Each item has 0 to 2 points, and summation of the 5 items constitutes the total score (Supplementary Table 2). The summation score ≥ 4 corresponds to cut-off for sarcopenia (12). The questionnaire including the SARC-F was administered at the date of admission for surgical treatment. The responses in SARC-F was not affected by the reference standard, as sarcopenia was classified after data collection. SARC-F+EBM was a combined scoring algorithm using SARC-F, age, and BMI. Age was determined at the date of the test for reference standard. Age ≥75 years was considered to be old (“E”lderly) (13). Body weight was measured using a digital scale. Height was measured using a fixed stadiometer. A body mass index (“BM”I) ≤ 21 kg/m² was considered to be underweight.[14,15]. For “E”lderly, zero point was assigned for age < 75 years, whereas 10 points was assigned for age ≥ 75 years. For “BM”I, zero point was assigned for those who were not underweight, whereas 10 points for those who were underweight. The resulting SARC-F+EBM score ranges from 0 to 30 points (Supplementary Table 2). We also considered SARC-F+E scoring, with combination of SARC-F plus old age (zero point for age < 75 years, whereas 10 points for age ≥ 75 years), ranging from 0 to 20 points (Supplementary Table 2).

Measurement of comorbidities

Data on baseline characteristics were collected to describe comorbidities of the study population. History of hypertension, cancer, chronic lung disease, heart disease, and stroke were determined by self-reported questionnaire. Heart disease was considered to be present if the patients answered “yes” to having either “heart attack,” “congestive heart failure,” or

SARC-F VALIDATION AND SARC-F+EBM DERIVATION IN MUSCULOSKELETAL DISEASE

Table 2
 Baseline characteristics in overall and by sarcopenia^a

	Total	Sarcopenia Positive	Sarcopenia Negative
	N = 959	N = 36	N = 923
Age, years	69.4 (9.3)	76.7 (8.7)	69.1 (9.2)
Female, n (%)	699 (72.9)	31 (86.1)	668 (72.4)
Body mass index, kg/m ²	24.7 (3.9)	20.2 (2.0)	24.9 (3.8)
Comorbidities, n (%)			
Cancer	77 (8.2)	2 (5.6)	75 (8.3)
<i>missing</i>	16	0	16
Chronic lung disease	8 (0.9)	1 (2.8)	7 (0.8)
<i>missing</i>	17	0	17
Heart disease	40 (4.2)	2 (5.6)	38 (4.2)
<i>missing</i>	16	0	16
Stroke	16 (1.7)	0 (0.0)	16 (1.8)
<i>missing</i>	18	0	18
Diabetes	151 (15.9)	6 (16.7)	145 (15.18)
<i>missing</i>	7	0	7
Chronic kidney disease	220 (23.0)	12 (33.3)	208 (22.6)
<i>missing</i>	2	0	2
Musculoskeletal disease eliciting surgical treatment, n (%)			
Knee disease	529 (55.2)	20 (55.6)	509 (55.2)
Hip disease	242 (25.2)	8 (22.2)	234 (25.4)
Spinal disease	188 (19.6)	8 (22.2)	180 (19.5)
Underlying knee disease, n (%) ^b			
Osteoarthritis	506 (95.7)	19 (95.0)	487 (95.7)
Necrosis	88 (16.6)	1 (5.0)	87 (17.1)
Underlying hip disease, n (%) ^b			
Osteoarthritis	231 (95.5)	7 (87.5)	224 (95.7)
Necrosis	17 (7.0)	1 (12.5)	16 (6.8)
Others	2 (0.8)	1 (12.5)	1 (0.4)
Underlying spinal disease, n (%) ^b			
Spinal canal stenosis	174 (92.6)	8 (100.0)	166 (92.2)
Spondylosis	76 (40.4)	5 (62.5)	71 (39.4)
Others	26 (13.8)	0 (0.0)	26 (14.4)
Appendicular skeletal mass index, kg/m ²	7.1 (1.2)	5.6 (0.5)	7.2 (1.2)
Handgrip strength, kg	24.0 (7.9)	16.7 (3.6)	24.3 (7.9)
Gait speed, m/s	1.04 (0.25)	0.92 (0.25)	1.05 (0.25)

a. Sarcopenia according to the definition of the Asian Working Group for Sarcopenia (10); b. Percentage was calculated within surgical treatment category. Multiple counts were allowed for underlying disease within the category (e.g., One patient has both osteoarthritis and necrosis of his/her knee).

“angina.” Chronic kidney disease was defined as estimated glomerular filtration rate (eGFR) values ≤ 60 mL/min/1.73 m², which were calculated based on the 3-variable Japanese equation using age, serum creatinine level, and sex as follows (16): $eGFR = 194 \times \text{serum creatinine}^{-1.094} \times \text{age}^{-0.287} \times 0.739$ (if female). Diabetes was defined as having glycosylated hemoglobin values of $\geq 6.5\%$ according to the National Glycohemoglobin Standardization Program.

Statistical analyses

All statistical analyses were conducted using Stata/SE, version 15 (Stata Corp., College Station, TX). Sociodemographic characteristics, comorbidities, underlying musculoskeletal disease, and outcomes (sarcopenia defined by two criteria, ASMI, handgrip strength, and gait speed) were described for all patients and separately by sarcopenia defined by AWGS. To examine the diagnostic test accuracy of the SARC-F questionnaire among the patients, the sensitivity and specificity of SARC-F score were estimated with provision of contingency table. To examine the overall diagnostic accuracy among SARC-F, SARC-F+E, and SARC-F+EBM, we fitted logistic regression models with sarcopenia being dependent variable and those scoring methods being independent variables. Receiver operating characteristic curves were then derived, and their areas under the curves (AUCs) were estimated and compared using the DeLong method (17). Larger AUC indicates better overall diagnostic accuracy. Cut-off points for the SARC-F+E and the SARC-F+EBM scores were determined using the Liu’s method (18). The sensitivities and specificities among the SARC-F, SARC-F+E, and SARC-F+EBM were compared using the McNemar’s test for those with sarcopenia and those without sarcopenia, respectively. (19). All the analyses were separately performed for sarcopenia determined by AWGS and EWGSOP2 criteria. All analyses were conducted by a well-trained epidemiologist (N.K.). Two-sided $P < .05$ was considered statistically significant.

Results

A total of 996 patients with musculoskeletal disease were invited. Figure 1 shows the flow of patients in this study, along with the outcome of sarcopenia defined by AWGS. After exclusion of 22 patients with incomplete response to the SARC-F and 15 patients with missing value for sarcopenia testing, 959 patients were evaluated. The mean age was 69.4 years, and three quarters of the patients were female (Tables 2). The median difference in dates between the reference and index tests was 31 days (interquartile range: 21-42 days).

The test results of the SARC-F for patients with and those without sarcopenia are shown in Supplementary Table S3. The prevalence of sarcopenia was 3.8% (36/959). When conventional cut-off point (≥ 4) was applied, only 41.7% of the patients with sarcopenia met this condition (sensitivity 41.7%, 95% confidence interval [95% CI] 25.5% to 59.2%), and 68.5%

of the patients without sarcopenia did not meet this condition (specificity 68.5%, 95% CI 65.4% to 71.5%). The ROC curve of the SARC-F to identify sarcopenia is presented in Figures 2, and the AUC was 0.557 (95% CI 0.452 to 0.662) (Tables 3).

Figure 1
Study flowchart

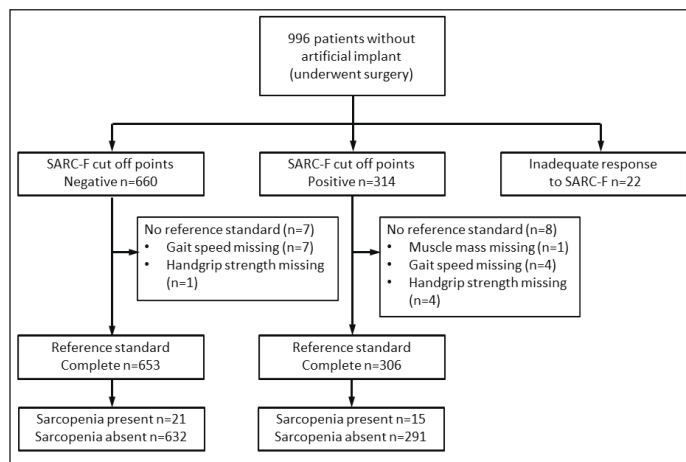
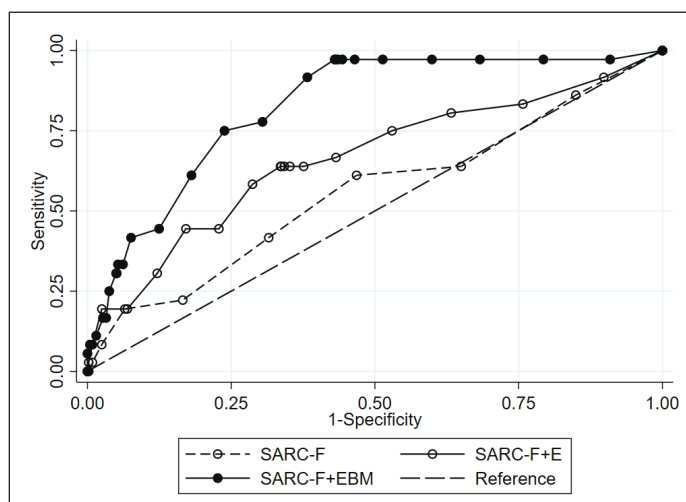


Figure 2

Receiver operating characteristics curves for SARC-F, SARC-F+E, and SARC-F+EBM for identifying sarcopenia by AWGS criteria



Dashed line with hollow circle symbols indicates receiver operating characteristics (ROC) curve for SARC-F scoring. Solid line with hollow circle symbols ROC curve for SARC-F+E scoring. Solid line with solid circle symbols indicates ROC curve for SARC-F+EBM scoring.

The ROC curve of the SARC-F+EBM to identify sarcopenia is presented in Figure 2. The AUC of the SARC-F+EBM was 0.824 (95% CI 0.762-0.886), which is superior to that of the SARC-F (Tables 3). The optimal cut-off point for sarcopenia was ≥ 12 points. The sensitivity of the SARC-F+EBM was 77.8% (95% CI 60.8% to 89.9%), which is superior to that of the SARC-F ($P < 0.001$) (Tables 3). The specificity of SARC-

SARC-F VALIDATION AND SARC-F+EBM DERIVATION IN MUSCULOSKELETAL DISEASE

Table 4
Sensitivity and specificity of the SARC-F, SARC-F + E, and SARC-F + EBM^a

Scoring	Sensitivity		P value for equality		Specificity		P value for equality	
	Point estimate	(95% CI)			Point estimate	(95% CI)		
1) SARC-F cut-off point ≥ 4	41.7%	(25.5% to 59.2%)			68.5%	(65.4% to 71.5%)		
2) SARC-F+E cut-off point ≥ 9	63.9%	(46.2% to 79.2%)	2) vs 1)	0.046	66.3%	(63.2% to 69.4%)	2) vs 1)	0.302
3) SARC-F+EBM cut-off point ≥ 12	77.8%	(60.8% to 89.9%)	3) vs 1)	<0.001	69.6%	(66.5% to 72.5%)	3) vs 1)	0.565

a. Sarcopenia according to the definition of the Asian Working Group for Sarcopenia (10)

F+EBM was 69.6% (95% CI 66.5% to 72.5%), which is similar to that of SARC-F (P=0.565).

Similarly, the ROC curve of SARC-F+E to identify sarcopenia is presented in Figure 2. The AUC of SARC-F+E was 0.663 (95% CI 0.561-0.765), which is superior to that of SARC-F (Tables 3). The optimal cut-off point for sarcopenia was ≥ 9 points. The sensitivity of the SARC-F+E was 63.9% (95% CI 46.2% to 79.2%), which is superior to that of the SARC-F (P=0.046) (Tables 4). The specificity of SARC-F+E was 66.3% (95% CI 63.2% to 69.4%), which is similar to that of SARC-F (P=0.302).

Similar results were obtained when sarcopenia by EWGSOP2 criteria was used as reference standard (Supplementary Tables 4 to 6).

Table 3

Overall diagnostic accuracy of the SARC-F, SARC-F + E, and SARC-F + EBM^a

Scoring	AUC		P value for equal AUC ^b	
	Point estimate	(95% CI)		
1) SARC-F	0.557	(0.452 to 0.662)		
2) SARC-F+E	0.663	(0.561 to 0.765)	2) vs 1)	0.031
3) SARC-F+EBM	0.824	(0.762 to 0.886)	3) vs 1)	<0.001

a. Sarcopenia according to the definition of the European Working Group on Sarcopenia in Older People in 2018 (3); b. P values were adjusted for multiple comparisons.

Discussion

We evaluated the diagnostic accuracy of the SARC-F questionnaire to classify sarcopenia as a screening purpose among patients with musculoskeletal disease and found that its overall diagnostic accuracy based on AUC was not acceptable. However, combined use of the SARC-F with “EBM” (Elderly plus BMI) resulted in excellent overall diagnostic accuracy and improved sensitivity. We propose that for sarcopenia screening, our “SARC-F+EBM” scoring should be applied to other ethnic or disease settings because of its simplicity.

With regard to the diagnostic accuracy of the SARC-F with conventional cut-off points (≥ 4), our study findings may partially corroborate those of previous research. Low sensitivity is shown in Chinese community-dwelling individuals (3.8% to 9.9%) (20) and in Brazilian community-dwelling individuals (33.0%) (21). In our study, regardless of sarcopenia definition, the sensitivity was under 50%. This low sensitivity may be attributed to the face validity of the SARC-F items, as previous researchers pointed out (21): the items ask about muscle strength and physical performance, but not about muscle quantity. The low specificity of this study (68.4% to 68.5%) contradicts the finding of a previous research (21, 20), and may be due to our unique disease population. Generally, specificity becomes larger if participants without reference standard are healthier (22). Our populations are composed of musculoskeletal disease requiring surgery, and they may have problems in physical performance, but they may not have sarcopenia. Previous studies showed that indices of lower limb performance are affected by pain among musculoskeletal disease (11). Therefore, participants without sarcopenia in our musculoskeletal disease settings demonstrated high false-positive rates for the SARC-F and contributed to low specificity.

Our findings have several implications for researchers and physicians. First, our “SARC-F+EBM” scoring is applicable to other ethnic and disease settings and expected to demonstrate high sensitivity. Age and BMI are inexpensive and easily obtainable information. In addition, “EBM” serves as items asking muscle quantity to some degree, as old age (age ≥ 75 years) corresponds to age-associated muscle loss and underweight (BMI ≤ 21 kg/m²) corresponds to undernutrition (3). Thus, incorporating such items in the SARC-F+EBM will foster to effectively screen sarcopenia in different populations. Even if weighing scale is not available in rural community settings, “SARC-F+E” can be used.

Second, the consistency of the magnitudes of the diagnostic accuracy of sarcopenia, which is based on two different criteria (AWGS and EWGSOP2), suggests that the observed diagnostic performance is unlikely to be affected by the use of gait speed for definition of reference standard. This is because only AWGS criteria require gait speed measurement to diagnose

sarcopenia.

Third, the fact that the specificities of the SARC-F and SARC-F+EBM are not high enough to rule in sarcopenia among musculoskeletal disease may be applicable to other diseases, which also cause impairment in physical performance. For example, patients with neurological diseases (such as Parkinson's disease and hemiplegia due to stroke) or organ failure (such as heart failure and chronic obstructive lung disease) have difficulty in walking, climbing stairs, or rising from a chair because of the disease itself and thus may tend to show false-positive results in the SARC-F scoring. Thus, to enhance the SARC-F, researchers may consider adding extra items asking about muscle quantity or muscle strength rather than about physical performance. However, this notion does not hamper the use of the SARC-F+EBM in clinical or community settings. Rather, we could show that the SARC-F+EBM has improved sensitivity (77.8% to 84.2%) and exhibited excellent overall diagnostic accuracy (AUC 0.824 to 0.876). Thus, the SARC-F+EBM can be used as a sarcopenia screening tool (i.e., tool preventing people without sarcopenia from taking unnecessary confirmatory workup for sarcopenia).

This study has several strengths. First, our study assessed the validity of the SARC-F to screen sarcopenia among a population with musculoskeletal disease, which is rarely investigated. Second, recruitment of a relatively large number of patients at a single-center enabled us to measure important variables (i.e., SARC-F as index test and sarcopenia [defined by ASMI, handgrip strength, and gait speed] as reference standard) with uniform measurement methods across all patients to effectively assess our research question. Third, our study reinforced the SARC-F by adding "EBM," using formal analyses required for diagnostic accuracy study.

This study has also several limitations. First, BIA was performed instead of volumetric analysis using computed tomography or magnetic resonance imaging to quantify muscle mass. However, BIA is a major approach of sarcopenia screening in population study settings and is recommended as an option for measuring muscle mass by the EWGSOP2 and AWGS (3, 10). Second, the prevalence of sarcopenia was unexpectedly low. A selection bias (i.e., referring bias) may exist with regard to the low prevalence, as was described already, as the patients were referred because they were expected to be amenable to surgery without complications.

Conclusions and Implications

The SARC-F+EBM improved the sensitivity and overall diagnostic accuracy of SARC-F to screen sarcopenia in musculoskeletal disease population. Thus, the SARC-F+EBM may enhance the SARC-F in the "find-cases" pathway (3). Further validation studies in different settings are warranted to elucidate the sensitivity and specificity of the SARC-F+EBM obtained in this population.

Brief summary

The SARC-F combined with elderly and body mass index (SARC-F+EBM) significantly improved the sensitivity and overall diagnostic accuracy of the SARC-F for screening sarcopenia in musculoskeletal disease settings.

Conflicts of Interest: The authors have nothing to disclose.

Ethical Standards: The study was conducted in accordance with the Declaration of Helsinki and the ethical guidelines for Ethical Guidelines for Medical and Health Research Involving Human Subjects in Japan.

Funding sources: This study was supported by JSPS KAKENHI (Grant Number: JP15K16518). The JSPS had no role in the study, except for providing funding.

Acknowledgments: The authors greatly thank the following research assistant and medical staff members for their assistance in collecting the clinical information used in this study: Takehiro Kaga, Tomohiro Oka, Yoriko Tamura, Hiroshi Nishi, Yuichi Isaji, Yutaka Sato, Tomohiro Takagi, Kaho Shibata, Maho Wakai, Chisato Shindoh, Kenta Hirose, Takuma Ota, Tatsuya Arita, Yuuki Ikawa, Tsuyoshi Fukui, Riuji Nakagawa, Taisuke Hayashida, Shuto Fujii, Keisuke Yoneya, Kazuaki Mori (Anshin Hospital, Kobe), Asako Tamura, Yuka Masuda (St. Marianna Medical University), Lisa Shimokawa (Fukushima Medical University Hospital, Fukushima-city, Fukushima).

References

1. Prince MJ, Wu F, Guo Y, Gutierrez Robledo LM, O'Donnell M, Sullivan R, Yusuf S. The burden of disease in older people and implications for health policy and practice. *The Lancet* 2015;385 (9967):549-562. doi:https://doi.org/10.1016/S0140-6736(14)61347-7
2. Brooks PM. The burden of musculoskeletal disease—a global perspective. *Clinical Rheumatology* 2006;25 (6):778-781. doi:10.1007/s10067-006-0240-3
3. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyere O, Cederholm T, Cooper C, Landi F, Rolland Y, Sayer AA, Schneider SM, Sieber CC, Topinkova E, Vandewoude M, Visser M, Zamboni M. Sarcopenia: revised European consensus on definition and diagnosis. *Age and ageing* 2019;48 (1):16-31. doi:10.1093/ageing/afy169
4. Godziuk K, Prado CM, Woodhouse LJ, Forhan M. The impact of sarcopenic obesity on knee and hip osteoarthritis: a scoping review. *BMC musculoskeletal disorders* 2018;19 (1):271. doi:10.1186/s12891-018-2175-7
5. Veronese N, Punzi L, Sieber C, Bauer J, Reginster J-Y, Maggi S. Sarcopenic osteoarthritis: a new entity in geriatric medicine? *Eur Geriatr Med* 2018;9 (2):141-148. doi:10.1007/s41999-018-0034-6
6. Arai H. Sarcopenia diagnostic criteria update by EWGSOP: what has been changed? *Eur Geriatr Med* 2018;9 (6):733-734. doi:10.1007/s41999-018-0117-4
7. Landi F, Cruz-Jentoft AJ, Liperoti R, Russo A, Giovannini S, Tosato M, Capoluongo E, Bernabei R, Onder G. Sarcopenia and mortality risk in frail older persons aged 80 years and older: results from the SIRENTE study. *Age and ageing* 2013;42 (2):203-209
8. Landi F, Liperoti R, Russo A, Giovannini S, Tosato M, Capoluongo E, Bernabei R, Onder G. Sarcopenia as a risk factor for falls in elderly individuals: results from the SIRENTE study. *Clin Nutr* 2012;31 (5):652-658
9. Kemmler W, Teschler M, Goisser S, Bebenek M, von Stengel S, Bollheimer LC, Sieber CC, Freiburger E. Prevalence of sarcopenia in Germany and the corresponding effect of osteoarthritis in females 70 years and older living in the community: results of the FORMOSA study. *Clin Interv Aging* 2015;10:1565-1573. doi:10.2147/CIA.S89585
10. Chen LK, Liu LK, Woo J, Assantachai P, Auyeung TW, Bahyah KS, Chou MY, Chen LY, Hsu PS, Krairit O, Lee JS, Lee WJ, Lee Y, Liang CK, Limpawattana P, Lin CS, Peng LN, Satake S, Suzuki T, Won CW, Wu CH, Wu SN, Zhang T, Zeng P, Akishita M, Arai H. Sarcopenia in Asia: consensus report of the Asian Working Group for Sarcopenia. *Journal of the American Medical Directors Association* 2014;15 (2):95-101. doi:10.1016/j.jamda.2013.11.025
11. Wada O, Kurita N, Kamitani T, Nakano N, Mizuno K. Influence of the severity of knee osteoarthritis on the association between leg muscle mass and quadriceps strength: the SPSS-OK study. *Clinical Rheumatology* 2019;38 (3):719-725. doi:10.1007/s10067-018-4337-2
12. Malmstrom TK, Morley JE. SARC-F: a simple questionnaire to rapidly diagnose sarcopenia. *Journal of the American Medical Directors Association* 2013;14 (8):531-532. doi:10.1016/j.jamda.2013.05.018
13. Ouchi Y, Rakugi H, Arai H, Akishita M, Ito H, Toba K, Kai I, Society obotJCoJG, definition JGSot, elderly cot. Redefining the elderly as aged 75 years and older: Proposal from the Joint Committee of Japan Gerontological Society and the Japan Geriatrics Society. *Geriatrics & gerontology international* 2017;17 (7):1045-1047. doi:10.1111/ggi.13118
14. Gaillard C, Alix E, Sallé A, Berrut G, Ritz P. Energy requirements in frail elderly people: A review of the literature. *Clinical Nutrition* 2007;26 (1):16-24. doi:https://doi.org/10.1016/j.clnu.2006.08.003
15. Volkert D, Beck AM, Cederholm T, Cruz-Jentoft A, Goisser S, Hooper L, Kiesswetter E, Maggio M, Raynaud-Simon A, Sieber CC, Sobotka L, van Asselt D, Wirth R, Bischoff SC. ESPEN guideline on clinical nutrition and hydration in geriatrics. *Clinical Nutrition* 2019;38 (1):10-47. doi:10.1016/j.clnu.2018.05.024

SARC-F VALIDATION AND SARC-F+EBM DERIVATION IN MUSCULOSKELETAL DISEASE

16. Imai E, Horio M, Nitta K, Yamagata K, Iseki K, Tsukamoto Y, Ito S, Makino H, Hishida A, Matsuo S. Modification of the Modification of Diet in Renal Disease (MDRD) Study Equation for Japan. *Am J Kidney Dis* 2007;50 (6):927-937. doi:http://dx.doi.org/10.1053/j.ajkd.2007.09.004
17. 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 (3):837-845. doi:10.2307/2531595
18. Liu X. Classification accuracy and cut point selection. *Statistics in medicine* 2012;31 (23):2676-2686
19. Trajman A, Luiz RR. McNemar chi2 test revisited: comparing sensitivity and specificity of diagnostic examinations. *Scandinavian journal of clinical and laboratory investigation* 2008;68 (1):77-80. doi:10.1080/00365510701666031
20. Woo J, Leung J, Morley JE. Validating the SARC-F: A Suitable Community Screening Tool for Sarcopenia? *Journal of the American Medical Directors Association* 2014;15 (9):630-634. doi:https://doi.org/10.1016/j.jamda.2014.04.021
21. Barbosa-Silva TG, Menezes AMB, Bielemann RM, Malmstrom TK, Gonzalez MC. Enhancing SARC-F: Improving Sarcopenia Screening in the Clinical Practice. *Journal of the American Medical Directors Association* 2016;17 (12):1136-1141. doi:https://doi.org/10.1016/j.jamda.2016.08.004
22. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HCW, Bossuyt PMM. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ open* 2016;6 (11):e012799-e012799. doi:10.1136/bmjopen-2016-012799