Sarcopenia: evaluation of different diagnostic criteria and its association with muscle strength and functional capacity

Sarcopenia: avaliação de diferentes critérios diagnósticos e sua associação com a força muscular e capacidade funcional

Abstract

Introduction: Sarcopenia has been associated with increased morbidity and mortality in older people. However, there is still no consensus about the best diagnostic criteria for it. Objective: This study aimed to evaluate different diagnostic criteria of sarcopenia and their relationship with muscle strength and functional capacity. Methods: A cross-sectional study was carried out in Santa Maria-RS, Brazil. Body composition was measured by dual-energy X-ray absorptiometry (DXA). Physical performance was evaluated by the Timed-get-up-and-go test (TGUG). Muscle strength was measured with a handgrip dynamometer. The proposed criteria of sarcopenia were evaluated as follows: A) appendicular lean mass index (aLM/ht²), cutoffs 5.5 kg/m², 2SD, 10th percentile, and 20th percentile of young reference; B) a linear regression model was used to adjust appendicular lean mass (aLM) for both height and fat mass, cutoff 20th percentile of elderly. Results: In total, 104 women agreed to participate in the research, 39 young (20 to 40 years) and 65 elderly (over 60 years). The prevalence of sarcopenia varied from 0% to 17.2% according to the diagnostic criteria. Regarding muscle strength, the aLM/ht² cutoff 10th percentile was the criterion with greater area under the ROC curve. However, there were no criteria with an area under the ROC curve greater than 0.5 when TGUG was the outcome. The reclassification showed that the criterion B reclassified 66.7% with normal grip strength as sarcopenic. Conclusion: Of these criteria, aLM/ht² cutoff 10th percentile had stronger correlation with muscle strength.

Key words: Sarcopenia. Postmenopause. Muscle Strength. Calibration. Discrimination.

http://dx.doi.org/10.1590/1809-9823.2015.14025
INTRODUCTION

The knowledge about sarcopenia as a syndrome has evolved in the last decades. Its association with increased morbidity and mortality is now well described. Nevertheless, there is still no consensus about the best diagnostic criteria for sarcopenia. The first criterion, proposed by Baumgartner et al. in 1998, used a young population as reference. According to this criterion, an individual is classified as sarcopenic when his/her appendicular lean mass index (appendicular lean mass/height\(^2\) (aLM/ht\(^2\))] is two standard deviations below the mean appendicular lean mass index (aLM/ht\(^2\)) of the young reference population. As this approach does not correct the appendicular muscle mass for fat mass, it might underestimate the prevalence of sarcopenia in obese people. Newman et al. suggested the use of linear regression adjusted for fat mass to deal with this problem. It uses the 20th percentile of a linear regression residual that includes only elderly subjects to define sarcopenia (the regression adjusts for height and total fat mass). Therefore the choice of the group of reference might have considerable influence on the sarcopenia diagnosis.

Despite the fact that sarcopenia is characterized by progressive and generalized loss of muscle mass and strength, the latter have been rarely used to evaluate the proposed diagnostic criteria. Both The European Working Group on Sarcopenia in Old People (EWGSOP) and the International Working Group on Sarcopenia (IWG) recommend that a sarcopenia definition should incorporate the low muscle mass and low muscle strength plus (or) low physical performance. One hypothesis is that the different diagnostic criteria might have different relationship with muscle strength and (or) physical performance. The aim of this study was to evaluate different diagnostic criteria of sarcopenia and its association with muscle strength and functional capacity.
METHODS

Study design and population

A cross-sectional study was carried out in the city of Santa Maria, Southern Brazil. Women were recruited from a Catholic Parish address lists from January to March 2012. Women aged between 20 to 40 years-old and over 60 years-old were invited to participate. Subjects with rheumatoid arthritis, bowel diseases, hepatic diseases, renal failure, thyroid diseases, neurodegenerative diseases, chronic obstructive pulmonary disease, and women using steroid hormones were excluded. The study was approved in 2011 by the Ethics Committee of the University of Santa Maria, CEP registration number 0311.0.243.000-11. The study protocol was in accordance with the ethical guidelines of the 1975 Declaration of Helsinki. All individuals provided an informed consent term.

The number of subjects necessary to find an appendicular lean mass index (aLM/ht²) average of 6.1 kg/m² and a standard deviation of 0.8 kg/m² was 38.5 The authors established a confidence level of 95% and a power of 90%.

Assessments

A standardized questionnaire was used to collect demographic data. Weight was measured in patients in light clothing without shoes using a balance beam scale while height was measured using a wall-mounted stadiometer. Body composition (whole body and regional measures of fat mass and lean mass) was measured by dual-energy X-ray absorptiometry (Lunar Prodigy Pro, GE Health Care, Madison, WI), according to the Official Positions of the International Society for Clinical Densitometry (ISCD).11 Participants were asked to lie flat on the table, with arms by their side, legs straight and side by side. All metal was removed as possible. For participants too large to fit within the dimensions of the scanning field, two hemi-scan were performed (2% of participants had a hemi-scan), one each side. The coefficients of variation (CV) were 1.5% for total body fat mass, 1.4% for total bone mineral densitometry (BMD) and 0.9% for total body lean mass. Appendicular lean mass (aLM) was measured as the sum of the lean soft-tissue masses for the arms and the legs.11

Physical performance was evaluated by the Timed get-up-and-go test (TGUG). All women were asked to sit in a 43-cm chair.12,13 The time to get up, walk a 3-meter distance, walk back and sit again was recorded. Muscle strength was measured with a handgrip dynamometer (Jamar Hydraulic Hand Dynamometer, Sammons Preston, Chicago, IL), according to the American Society for Hand Therapists Society recommendation.14 The test was performed in both hands in standing position with straight back, the shoulder adducted and in neutral rotation, elbow flexed 90°, and the lower arm and wrist in the neutral position. The tests were performed three times after a learning trial and rest interval between tests was 1 minute. The cutoff suggest by the European Working Group on Sarcopenia in Older People was used (body mass index [BMI] ≤23: ≤17 kg; 23.1< BMI ≤26: ≤17.3 kg; 26.1< BMI ≤29: ≤18 kg; BMI > 29: ≤21 kg).10 These tests were chosen because they were easy to perform at the clinical practice.

Criteria of sarcopenia

The proposed criteria of sarcopenia were evaluated as follows: firstly, it was calculated the relative aLM using the Baumgartner formula (aLM/ht²). It was used the classical Baumgartner definition of sarcopenia [two standard deviation (SD) [G1] and the Rosetta study reference [G2].4 As the best criterion for the Brazilian population is unknown, it was also used the 10th percentile [G3], and 20th percentile of the young [G4] population as cut-points to define
sarcopenia. Furthermore, aLM was calculated adjusting for fat mass in addition to height in the elderly women, as suggested by Newman. A linear regression model was used to adjust aLM for both height and fat mass. The percentile proposed by Newman (20th percentile) of the regression residuals was used as cut-point to define sarcopenia [G5].

Statistical analyses

The prevalence of sarcopenia was calculated using the proposed criteria. Student $t$ test and Chi-square test were used to compare younger and elderly women. Logistic regression was used to evaluate the association between sarcopenia and muscle strength and the TGUG. The Hosmer-Lemeshow goodness-of-fit statistic and corresponding $p$ value for each model were calculated to determine the best model. The proposed sarcopenia definitions were compared using the ROC curve and reclassification. Furthermore, the prevalence of sarcopenia was calculated using the Newman 20th percentile criterion plus muscle strength cut-off as suggested by the European consensus. Differences were found significant when the two-tailed $p$ value was $<0.05$, and confidence intervals are provided where appropriate.

RESULTS

In total, 227 families were identified from the Catholic Parish address list. Of those, 177 families returned the contact. There were 158 eligible women in these families. After excluding participants who did not fulfil the research criteria, 104 women agreed to participate in the research [39 young women (age from 20 to 40 years) and 65 old women (age over 60 years)], as shown in figure 1.

![Figure 1. Study flowchart. Santa Maria-RS, 2012.](image-url)
The demographic characteristics of these women are described in table 1. Young women were significantly slimmer and taller than elder women. They also have greater muscle strength and lower TGUG time than elder women. Although fat mass was lower in young women, all other body composition parameters were no different both groups (table 1).

Table 1. Characteristics of the study population. Santa Maria-RS, 2012.

<table>
<thead>
<tr>
<th></th>
<th>Younger (n= 39)</th>
<th>Older (n= 65)</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>29.28 ± 6.54</td>
<td>67.85 ± 6.68</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Race (white)</td>
<td>92%</td>
<td>96%</td>
<td>0.375</td>
</tr>
<tr>
<td>Smoker</td>
<td>5%</td>
<td>5%</td>
<td>0.554</td>
</tr>
<tr>
<td>Weight (Kg)</td>
<td>66.62 ± 13.35</td>
<td>70.06 ± 11.62</td>
<td>0.172</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.64 ± 0.05</td>
<td>1.57 ± 0.06</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Body mass index (Kg/m²)</td>
<td>24.96 ± 5.16</td>
<td>28.23 ± 4.21</td>
<td>0.001</td>
</tr>
<tr>
<td>Handgrip (right and left hand, Kg)</td>
<td>27.93 ± 7.24</td>
<td>22.18 ± 5.78</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TGUG† (seconds)</td>
<td>6.08 ± 0.99</td>
<td>8.98 ± 2.16</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>aLM‡ (Kg)</td>
<td>16.46 ± 2.55</td>
<td>15.98 ± 2.62</td>
<td>0.364</td>
</tr>
<tr>
<td>Relative aLM§ (Kg/m²)</td>
<td>6.15 ± 0.92</td>
<td>6.43 ± 0.87</td>
<td>0.126</td>
</tr>
<tr>
<td>Lean mass (Kg)</td>
<td>38.29 ± 4.71</td>
<td>38.34 ± 5.80</td>
<td>0.959</td>
</tr>
<tr>
<td>Fat mass (Kg)</td>
<td>25.35 ± 10.38</td>
<td>29.06 ± 7.02</td>
<td>0.033</td>
</tr>
<tr>
<td>Total mass (Kg)</td>
<td>66.24 ± 13.35</td>
<td>69.62 ± 11.52</td>
<td>0.178</td>
</tr>
</tbody>
</table>

*Timed get-up-and-go test; †Timed get-up-and-go test; ‡appendicular lean mass; §appendicular lean mass/ height² (aLM/ht²).

The cut-off of sarcopenia calculated using 2SD of the young population as suggested by Baumgartner [G1] was 4.29 kg/m² and the prevalence was 0%. The prevalence of sarcopenia using the Newman criterion 20th [G2] percentile cutoff plus muscle strength cutoff as suggested by the European consensus was 10.7%. Additionally, the prevalence of sarcopenia using the cutoff of 5.67 kg/m² plus muscle strength as indicated by the IWG was 7.8% (data not shown); all other calculated frequencies are displayed in table 2.

Regarding muscle strength, the criteria that showed the greater area under the curve in the ROC analysis was the criterion proposed by Baumgartner with the 10th [G3] percentile as cutoff (table 2). When the TGUG was used as outcome, there were no criteria with an area under the ROC curve greater than 0.5 (data not shown).
In the logistic regression, the muscle strength was strongly associated to all proposed diagnostic criteria (table 3). Nevertheless, the TGUG was only associated with the modified Baumgartner (10th percentile) [G3], \( p < 0.05 \) (table 3). The model that presented the better fit was Baumgartner with the 10th percentile as cutoff [G3]. The model using Baumgartner with the 10th percentile [G3] was better to predict muscle strength (table 2).

The reclassification showed that Newman criterion [G5] reclassified 66.7% of patients with normal grip strength as sarcopenic. All other diagnostic criteria agreed in relation to muscle strength and sarcopenia. Nevertheless, 20% of patients classified as non-sarcopenic criteria by the G2, G4 and G5 have muscle strength decreased (data not shown). The G2 and G3 criteria agreed on the diagnosis of sarcopenic subjects. In both criteria, there were 25.4% of lost individuals with muscle strength decreased. These data are displayed in table 4.

**Table 2.** Comparison of different diagnostic criteria of sarcopenia. Santa Maria-RS, 2012.

<table>
<thead>
<tr>
<th>Diagnostic criteria</th>
<th>Prevalence (%)</th>
<th>ROC curve (Confidence interval)</th>
<th>Hosmer-Lemeshow test</th>
<th>Hosmer-Lemeshow test</th>
</tr>
</thead>
<tbody>
<tr>
<td>[G2] Baumgartner - (Rosetta)</td>
<td>14.1</td>
<td>0.65 (0.45-0.85)</td>
<td>6.26, ( p = 0.62 )</td>
<td>13.1, ( p = 0.10 )</td>
</tr>
<tr>
<td>[G3] Baumgartner 10th</td>
<td>7.8</td>
<td>0.88 (0.79-0.97)</td>
<td>0.77, ( p = 0.99 )</td>
<td>8.3, ( p = 0.40 )</td>
</tr>
<tr>
<td>[G4] Baumgartner 20th</td>
<td>10.9</td>
<td>0.73 (0.53-0.94)</td>
<td>8.93, ( p = 0.35 )</td>
<td>9.9, ( p = 0.27 )</td>
</tr>
<tr>
<td>[G5] Newman</td>
<td>17.2</td>
<td>0.70 (0.52-0.88)</td>
<td>4.18, ( p = 0.84 )</td>
<td>16.0, ( p = 0.04 )</td>
</tr>
</tbody>
</table>

* Data is shown as area under the curve (area under the Receiver Operating Characteristic (ROC) curve (confidence interval 95%); **data is shown as Chi-square value, \( p \)value; ***Timed get-up-and-go test; \( \dagger \)Baumgartner – Rosetta = Baumgartner definition of sarcopenia and the Rosetta study reference (5.5 kg/m²); \( \ddagger \) Baumgartner 10th = Baumgartner definition of sarcopenia and percentile 10th (5.21 kg/m²); \( \S \) Baumgartner 20th = Baumgartner definition of sarcopenia and percentile 20th (5.67 kg/m²); Newman = 20th percentile of the regression residuals was used as cut-point to define sarcopenia (-1.42).

**Table 3.** Logistic regression models using different diagnostic criteria as predictors. Santa Maria-RS, 2012.

<table>
<thead>
<tr>
<th>Diagnostic criteria</th>
<th>Muscle Strength OR (95% CI)*</th>
<th>TGUG OR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>[G2] Baumgartner - (Rosetta)</td>
<td>0.83 (0.70-0.97), ( p = 0.02 )</td>
<td>1.24 (0.94-1.64), ( p = 0.13 )</td>
</tr>
<tr>
<td>[G3] Baumgartner 10th</td>
<td>0.50 (0.29-0.84), ( p = 0.01 )</td>
<td>1.45 (1.03-2.04), ( p = 0.03 )</td>
</tr>
<tr>
<td>[G4] Baumgartner 20th</td>
<td>0.72 (0.57-0.92), ( p = 0.01 )</td>
<td>1.30 (0.97-1.77), ( p = 0.07 )</td>
</tr>
<tr>
<td>[G5] Newman</td>
<td>0.75 (0.61-0.91), ( p = 0.01 )</td>
<td>1.30 (0.98-1.73), ( p = 0.07 )</td>
</tr>
</tbody>
</table>

* Data is shown Odds Ratio (95% confidence interval); \( \dagger \) Baumgartner – Rosetta = Baumgartner definition of sarcopenia and the Rosetta study reference (5.5 kg/m²); \( \ddagger \) Baumgartner 10th = Baumgartner definition of sarcopenia and percentile 10th (5.21 kg/m²); \( \S \) Baumgartner 20th = Baumgartner definition of sarcopenia and percentile 20th (5.67 kg/m²); Newman = 20th percentile of the regression residuals was used as cut-point to define sarcopenia (-1.42).
Table 4. Reclassification using decreased muscle strength as outcome. Baumgartner definition of sarcopenia and percentile 10th versus Newman - 20th percentile of the regression residuals. Santa Maria-RS, 2012.

<table>
<thead>
<tr>
<th>Baumgartner 10th [G3]</th>
<th>Sarcopenia</th>
<th>Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100% (5/5)*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>33.3% (2/6)</td>
<td>22.6% (12/53)</td>
</tr>
</tbody>
</table>

*number of subjects with decreased muscle strength / total of subjects in the stratum.

DISCUSSION

This study evaluated the ability of the different criteria and its different cutoffs predicting muscle strength or functional capacity. The results have shown that the classical Baumgartner definition of sarcopenia but with a 10th percentile cutoff [G3] presented the best calibration and discrimination regarding muscle strength. Although all proposed criteria were apparently well calibrated, none of those shown a good discrimination of functional capacity evaluated by the TGUG.

Differences in prevalence of sarcopenia according to different diagnostic criteria are reinforced by our findings. The prevalence of sarcopenia in this study ranged from 0% to 17.2%, depending on the operational definition. In parallel to this study, Domiciano et al. have found a prevalence varying from 3.7% (using the classical Baumgartner definition with the Rosetta study reference) to 19.9% (classical Newman criterion [20th percentile cutoff]) in 611 community-dwelling women living in Sao Paulo City, Brazil. The Fourth Korean National Health Survey has also found a prevalence ranging from 0.1% (using the classical Baumgartner definition with 2 SD cutoff) to 11.8% (using appendicular skeletal muscle mass adjusted by body weight) in the elderly women.17 Dam et al. compared nine important cohort studies [the Framingham Heart Study Original cohort and its Offspring cohort; the Study of Osteoporotic Fractures (SOF); the InCHIANTI study; the Rancho Bernardo Study, the Health, Aging, and Body Composition Study (HABC); the Osteoporotic Fractures in Men Study (MrOS); the Age, Gene and Environment Susceptibility-Reykjavik Study (AGES); and the Boston Puerto Rican Health Study (BPRHS)] and six clinical trials regarding the operational criterion agreement. They found poor agreement (varying from 4.0% to 19.9% in women, evaluated by Cohen’s kappa test) for the positive results of the studies.

In this study, the method that showed the best sensitivity and specificity for muscle strength was the classical Baumgartner definition of sarcopenia but with a 10th percentile cutoff [G3]. Interestingly enough, no method presented sufficient sensitivity and specificity for physical performance. Although there was some early concern about the TGUG test-retest reliability further studies demonstrated that TGUG is a good measure of physical capability in older adults. Nevertheless, the lack of association between sarcopenia and the TGUG
test has been found by others.\textsuperscript{40,41} Merriwether et al.\textsuperscript{41} have studied 154 community-dwelling older adults (72\% women) and found no relationship between sarcopenia defined as Baumgartner suggested and physical performance.

Although the criterion suggested by Newman [G5] is useful to minimize misdiagnoses in the obese population,\textsuperscript{42} it appears to super-classify individuals with normal muscle strength as sarcopenic in this study. Furthermore, the proportion of subjects with normal muscle strength classified as sarcopenic by this criterion was considerably high in our study. On the other hand, all other proposed criteria in the present study classified as normal (non-sarcopenic) at least 20\% of individuals with diminished muscle strength. This result implies that the use of the measurement of muscle mass by densitometry method alone is insufficient to diagnose sarcopenia as suggested by EWGSOP and IWG.

This study has some limitations: firstly, the small sample size. Although it was calculated based on the average appendicular lean mass described in the previous studies,\textsuperscript{39} it could not have enough power to find small differences between the criteria. Secondly, functional capacity was evaluated only by the TGUG test. It was chose the TGUG over the gait speed test because both tests appear to have the same predictive ability\textsuperscript{39} and the TGUG was considered more feasible in clinical practice. Finally, the main restriction of our study is due to the study design. Because of its cross-sectional nature, it can only evaluate associations. Furthermore, the outcomes such as fall, fracture, hospitalization, and/or death were not evaluated.

This study has also some strengths: the young reference population is well matched to the elderly population. To the best of the authors’ knowledge, this is the first study comparing calibration and discrimination of the diagnostic criteria of sarcopenia. These statistical procedures are important to establish the best diagnostic criterion of sarcopenia.

CONCLUSIONS

The best diagnostic criterion in this study was the criterion proposed by Baumgartner with the 10\textsuperscript{th} percentile of the young women as cutoff [G3]. Despite the fact that all diagnostic operational definitions in the present study were associated with muscle strength, recalibration showed that all of them miss subjects with decreased muscle strength. The clinical implications of these findings are not clear but some sarcopenic women may be underdiagnosed with all proposed criteria.

Large epidemiological studies that evaluate the calibration and discrimination of these diagnostic criteria on hard outcomes are needed to answer these questions.

REFERENCES


