An Evidence-Based Comparison of Operational Criteria for the Presence of Sarcopenia

Thuy-Tien Dam, Katherine W. Peters, Maren Fragala, Peggy M. Cawthon, Tamara B. Harris, Robert McLean, Michelle Shardell, Dawn E. Alley, Anne Kenny, Luigi Ferrucci, Jack Guralnik, Douglas P. Kiel, Steve Kritchevsky, Maria T. Vassileva, and Stephanie Studenski

1Department of Medicine, Columbia University College of Physicians and Surgeons, New York, New York.
2California Pacific Medical Center Research Institute, San Francisco.
3Intramural Research Program, National Institute of Aging, National Institutes of Health, Bethesda, Maryland.
4Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts.
5Department of Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore.
6Department of Medical Sciences, University of Central Florida, Orlando.
7The Sticht Center on Aging and Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina.
8The Biomarkers Consortium, Foundation for the National Institutes of Health, Bethesda, Maryland.
9Department of Medicine, School of Medicine, University of Pittsburgh, Pennsylvania.

Address correspondence to Thuy-Tien Dam, MD, Columbia University, 5141 Broadway Avenue, 3 Field West, Room 13, New York, NY 10034. Email: td2265@columbia.edu

Background. Several consensus groups have previously published operational criteria for sarcopenia, incorporating lean mass with strength and/or physical performance. The purpose of this manuscript is to describe the prevalence, agreement, and discrepancies between the Foundation for the National Institutes of Health (FNHI) criteria with other operational definitions for sarcopenia.

Methods. The FNIH Sarcopenia Project used data from nine studies including: Age, Gene and Environment Susceptibility-REykjavík Study; Boston Puerto Rican Health Study; a series of six clinical trials from the University of Connecticut; Framingham Heart Study; Health, Aging, and Body Composition Study; Invecchiare in Chianti; Osteoporotic Fractures in Men Study; Rancho Bernardo Study; and Study of Osteoporotic Fractures. Participants included in these analyses were aged 65 and older and had measures of body mass index, appendicular lean mass, grip strength, and gait speed.

Results. The prevalence of sarcopenia and agreement proportions was higher in women than men. The lowest prevalence was observed with the FNIH criteria (1.3% men and 2.3% women) compared with the International Working Group and the European Working Group for Sarcopenia in Older Persons (5.1% and 5.3% in men and 11.8% and 13.3% in women, respectively). The positive percent agreements between the FNIH criteria and other criteria were low, ranging from 7% to 32% in men and 5% to 19% in women. However, the negative percent agreement were high (all >95%).

Conclusions. The FNIH criteria result in a more conservative operational definition of sarcopenia, and the prevalence was lower compared with other proposed criteria. Agreement for diagnosing sarcopenia was low, but agreement for ruling out sarcopenia was very high. Consensus on the operational criteria for the diagnosis of sarcopenia is much needed to characterize populations for study and to identify adults for treatment.

Key Words: Muscle—Sarcopenia—Lean mass.

Received June 19, 2013; Accepted January 14, 2014

Decision Editor: Roger Fielding, PhD

LOW muscle mass and weakness are potential contributors to disability in older persons. Although the term “sarcopenia” has become widespread, the criteria for an operational definition vary among studies and experts. Initial work on defining sarcopenia was based on measures of muscle mass alone, and the prevalence of sarcopenia when compared with a young reference population ranged between 13% and 24% among adults younger than 70 years to more than 50% among adults older than 80 years (1). However, a growing body of research suggests that there is a disconnect between muscle mass and strength. Thus, recent definitions of sarcopenia have incorporated elements...
of strength and physical performance in addition to muscle mass in the criteria for sarcopenia (2–5). However, these consensus statements were based on expert opinions and lacked access to large data sets to validate their recommendations. Thus, the goal of the Foundation for the National Institutes of Health (FNIH) Sarcopenia Project was to create a data-driven set of criteria for clinically relevant weakness and low lean mass using pooled data from multiple studies.

This is the fifth report of the FNIH Sarcopenia Project. The first manuscript describes the rationale for the FNIH Sarcopenia Project and characteristics of the participating studies. The second and third manuscripts describe in detail the development of cutpoints for weakness and low lean mass; and the fourth manuscript demonstrates the predictive validity of these cutpoints. The purpose of the analyses presented here is to compare the criteria developed by the FNIH project to other published criteria, in order to assess prevalence, agreement, and discrepancies between candidate criteria. Our goal is to provide data-driven evidence to the field in order to advance professional consensus regarding clinically relevant cutpoints and terminology.

**Methods**

**Participants**

The studies participating in the FNIH Sarcopenia Project are described in the first manuscript in this series (6). They include: Age, Gene and Environment Susceptibility-Reykjavik Study (AGES) (7); Boston Puerto Rican Health Study (BPRHS) (8); six clinical trials at University of Connecticut (UCONN) (9–14); Framingham Heart Study (FHS) Original cohort (15) and the Offspring cohort (16); Health, Aging, and Body Composition Study (HABC) (17); Invecchiare in Chianti (InChianti) (18); Osteoporotic Fractures in Men Study (MrOS) (19,20); Rancho Bernardo Study (RBS) (21); and Study of Osteoporotic Fractures (SOF) (22,23). To be included in these analyses, participants must be aged 65 and older and must have completed, at a single time point, the following measures: objectively measured body mass index (BMI), appendicular lean mass (ALM: sum of lean mass in the arms and legs), grip strength, and gait speed. Participants from RBS and AGES were excluded because RBS did not measure walking speed and AGES measured body composition with bioelectrical impedance (BIA). A total of 7,113 men and 2,950 women were included in the analyses presented here.

**Measurement of Lean Mass, Strength, and Performance**

Gait speed was measured as the length of the walking course (4 or 6 m) divided by the time it took participants to walk the course at their usual pace. Walking courses that were longer or shorter were converted to a speed that would have been achieved on a 4- or 6-m course using previously published equations (24). If more than one test was administered, the average gait speed (m/s) was used. Grip strength was measured by a handheld dynamometer, and the maximum value of either hand was analyzed. Total body fat mass and total bone-free lean mass (kg) were acquired using dual energy x-ray absorptiometry (DXA) on Hologic (Waltham, MA) or Lunar/GE Healthcare (Madison, WI) machines.

**Operational Definitions of Lean Mass, Strength, and Other Factors**

The FNIH cutpoints for grip strength and lean mass, derived from classification and regression tree analysis, were reported in two accompanying articles (25,26). Participants with gait speed less than or equal to 0.8 m/s were classified as having slow walking speed. Men with grip strength less than 26 kg and women with a grip strength less than 16 kg were defined as weak (25). We used ALM divided by body size (ALM_{BMI}) to determine lean mass; men with ALM_{BMI} less than 0.789 and women with ALM_{BMI} less than 0.512 were classified as low lean mass (26). We also examined our alternative definitions for low lean mass using absolute ALM (not corrected for body size); men with ALM less than 19.75 kg and women with ALM less than 15.02 were classified as low lean mass (data not shown). Using these cutpoints, we examined two possible FNIH definitions: (i) clinically relevant weakness and low lean mass (low grip strength + low ALM_{BMI}) or (ii) clinically relevant slowness with weakness and low lean mass (slow gait speed + low grip strength + low ALM_{BMI}). These definitions were used to compare with other proposed definitions for sarcopenia.

Several groups have previously published operational criteria to define sarcopenia, including: (i) International Working Group (IWG) (4); (ii) European Working Group on Sarcopenia Older Persons (EWGSOP) (3); (iii) European Society for Clinical Nutrition and Metabolism Special Interest Group on cachexia-anorexia in chronic wasting diseases (ESPEN) (2); and (iv) Society of Sarcopenia, Cachexia, and Wasting Disorders (SCWD) (5). These recommendations combined lean mass with a strength and/or physical performance measure. The EWGSOP suggested that sarcopenia be defined as low lean mass + low strength and/or low performance. The EWGSOP differentiated presarcopenia (low mass) from sarcopenia (low mass + low strength or low performance) and severe sarcopenia (low mass + low strength + low performance). Several possible performance measures (grip strength, chair stand, gait speed), lean mass assessment methods (DXA, bioimpedance, computed tomography, and magnetic resonance imaging), and different cutpoints were suggested by the EWGSOP. For the FNIH analyses, we used ALM by DXA, grip strength, and gait speed as measures of muscle mass, strength, and physical performance (3) and used similar cutpoints that were recently published to operationalize the EWGSOP criteria (27–29). The IWG recommended gait
speed as a measure of physical performance and defined gait speed less than 1 m/s as slow (4). The recommendations for ESPEN and SCWD were similar to EWGSOP and IWG, thus were not analyzed separately. The operational cutpoints used in these analyses are summarized in Table 1.

**Statistical Analysis**

Data were analyzed for men and women separately. Descriptive statistics were examined across sets of criteria. Sensitivity and specificity were not determined because they are not applicable in the absence of a gold standard criterion to define sarcopenia. Furthermore, positive predictive value, negative predictive value, and likelihood ratios cannot be computed because a participant’s status (as determined by a reference standard) is unknown. Therefore, we described the agreement between the FNIH criteria with other proposed criteria using several different statistical measures (30), including: (i) positive percent agreement: the proportion of participants who were categorized as having the condition by both the FNIH criteria and a second set of criteria divided by the number of participants who were categorized as having the condition by the second set of criteria. This is analogous to a sensitivity calculation; (ii) negative percent agreement: the proportion of participants who were categorized as not having the condition by both the FNIH criteria and a second set of criteria divided by the number of participants who were categorized as not having the condition by the second set of criteria. This is analogous to a specificity calculation; and (iii) Cohen’s kappa (κ). Kappa (κ) values less than 0.40 are considered poor reliability, 0.40–0.75 are considered fair-to-good reliability, and greater than 0.75 are considered excellent reliability (30).

**RESULTS**

These analyses included 10,063 participants (7,113 men and 2,950 women). Mean ± SD for gait speeds were 1.23±0.24 m/s and 0.97±0.24 m/s; grip strengths were 40.7±8.8 kg and 21.1±5.9 kg; ALM_BMI were 0.88±0.11 kg and 0.60±0.10 kg; and ALM were 40.23±8.92 kg and 20.58±5.81 kg, for men and women, respectively. The proportion of participants who fell below the FNIH cutpoints for gait speed, grip strength, lean mass, and the multiple combinations are presented in Figure 1. Compared with ALM divided by height squared (ALM/ht²), participants with low lean mass by the FNIH criteria (ALM_BMI) were heavier with higher BMI and ALM; more participants were overweight or obese and reported a history of diabetes and heart failure (Supplementary Table 1). Despite their larger body size, adults with low lean mass by the FNIH ALM_BMI cutpoint had poorer physical function including weaker grip strength, slower mean walking speed, and a higher percentage with walking speed less than or equal to 0.8 m/s compared with participants with low lean mass defined by ALM/ht².

The various proposed operational definitions and prevalence of sarcopenia are presented in Table 1. Note that primary indicator of lean mass for the FNIH project was ALM_BMI; whereas all other proposed criteria used ALM/ht². These sets of candidate definitions largely differed from

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Physical Performance</th>
<th>Muscle Strength</th>
<th>ALM</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Men</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(n = 7,113)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operational Definition</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
</tr>
<tr>
<td>Foundation of NIH Sarcopenia Project</td>
<td></td>
</tr>
<tr>
<td>Weakness and low lean mass</td>
<td>Grip strength</td>
</tr>
<tr>
<td>Men: &lt;26 kg</td>
<td>Men: &lt;0.789</td>
</tr>
<tr>
<td>Women: &lt;16 kg</td>
<td>Women: &lt;0.512</td>
</tr>
<tr>
<td>Slowness with weakness and low lean mass</td>
<td>Grip strength</td>
</tr>
<tr>
<td>Men: &lt;26 kg</td>
<td>Men: &lt;0.789</td>
</tr>
<tr>
<td>Women: &lt;16 kg</td>
<td>Women: &lt;0.512</td>
</tr>
<tr>
<td>International Working Group</td>
<td>Grip strength</td>
</tr>
<tr>
<td>Gait speed: ≤1.0 m/s</td>
<td>ALM_BMI</td>
</tr>
<tr>
<td>Men: ≤7.23 kg/m²</td>
<td>Men: ≤5.67 kg/m²</td>
</tr>
<tr>
<td>Women: ≤5.67 kg/m²</td>
<td>Women: ≤5.67 kg/m²</td>
</tr>
<tr>
<td>European Working Group on Sarcopenia Older Persons</td>
<td></td>
</tr>
<tr>
<td>Sarcopenia</td>
<td>Grip strength</td>
</tr>
<tr>
<td>Gait speed: ≤0.8 m/s</td>
<td>ALM_BMI</td>
</tr>
<tr>
<td>Men: ≤7.23 kg/m²</td>
<td>Women: ≤5.67 kg/m²</td>
</tr>
<tr>
<td>Severe sarcopenia</td>
<td>Grip strength</td>
</tr>
<tr>
<td>Gait speed: ≤0.8 m/s</td>
<td>ALM_BMI</td>
</tr>
<tr>
<td>Men: ≤7.23 kg/m²</td>
<td>Women: ≤5.67 kg/m²</td>
</tr>
</tbody>
</table>

*Note: ALM_BMI = ratio of appendicular lean mass over body mass index; ALM/ht² = ratio of appendicular lean mass over height squared.*
each other in regards to the cutpoint for slow gait speed and whether or not to include a measure of weakness. The prevalence of sarcopenia was higher in women than men. In men, the prevalence was 1.3% for the FNIH criteria, 5.1% for IWG, and 5.3% for EWGSOP. In women, the prevalence was 2.3% for FNIH, 11.8% for IWG, and 13.3% for EWGSOP. When we included gait speed with grip strength and lean mass in the FNIH definition, the proportion of participants who met all three was lower: 0.5% in men and 1.8% in women. Furthermore, the prevalence of severe sarcopenia by the EWGSOP was 0.7% in men and 2.9% in women. In general, the FNIH criteria, compared with other proposed definitions, identified participants who were older with higher BMI and higher lean mass, but were functionally more impaired, including a higher proportion with slow gait and inability to rise from a chair (Supplementary Tables 2 and 3).

Table 2 presents agreement proportions between the various criteria. In general, agreement was higher in men than in women and higher between the FNIH and EWGSOP criteria. The positive percent agreements between the FNIH criteria and other criteria were low, ranging from 4.3%
to 32% in men and 4.0% to 19.9% in women. However, the negative percent agreement was very high (all >95%). Our results demonstrated that kappa values were modest, with a range of 0.11–0.53 in men and 0.04–0.17 in women. Between the IWG and EWGSOP definitions, the positive percent agreement for sarcopenia was 52.1% and 61.7% in men and women, respectively. The negative percent agreement for sarcopenia was 97.1% and 96.4% in men and women, respectively.

**Discussion**

The prevalence of sarcopenia varies greatly depending on the criteria used for diagnosis. Based on the presence of lean mass alone, initial prevalence of sarcopenia ranged between 7% and 50%. The large range was due to differences in the criterion used to diagnose sarcopenia, including (i) definitions that only included lean mass with or without correction for height (1,31), body mass (32), or body height and body fat (33); (ii) methodological differences to measure muscle mass (DXA or BIA) (1,32–35); and (iii) differences in the reference population used to establish normative data. In this study, the prevalence of sarcopenia incorporating both low lean mass and poor function were much lower—between 0.5% and 5.3% in men and 1.8% and 13.3% in women compared with definitions based on muscle mass alone.

When comparing the FNIH, EWGSOP, and IWG criteria in this pooled sample, the positive percent agreements were low, but the negative percent agreements were high (all >95%), suggesting that there was good agreement on the absence of the condition. The prevalence of sarcopenia was highest with the EWGSOP criteria in both men and women (5.3% and 13.3%, respectively). Given that the FNIH criteria were more restrictive, it was not surprising that the prevalence of sarcopenia was lower than either the EWGSOP or IWG criteria. Adults considered sarcopenic by both the FNIH and EWGSOP or by both FNIH and IWG criteria had little overlap as evidenced by the modest positive percent agreement and kappa values. However, the FNIH criteria had better concordance with the EWGSOP criteria. This was because the FNIH and EWGSOP had similar conceptual frameworks and both included measures of lean mass, strength, and/or performance. We believe the lack of agreement between these two criteria is explained by differences in the participants categorized as having low lean mass: The FNIH criteria used $\text{ALM}_{\text{BMI}}$ whereas both EWGSOP and IWG criteria used $\text{ALM}_{\text{ht}}$. Although adults with $\text{ALM}_{\text{ht}}$ less than 5.67 kg/m² in women and less than 7.23 kg/m² in men did have low lean mass relative to a young reference population, this amount of lean mass was associated with faster gait speeds, stronger grip strength, and lower rates of obesity compared with participants who had low lean mass by $\text{ALM}_{\text{BMI}}$. On the other hand, adults who met the FNIH criteria for low lean mass ($\text{ALM}_{\text{BMI}}$) were slightly more impaired with slower walking speeds and lower grip strength even though they had higher BMI and ALM, and proportionately more obesity. These data suggests that $\text{ALM}_{\text{BMI}}$ lean mass corrected for body size, is a good discriminator for low lean mass and is likely capturing adults who were unable to generate enough strength or function relative to their body size (sarcopenic obesity), and $\text{ALM}_{\text{BMI}}$ may be a good measure for low muscle quality or efficiency.

In this study, there were several limitations. In particular, the prevalence and agreement rates may have been affected by several factors, including (i) the conceptual model, (ii) the strength or performance measure and cutpoints used, (iii) the method of assessment, and (iv) the study population. First, mobility is an important predictor and indicator of functional independence and disability. Therefore, our conceptual framework and statistical approach was based on using mobility impairment as the clinically relevant functional state to determine meaningful weakness and low lean mass. In this series, we provided two different possible FNIH criteria: (i) weakness + low lean mass or (ii) slowness + weakness + low lean mass. Although this framework resulted in a low prevalence of sarcopenia, we chose to combine gait speed with grip strength and lean mass because the goal of the FNIH Sarcopenia Project was to develop criteria that were conservative with few false positives in order to identify individuals who were clearly abnormal. Second, there were many different measures of strength or performance that could have been used (eg, grip strength, leg power, dynamic leg strength, short physical performance, gait speed, and chair stand). However, we chose gait speed as a measure of mobility because it was available in all but one of the pooled studies, it has been reliably measured in clinical studies, and has been closely linked to function. Additionally, we chose gait speed less than or equal to 0.8 m/s because it has been associated with survival (36), and the prevalence of gait speed less than or equal to 0.6 m/s was very rare in our pooled sample. In one study, among 70- to 80-year-old Finnish women, the prevalence of sarcopenia was 2.7% vs. 0.9% when gait speed less than 1.0 m/s was used instead of gait speed less than 0.8 m/s (28). We chose grip strength as a measure for muscle strength because it is easy to use in both clinical and community settings and was available across all of the studies participating. Selection of other measures like short physical performance battery or chair stand could have increased or decreased the prevalence. However, among the participating studies that had chair stands in our pooled data set, the proportion of participants unable to complete a chair stand was similar to the proportion with gait speed less than or equal to 0.8 m/s. Since the EWGSOP consensus statement recommended several different cutpoints for different measures and did not mandate specific measures or cutpoints, the prevalence of sarcopenia using the EWGSOP criteria could have been influenced by our operational decisions. In fact, recent publications operationalizing the
EWGSOP and IWG criteria have reported different prevalence rates. For example, Landi and coworkers used the lowest tertile of skin fold thickness for low lean mass, the prevalence was 21.8% among Italians aged 80–85 years (29). In another study, among 103 community-dwelling men in the UK Hertfordshire Sarcopenia Study, the prevalence for sarcopenia was 6.8% and 7.8% when the lowest tertile of DXA-based lean mass was used instead of the lowest tertile of skin fold–based fat free mass, respectively (27). Furthermore, the positive percent agreement between the FNIH and EWGSOP are likely overestimated because we used similar gait speeds (<0.8 m/s) and grip strength cutpoints (<26 vs <30 kg in men and <16 vs <20 kg in women). Third, different brands/methods of DXA were used in the studies participating in the FNIH Sarcopenia Project, and therefore potential bias may lead to different results. However, we tried to account for these differences between studies by using a random effects term in all our analyses that evaluated the association between our definitions and outcomes.

Finally, the FNIH Sarcopenia Project pooled data sets from nine different studies for pooled analyses—the largest to ever be studied in this area and are generalizable because the data set had broad representation of community-dwelling older adults. However, the pooled data set included primarily healthy community-dwelling populations with few comorbidities. This prevalence may be lower compared with more vulnerable populations (eg, assisted living, nursing home, or hospitalized adults), where disability rates are higher. However, these vulnerable populations are more heterogeneous, and the factors contributing to slow gait are more numerous including cognition, osteoarthritis, pain, disuse atrophy, and cachexia. Whether the relationship between lean mass, strength, and mobility are the same in more vulnerable populations is not known.

This large variation in the prevalence may lead to different conclusions and implications for treatment. In particular, it is not clear whether treatment of weakness and low lean mass, especially with interventions that only target improving muscle strength and mass, in different populations is beneficial. Thus, the work presented in this series is a work in progress, and many more questions and studies are needed. However, the FNIH Sarcopenia Project provides evidence-based and data-driven cutpoints that will help the field come to a consensus on a diagnostic criteria. Future studies will need to address whether: (i) the prevalence of low lean mass, weakness, and poor physical performance with the FNIH criteria is higher among different populations; (ii) the associations between mass, strength, and disability are similar or stronger in more disabled or sick populations; (iii) the FNIH criteria is useful in identifying participants for clinical trials; and (iv) these criteria allow clinicians to recognize and potentially treat this disabling condition. We envision that the FNIH criteria for clinically relevant weakness and low lean mass might be used to plan prevention studies in which older persons with weakness and low lean mass, but have not yet developed mobility limitations, would receive interventions designed to reduce the incidence or increase the time to onset of mobility impairment. Alternatively, the FNIH criteria for clinically relevant slowdown with weakness and low lean mass may be used to identify candidates who already have mobility impairment, weakness, and low lean mass for recruitment in treatment studies that may look at outcomes like maintaining independence, preventing disability, delaying transitions from home to long-term care, quality of life, and/or survival.

In conclusion, the current work taps into the expertise in the field and utilizes the largest sample of community-dwelling older adults to build upon and validate prior recommendations. These data provide comparison of the different definitions and suggest that the definitions have good negative, but poor positive agreement. Thus, future studies should examine the predictive validity of these different definitions with important clinical outcomes (eg, disability, mortality) among different populations that may benefit from the diagnoses of sarcopenia.

Supplementary Material

Supplementary material can be found at: http://biomedgerontology.oxfordjournals.org/

Funding

Funding support for the conference and the work of the consortium was provided by the National Institute on Aging (U13AG04158 and P30 AG024827), the Food and Drug Administration, and through grants from the Foundation of the National Institute of Health (FNIH) made possible by funding from Abbott Nutrition, Amgen, Eli Lilly, Merck, Novartis, and the Dairy Research Institute. This research was supported in part by the Intramural Program of the NIH, National Institute on Aging. Additional acknowledgments for each contributing cohort and members of the FNIH Sarcopenia Project can be found in an online supplement. Dr. Dam was supported by the National Institute on Aging through grant number K23 AG040168.

References


