Body Composition Tools for Assessment of Adult Malnutrition at the Bedside: A Tutorial on Research Considerations and Clinical Applications

Carrie P. Earthman, PhD, RD, LD

Abstract
Because of the key role played by the body’s lean tissue reserves (of which skeletal muscle is a major component) in the response to injury and illness, its maintenance is of central importance to nutrition status. With the recent development of the Academy of Nutrition and Diетetics/American Society for Parenteral and Enteral Nutrition diagnostic framework for malnutrition, the loss of muscle mass has been recognized as one of the defining criteria. Objective methods to evaluate muscle loss in individuals with acute and chronic illness are needed. Bioimpedance and ultrasound techniques are currently the best options for the clinical setting; however, additional research is needed to investigate how best to optimize measurements and minimize error and to establish if these techniques (and which specific approaches) can uniquely contribute to the assessment of malnutrition, beyond more subjective evaluation methods. In this tutorial, key concepts and statistical methods used in the validation of bedside methods to assess lean tissue compartments are discussed. Body composition assessment methods that are most widely available for practice and research in the clinical setting are presented, and clinical cases are used to illustrate how the clinician might use bioimpedance and/or ultrasound as a tool to assess nutrition status at the bedside. Future research needs regarding malnutrition assessment are identified.

Keywords

body cell mass; nutrition status; malnutrition; fat-free mass; lean body mass; muscle mass; muscle loss; skeletal muscle; intracellular water; bioimpedance; spectroscopy; impedance ratio; phase angle

There has been long-standing interest in the assessment of lean tissue as a key parameter of nutrition status. Specifically, lean tissue mass (of which skeletal muscle is a major component) plays a central role in the body’s ability to respond to acute and chronic illness by serving as a vitally important reservoir of amino acids that can be redirected to the tasks of injury repair and the immune response when needed. Albeit highly functional, the process of protein catabolism under these circumstances leads to loss of lean tissue. The loss of lean tissue and overt sarcopenia (defined as loss of muscle mass and strength) that often occur in the face of chronic and acute illness carry significant ramifications in terms of clinical outcomes, including increased incidence of infections, increased length of stay, and increased morbidity and mortality. Indeed, we now have published cutpoints defining sarcopenia by several body composition assessment methods (see Table 1); these are also used as part of the defining characteristics of cachexia, a term used to describe the significant weight loss, protein catabolism, and muscle and fat tissue loss that occur due to underlying disease processes. Furthermore, with the recent publication of the Academy of Nutrition and Diетetics (AND)/American Society for Parenteral and Enteral Nutrition (A.S.P.E.N.) consensus statement defining diagnostic criteria for malnutrition, there has been renewed focus on the assessment of muscle loss as a key component of nutrition status. As the consensus malnutrition diagnostic framework has been developed, much attention has been paid to the assessment of muscle loss through physical examination techniques; however, the subjective nature of physical examination is a potential limitation. One of the major limitations of any subjective method is poor reliability due to interobserver and intraobserver variability. Although intensive and standardized training programs may be able to minimize that variability, the potential for error remains high as the technique is broadly applied across medical centers. Furthermore, lean tissue loss can precede overt weight loss and may be masked by excess extracellular water (ECW), thus making it difficult to detect through visual techniques. A recent

From the 1Department of Food Science and Nutrition, University of Minnesota–Twin Cities, St Paul, Minnesota.

Financial disclosure: C.P.E. received loaner bioimpedance devices and a small gift grant in support of her research from Bodystat Ltd (UK), a company that manufactures bioimpedance devices.

Received for publication February 6, 2015; accepted for publication June 17, 2015.

Corresponding Author:
Carrie P. Earthman, PhD, RD, LD, Professor, Department of Food Science & Nutrition, University of Minnesota, Food Science & Nutrition 225, 1334 Eckles Ave, St Paul, MN 55108-6099, USA. Email: cearthma@umn.edu
Skeletal muscle index (height) by SF-BIA clinical implications. For all these reasons, a method to problem of sarcopenia in the presence of obesity has significant on sex-specific lowest 20% of study group in the Health, Aging, and Body Composition Study.

Skeletal muscle index (SMI) percent values from young adults from the National Health and Nutrition Examination Survey III (NHANES III), and cutpoints for sex, male = 1 and female = 0, divided by weight in kilograms multiplied by 100. Cutpoints for class I sarcopenia are 1–2 SD below the mean skeletal muscle mass (kg) = \[
\text{[(Height2/R × 0.401) + (Sex × 3.825) + (Age × –0.071)]} + 5.102,\]
where height is in centimeters, R is resistance in ohms, and cutpoints are based on mortality using optimum stratification.

CT, computed tomography; DXA, dual-energy X-ray absorptiometry; SF-BIA, single-frequency bioelectrical impedance analysis.

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference Cutpoint Males</th>
<th>Reference Cutpoint Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendicular skeletal muscle index by DXA</td>
<td>&lt;7.26</td>
<td>&lt;5.45</td>
</tr>
<tr>
<td>(Baumgartner et al(^{10}), (^{2}) kg/m(^2))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendicular lean muscle index by DXA</td>
<td>&lt;7.25</td>
<td>&lt;5.67</td>
</tr>
<tr>
<td>(Delmonico et al(^{9}), (^{3}) kg/m(^2))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumbar skeletal muscle index by CT (Prado et al(^{11}), (^{4}) cm(^2)/m(^2))</td>
<td>&lt;52.4</td>
<td>&lt;38.5</td>
</tr>
<tr>
<td>Skeletal muscle index (%weight) by SF-BIA</td>
<td>Class II: &lt;31</td>
<td>Class II: &lt;22</td>
</tr>
<tr>
<td>(Janssen et al(^{12}), (^{5}) %)</td>
<td>Class I: 31–37</td>
<td>Class I: 22–28</td>
</tr>
<tr>
<td>Skeletal muscle index (height) by SF-BIA</td>
<td>High risk: ≤8.50</td>
<td>High risk: ≤5.75</td>
</tr>
<tr>
<td>(Janssen et al(^{13}), (^{6}) kg/m(^2))</td>
<td>Moderate risk: 8.51–10.75</td>
<td>Moderate risk: 5.76–6.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Reference Cutpoint Males</th>
<th>Reference Cutpoint Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendicular skeletal muscle index by DXA</td>
<td>&lt;7.26</td>
<td>&lt;5.45</td>
</tr>
<tr>
<td>(Baumgartner et al(^{10}), (^{2}) kg/m(^2))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendicular lean muscle index by DXA</td>
<td>&lt;7.25</td>
<td>&lt;5.67</td>
</tr>
<tr>
<td>(Delmonico et al(^{9}), (^{3}) kg/m(^2))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lumbar skeletal muscle index by CT (Prado et al(^{11}), (^{4}) cm(^2)/m(^2))</td>
<td>&lt;52.4</td>
<td>&lt;38.5</td>
</tr>
</tbody>
</table>

This is not an exhaustive list of cutpoints from the literature.

Appendicular skeletal muscle index calculated as the sum of arm and leg lean soft tissue mass divided by height in meters squared, and cutpoints are 2 SD below the mean values from young adults in the Rosetta Study.

Appendicular lean muscle index calculated as the sum of arm and leg lean soft tissue mass divided by height in meters squared, and cutpoints are based on sex-specific lowest 20% of study group in the Health, Aging, and Body Composition Study.

Lumbar 3 (L3) skeletal muscle index was calculated as the total area of L3 skeletal muscle in centimeters squared, divided by height in meters squared, and cutpoints are based on mortality using optimum stratification.

Skeletal muscle index (%weight) calculated by the Janssen SF-BIA skeletal muscle mass equation developed from magnetic resonance imaging data\(^{17}\). Skeletal muscle mass (kg) = \[
[(\text{Height}^2/R \times 0.401) + (\text{Sex} \times 3.825) + (\text{Age} \times –0.071)] \times 5.102,\]
where height is in centimeters, R is resistance in ohms, and for sex, male = 1 and female = 0, divided by weight in kilograms multiplied by 100. Cutpoints for class I sarcopenia are 1–2 SD below the mean skeletal muscle index (SMI) percent values from young adults from the National Health and Nutrition Examination Survey III (NHANES III), and cutpoints for class II sarcopenia are 2 SD below the mean SMI percent values from young adults.

Appendicular muscle index (height) calculated by the above Janssen skeletal muscle mass SF-BIA equation, divided by height in meters squared, and cutpoints are based on high and moderate risk for physical disability in adults aged ≥60 years from the NHANES III (1988–1994).

This article by Sheean et al\(^{19}\) would attest to this. They reported that 60% of individuals with respiratory failure who were deemed normally nourished at intensive care unit (ICU) admission by subjective global assessment (administered by experienced clinicians) were actually sarcopenic as defined by computerized tomography (CT). Approximately 33% of individuals who were misclassified were overweight or obese. Although there is no clear consensus yet on how best to define and identify it,\(^{21,22}\) the problem of sarcopenia in the presence of obesity has significant clinical implications.\(^{7,9,23}\) For all these reasons, a method to objectively and reliably evaluate lean tissue loss at the bedside is highly desirable but has been somewhat elusive.

In the first part of this tutorial, basic terminology around the assessment of lean tissue body composition is introduced. In the second part, key concepts around validity of field (which we will henceforward term bedside) techniques, as they compare with reference methods, are presented. Subsequently, the statistical methods that are typically used in the validation of bedside techniques are reviewed to gain perspective on the interpretation and application of the research literature. In the third and fourth parts, body composition assessment methods that are most widely available for practice and research in the clinical setting are subsequently discussed, including multiple dilution, dual-energy X-ray absorptiometry (DXA), and CT (reference methods), as well as bioimpedance and ultrasound (bedside methods). In the fifth part, clinical cases are used to illustrate how the clinician might ultimately use bioimpedance and ultrasound to assess lean tissue and/or nutrition status at the bedside. Finally, the sixth part addresses gaps in the research literature around malnutrition assessment where additional clinical research studies are needed.

### Introduction to Lean Tissue Terminology

Multiple terms to describe the lean tissue compartment of the body, based on various conceptual models of body composition, have been well described elsewhere.\(^{24,25}\) At the most basic level, the 2-component model of body composition describes the body as the sum of fat mass and fat-free mass (FFM). FFM thus is a rather broad term that includes skeletal muscle, non-skeletal muscle, organs, connective tissue, total body water (TBW, including ECW and intracellular water [ICW]), and bone. FFM can be estimated from TBW measurements using the following equation:

\[
\text{TBW (kg)}/0.73 = \text{FFM (kg)}, \text{based on the standard hydration constant for FFM.}\]

This hydration constant has been shown to be remarkably stable under euhydrated, healthy conditions but varies under clinical conditions where hydration may be altered and has been shown to be somewhat higher (>0.75) in individuals with extreme obesity.\(^{27,28}\) The lean body mass (LBM), which is more appropriately termed lean soft tissue (LST) when measured by DXA, is a slightly more specific compartment that includes all of the aforementioned components except for bone.\(^{29}\) The body cell mass (BCM) is a term originally described by Moore and
Boydend as the total mass of cells in the body that consume oxygen and produce work; it is the nonfat cellular portion of tissues, of which the primary components are skeletal muscle, organ tissue mass, blood, and the brain. As such, the BCM has been considered by many to be the primary lean tissue compartment to assess nutrition status and to serve as a key target for nutrition interventions.

There is no direct way to quantify BCM, but there are several ways it can be estimated, including from total body nitrogen determined by neutron activation analysis (NAA), total body potassium through total body potassium counting (TBK), and ICW measured by multiple dilution.  Given that NAA exposes individuals to radiation, and both NAA and TBK are highly technical and expensive, ICW through multiple dilution is one of the more accessible reference methods to clinical researchers and depends on the underlying observations that a large proportion of the BCM is represented by the water content of cells, and acute changes in body protein occur primarily within cells, and these changes are generally accompanied by changes in ICW. Nevertheless, despite differences in their level of specificity to core functional processes, all of the aforementioned lean tissue compartments, including FFM, LST, and BCM (and/or ICW itself), may be measured as the target variable in methods validation, as well as nutrition assessment and intervention studies, depending on the available body composition assessment methods.

Validation of Bedside Techniques to Assess Lean Tissue in Clinical Settings

Several excellent reviews describe the various methods that are available to evaluate lean tissue compartments. Bedside techniques need to be validated against more established reference techniques before they can be applied with confidence in the clinical setting. It should be noted that all methods to evaluate body composition are indirect, requiring various assumptions that may or may not hold true in acute and/or chronic illness, and none is completely free from error. Thus, the term gold standard, which has been used in the past, has been replaced with the term reference to describe the more established techniques against which bedside methods are validated. For the purposes of this discussion, the focus is primarily on bioimpedance and ultrasound techniques as promising currently available bedside assessment options. The most common reference methods against which bioimpedance and ultrasound have been compared in clinical studies include multiple dilution, DXA, and CT. Although magnetic resonance imaging (MRI) is also an important reference method for validation studies, it is generally less widely available to clinician researchers and thus will not be discussed in this tutorial. However, the reader is referred to a recent review that includes discussion of MRI. Before these methods are briefly described, however, it is important to understand the key terminology and statistical approaches used in methods validation.

Basic Terminology Regarding Validity

There are several important concepts to keep in mind as we talk about the validation of bedside methods. At the heart of it, the effort to prove the validity of a bedside method is undertaken to establish that method as an acceptable alternative to a more expensive and/or technical and, usually, less widely available technique. Theoretically speaking, the definition of validity is the extent to which a measurement actually measures the “true” value of a particular body composition parameter; however, given that the “true” quantification of body composition components may only be known through cadaver analysis, and there are no completely error-free “gold standards” among body composition assessment methods applied to living humans, it is virtually impossible to determine that “true” value. The best we can do is to compare the bedside method with the best reference method that is available to measure the compartment of interest, and if the methods produce sufficiently comparable measures with acceptable measurement error, we can then call the bedside method valid. Thus, in most of the body composition literature, the term validity is used more generally and encompasses such concepts as precision, bias, and accuracy, among others. In general, measurement error can be conceived of as falling into 1 of 2 categories: (1) random error due to biological variation and other factors and (2) systematic error that occurs from constant fixed and/or proportional bias. From a practical standpoint, investigators aiming to validate a device or technique should strive to evaluate (1) how close do the values obtained by the bedside method agree with the values obtained by the reference method, (2) how often are the values within an acceptable range of difference, and (3) whether there is a consistent tendency for the bedside method to over- or underestimate the body composition compartment of interest compared with the reference method. To answer these questions, it is important to know the precision of the reference method and to evaluate the precision, accuracy, and bias of the bedside method.

Precision and reliability. The term precision as it applies to body composition measurement is typically used to describe the degree of agreement among repeated measurements by a particular assessment method (ie, how well does a particular method produce the same result on multiple occasions). There is another use of the word that is sometimes seen in the literature, although it deviates from the true definition; precision is sometimes loosely used to describe how well a bedside technique produces measurements sufficiently close to those produced by a reference method in a group of individuals (ie, how variable are individual measures between methods, defined in part by the limits of agreement or 95% confidence interval, which will be discussed later) but not necessarily taking into consideration inter- or intraobserver errors that would typically be used to describe precision in terms of repeatability and reproducibility. We will assume the first definition when precision is discussed and will indicate when the alternative
meaning is intended by using “precision.” Reliability is a term that is used to describe a method that is considered to have high precision. Random error will affect the precision and reliability of a method.41 For a method to be considered reliable, it should have high repeatability and reproducibility. Repeatability is a term that refers to the variability between repeated, independent measurements of a particular variable (eg, FFM) that are made by the same operator or observer using a single device, in the same individual, under the same conditions. Reproducibility, on the other hand, can be used to describe the variability in measurements of a particular variable taken by that particular device operated by different observers in the same institution or in different institutions. Imprecision has been defined as the variability of repeated measurements due to intra- and interobserver measurement differences.42 Interobserver measurement error (ME) refers to the variability in a particular variable by the same method when measurements are made by more than one operator. Intraobserver ME refers to the variability in a particular variable by the same method when more than one measurement is made by the same operator. ME is typically referred to as technical error in the anthropometric literature.42 To calculate intraobserver ME for 2 measurements and interobserver ME for 2 operators, the same equation may be used42:

\[
\text{ME} = \sqrt{\left(\sum D^2\right)/2n}, \text{ where } D \text{ is the difference between the paired measurements and } n \text{ is the number of individuals measured.}
\]

Because ME can sometimes be positively associated with the size of measurement, it is advantageous to convert the absolute ME to a relative ME term, which can be done by calculating the %ME using the following formula42:

\[
(\text{ME} / \text{mean of measurements}) \times 100.
\]

Calculating the %ME in this way is a variation on the measurement of the coefficient of variation (CV), which is typically calculated by dividing the standard deviation (SD) by the mean of the set of measurements and then converting to a percent (ie, \( \text{CV} = \text{SD/mean} \times 100 \)). In fact, it could be said that all expressions of error are ideally described as a percentage relative to the actual size of the compartment being measured, although it is important to remember that there will be differences in relative magnitude between various error terms. A method that is subject to high inter- and intraobserver errors is judged to have low precision. Simple anthropometry using skinfolds and circumferences is an example of a method that has been shown to have low precision, with %ME approaching 8% and CVs as high as 45% for some skinfold measurements.42-44

To summarize, random error (caused by biological variability, as well as errors in the measurement procedure caused by protocol deviations, random instrument error, operator error, and environment, among other factors) as it relates to precision can be expressed in several different ways, but the bottom line is we are attempting to capture the degree of variability in measurements not due to fixed systematic error. You will see derivations of the same concept applied to calculate other error terms as we go along, with minor adjustments to the formulae depending on the statistical approach. Statistical terms that are commonly used to describe reliability and precision include test-retest correlation coefficients (same day or between day), intraclass correlation coefficient (ICC), SD, and CV for a given set of measurements. ICC values approaching 1.0 would indicate low variability between repeated measures of the same subject (ie, low ME). Root mean squared error (RMSE) can also be calculated to estimate random error in a bedside method, and although there are slight deviations that might be applied when using multiple regression analysis,45 RMSE (sometimes termed pure error) can be calculated as follows46:

\[
\text{RMSE} = \sqrt{\left(\sum (Y - X)^2\right)/n}, \text{ where } Y \text{ is the reference measure, } X \text{ is the bedside measure, and } n \text{ is the number of subjects measured.}
\]

Accuracy and bias. The term accuracy is used to indicate the closeness of agreement in a particular variable between 2 assessment methods (ie, how close are the values of a particular variable by a bedside method to those generated by an accepted reference method). The term bias is a general term to describe the systematic error in a method (ie, the difference between the measurements made by the bedside method and those made by the reference method). Bias is typically calculated as the mean of the differences in a particular measurement variable between the 2 methods being compared. Fixed (or constant) bias refers to the type of systematic error that occurs when a method yields measurements of a particular variable that are consistently higher or lower than those taken by the reference method. The method can be highly precise in terms of repeatability and reproducibility yet still yields values that are predictably and consistently lower or higher than those made by a reference method (ie, low accuracy). For example, a single-frequency bioelectrical impedance analysis (SF-BIA) device shown to have high precision can be used to generate data applied to an equation that might have been developed from deuterium dilution data to predict TBW; when that equation-generated TBW value is converted to FFM using standard assumptions (eg, TBW/0.73 = FFM) and those values are subsequently compared with FFM values generated by DXA, it can lead to a type of fixed constant bias (ie, systematic error) termed scaling error. In this case of high precision but low accuracy, it is possible that the fixed bias might be addressed through the use of a correction factor. Unfortunately, errors between methods are rarely that simple to address. Proportional or positional bias refers to errors that are proportional to the value of the variable being measured. An example of this type of error is the proportional bias observed in bioimpedance measurements of lean tissue in individuals with extreme obesity. The errors in ICW and other lean tissue compartments by bioimpedance techniques...
have been well demonstrated to increase with increasing adiposity. 47-49 Thus, the accuracy of a single measurement in a single individual is impossible to ascertain with 100% certainty, but it may be inferred from estimates of the precision of the method as well as from an understanding of the systematic error of the method.

In methods validation, there are 2 primary ways of thinking about accuracy, in part based on the statistical methods used to evaluate it. The first school of thought would argue that a bedside method is accurate if the average of all measurements taken by that method is close to the average of all measurements taken by the reference method (ie, the mean difference between methods is close to 0); the closeness of agreement between methods at the individual level is not taken into consideration. In this line of reasoning, the primary concern is that a method has low bias (ie, a method could have low precision and individual values by the 2 methods could be quite different, and the method could still be called accurate as long as low bias at the mean level is evident). An example of when this definition of accuracy is most commonly (and probably appropriately) applied is in large-scale epidemiologic studies that compare measurements of a variable (eg, FFM) by a field method such as bioimpedance with a reference method such as DXA, in order to establish bioimpedance as an acceptable method for describing population-level mean differences in FFM across different ethnic groups. The statistical approaches used to establish accuracy based on this definition include correlation and paired t test statistics. We will revisit these methods in the following sections.

The second school of thought would argue that a technique cannot be considered accurate unless measurements of a particular variable by 2 methods agree closely with one another across all individuals. The statistical approach typically used to establish accuracy based on this definition is Bland-Altman analysis, including calculations of various error terms and limits of agreement (defined as mean difference between methods ±1.96 SD) that can help us to determine if a method is truly acceptable to replace another more established method for the individual at the bedside. There are no clear guidelines on what should be considered an acceptable range for the limits of agreement to define a method as valid. It is up to the individual conducting the study to determine what defines acceptable in terms of the width of the limits of agreement. Therein lay one of several difficulties in the interpretation of validation studies. We will get back to Bland-Altman analysis and its application and interpretation in a subsequent section.

Sensitivity and specificity. Although not typically used to describe methods comparisons, for our purposes, it is useful to review what is intended by the terms sensitivity and specificity, particularly as they pertain to the identification of malnutrition by a particular technique. Sensitivity can be defined as the percentage of truly malnourished individuals who are correctly identified as such, calculated as follows:

\[
\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \times 100.
\]

Specificity, on the other hand, can be defined as the percentage of well-nourished individuals correctly identified as not being malnourished, calculated as follows:

\[
\text{Specificity} = \frac{\text{True Negatives}}{\text{True Negatives} + \text{False Positives}} \times 100.
\]

If a body composition assessment technique is to be useful to identify individuals with malnutrition, it will need to have high sensitivity and specificity (>90%, ideally—particularly for sensitivity).

Statistical Techniques Used to Evaluate Agreement Between Methods in Validation Studies

Multiple statistical approaches can be applied to the problem of establishing the validity of a particular bedside method compared with a reference method. It is highly recommended that a combination of approaches be used to evaluate both mean-level and individual-level agreement. Regression and correlation analyses and paired t tests are the most commonly applied statistical approaches used to evaluate agreement between 2 body composition assessment methods in terms of mean-level agreement.

Regression analysis. Simple linear regression and correlation analyses are statistical approaches that have traditionally been used to evaluate the relationship between 2 different (eg, bedside and reference) body composition assessment methods as a way of validating the bedside technique. While correlation analysis simply tells us the strength of the interrelatedness of 2 sets of data, linear regression evaluates the nature of the relationship. 52 A strong linear relationship and high correlation between 2 sets of data may be interpreted to mean “good agreement” between methods, although this conclusion may be faulty. To use linear regression analysis on a set of data, several underlying assumptions should be met, including normal distribution of the continuous variable data, among others. Typically, measurements of a body composition compartment (eg, TBW) by a bedside method (eg, an SF-BIA equation) are regressed against those measured by a reference method (eg, deuterium dilution) to create a least squares regression line describing the relationship between the 2 methods, reflected by the general equation \( y = mx + c \), where \( m \) is the slope of the line and \( c \) is the intercept (ie, the value of \( y \) when \( x = 0 \)). The 45° line of identity (or line of equality) is used to represent perfect unity and can be represented by the equation \( y = x \), where the slope is 1 and the line goes through the origin.

For the purpose of illustration, Figure 1 represents various scenarios resulting from linear regression analyses applied to
body composition data. Plots A and C of Figure 1 show the regression line to be very close to the line of identity, when a bedside method is either highly precise and unbiased (A) or not very precise but unbiased (C). However, you would not usually expect the regression line ($Y$ on $X$) to go through the origin and have a slope of 1; in fact, more commonly, measurement errors in $X$ will reduce the slope of the line, such that the slope will be $<1$, and the intercept will be $>0$. The correlation coefficient (eg, the Pearson product-moment correlation coefficient), referred to as $r$ (along with the $P$ value indicating statistical significance), is used to reflect the strength of the association between the 2 methods compared with the line of identity.\(^{41}\) Interpretation varies depending on the application\(^{53}\); the closer the $r$ value is to 1.0 (+1 or −1), the stronger

Figure 1. Interpretation of linear regression for evaluation of methods agreement. (A) Regression line suggesting good agreement in hypothetical fat-free mass (FFM) measures by single-frequency bioelectrical impedance analysis (SF-BIA) and dual-energy X-ray absorptiometry (DXA), with low bias and high precision. (B) Regression line suggesting poor agreement between SF-BIA and DXA methods, with high fixed bias and high precision. (C) Regression line suggesting poor agreement between SF-BIA and DXA methods, with low bias and low precision. (D) Regression line suggesting poor agreement between SF-BIA and DXA methods, with high bias and low precision. Figure reproduced with permission from David Frankenfield.
the association (positive or negative, respectively), with values in the range of 0.9–1.0 typically interpreted as strong and values <0.50 interpreted as weak associations between body composition methods. Squaring $r$ to get the coefficient of determination ($R^2$) is useful, in that it can be interpreted directly as the percentage of variation in the measurements by one method that is related to the variation in the other.\textsuperscript{53,54} For example, an $r$ of 0.99 would yield an $R^2$ value of 0.98, which could be interpreted as “98% of the variation in TBW predicted by the SF-BIA equation is related to the variation in TBW measured by deuterium dilution"; such a high value for $r$ and $R^2$ could further be interpreted to suggest that the BIA approach might be considered a valid way to estimate TBW, although other considerations should be made before that conclusion is drawn. For example, the degree of error in the measurement needs to be evaluated. How far the measurements plotted on the $y$-axis vary around the regression line is typically described as the degree of error and may be represented by the term standard error of the estimate (SEE). SEE can be calculated by taking the square root of the sum of the squared differences between measured and predicted values divided by $n$, the number of pairs of scores, minus 2; this is represented by a very similar equation to that used to compute RMSE:\textsuperscript{55}

$$\text{SEE} = \sqrt{\left(\sum (Y - X)^2\right) / n - 2}.$$ 

It is also useful to evaluate SEE for a bedside method in terms of percentage error compared with the mean body composition compartment measurement by the reference method; for example, if we were comparing FFM measured by an SF-BIA approach against DXA, we would calculate

$$\text{SEE/(mean FFM by DXA) \times 100}.$$ 

Linear regression analysis has been criticized as a poor approach to establish the validity of a body composition assessment method, due to several inherent methodological limitations, and if used as the sole method to evaluate agreement between methods, it will more often than not lead to faulty conclusions. For example, if you have a high degree of heterogeneity and low precision in your sample, it is quite easy to obtain a high value of $r$.\textsuperscript{56} On the other hand, highly precise, homogeneous data generated by a bedside method can also yield a high value of $r$, despite having substantial systematic error (eg, high bias).\textsuperscript{57} The reader is again referred to Figure 1, which can further illustrate some of these points, based on hypothetical FFM data generated by a bioimpedance technique and a DXA instrument. If the bioimpedance method is highly precise and unbiased, the regression line for the relationship between that method and DXA might appear as shown in Figure 1A, with a statistically significant ($P < .05$) and very high value for $r$ and $R^2$. On the other hand, the bioimpedance method might be highly precise but significantly biased (see Figure 1B), and you could still have a strong correlation signified by a high value for $r$ and $R^2$ with a high degree of statistical significance. Such a situation can easily arise, as you can imagine when you look at 2 data sets that differ in value by 50% (eg, $\{1,2,3,4,5,6\}$ and $\{2,4,6,8,10,12\}$); comparison of these data by linear regression would yield a perfect correlation of 1 and an SEE of 0, despite the significant bias present. Figure 1C illustrates a highly imprecise method that produces values that on average are unbiased compared with the average of the values produced by DXA. This kind of result would meet our first definition of accuracy based on mean-level agreement but, due to the imprecision of the method, would not meet our second definition of accuracy based on individual-level agreement. Yet another possibility, illustrated in Figure 1D, can be seen when a bedside method is both imprecise and highly biased; in this case, the poor agreement between methods would be reflected by a low correlation coefficient.

The commonly applied least squares linear regression approach is partly based on the assumption that the reference method has minimal error compared with the range of the measurements. Given that there is error in both reference and bedside methods, this assumption may be faulty. Orthogonal least squares regression (eg, Deming regression\textsuperscript{58,59} or Passing-Bablok regression\textsuperscript{59,63}) assumes measurement errors are produced by both methods and are independent and normally distributed; this approach has been advocated as an alternative\textsuperscript{58,64,65} but has not yet been widely applied in the body composition assessment field. Finally, concordance correlation coefficient (CCC) and ICC (which is sometimes applied to evaluate precision, as mentioned earlier) analyses have been advocated as superior to simple linear regression analyses, because they yield a value approaching 1.0 only if there is minimal bias and the paired measurements are in close agreement. Thus, the CCC not only provides a measure of association between method variables but also indicates how close values are to the line of identity. Although CCC has been advocated to be superior to Pearson’s product-moment correlation because it considers both systematic and random error,\textsuperscript{66–68} it has also been criticized for relying on similar assumptions to those underlying Pearson’s correlation, such that highly heterogeneous and low precision data can lead to high $r$ values and erroneous interpretation as good agreement between methods.\textsuperscript{56} That said, there may be occasions when some form of regression analysis is the only procedure that can reasonably be applied to 2 sets of data. The particular example that comes to mind is when one is attempting to validate a method that measures a distinctly different lean tissue compartment from that measured by a reference method (eg, ultrasound measures of the upper quadriceps muscle compared with LST measures by DXA). It would be impossible to compare these directly, given the magnitude of difference between them, but a strong correlation would suggest that there might be utility in the ultrasound measures, particularly if the SEE were relatively
small. In this validation scenario, it would be recommended to take longitudinal measures to ascertain measures of change in the compartment of interest, assuming that relatively high precision and reliability (ie, minimal random error) in the bedside assessment method could be achieved. Strong correlation between measures of change in this case could be interpreted as good agreement between the techniques. Another situation that frequently arises in the nutrition assessment field is the evaluation of a variable or set of variables for the ability to identify malnutrition. For example, in the bioimpedance field, select raw data parameters have been evaluated for their ability to identify malnutrition. The only way that this kind of relationship can be established is through the use of regression analysis, most commonly taking a multivariate approach. Nevertheless, in most instances, it is not recommended to rely solely on the evaluation of the linear association between methods but rather to use a combination of techniques to evaluate both mean-level and individual-level agreement. Although analysis of variance (ANOVA) techniques would be most appropriate when comparing more than 2 methods, our focus here is on the comparison of two, a bedside method to a reference method, and thus, ANOVA will not be discussed. Paired t test is another procedure that is frequently used to determine mean-level agreement between a bedside and reference method.

**Paired t test.** The paired t test procedure can be used to test the difference between the means of 2 sets of measurements on each of a group of individuals. Put simply, the difference between each set of measurements is computed, and the test is applied to determine whether the mean difference is significantly different from zero. Because you are essentially testing the null hypothesis that the 2 sets of data produce equal means, achieving statistical significance for this test is interpreted as “the 2 methods do not produce equivalent values.” If, on the other hand, statistical significance is not reached, and thus the null hypothesis is not rejected, the conclusion cannot be assumed to be that there is no difference between the methods, only that no difference could be found. The statistical significance level (α) is typically set at 0.05, and the 95% confidence interval of the differences between the methods (defined as the mean difference ±1.96 SD) is calculated in the procedure. If the 95% confidence interval of the differences excludes zero as a possibility, then we conclude that bias is present in the method, and the null hypothesis is rejected (P < .05). On the other hand, if zero is included in the 95% confidence interval, the null hypothesis is not rejected (P > .05), and it might be concluded that the methods are not different, although this could be an erroneous conclusion.

A major limitation of using the paired t test to compare a bedside method with a reference method for agreement is that the results are quite difficult to interpret. For example, if statistical significance is reached when a paired t test is conducted, the finding of difference between methods could be due to random and/or systematic error. It would be impossible to ascertain which of these (and to what magnitude) is primarily in evidence. Furthermore, a nonsignificant result from a paired t test can easily occur when there is equally distributed imprecision in a set of measurements, with both negative and positive intermethod differences; these differences will cancel each other out when averages are compared, and the mean will appear to be close to zero. In fact, one conservative approach to method comparisons, particularly with small sample sizes, is to set statistical significance (α) at .10 rather than the typical .05. One might choose to do this to minimize the possibility of making a type II error (ie, to avoid the error of falsely accepting the null hypothesis of no difference between the 2 methods).49

One reasonable approach to minimize erroneous conclusions from a paired t test is to apply it in combination with linear regression and correlation analyses; a high degree of correlation and a highly nonsignificant P value for the paired t test would suggest good agreement (at least at the mean level) between methods. However, to ascertain fixed and proportional bias (ie, systematic error), as well as the level of agreement on the individual level, it is advisable to take additional steps, including analysis of the data as recommended by Bland and Altman.57,69,70

**Bland-Altman analysis.** Although it is not the only methodological approach that may be taken and is not universally advocated,71,72 Bland and Altman’s seminal paper in 198648 proposing their method of plotting the differences between values generated by 2 methods of measurement on the y-axis against the average of the values produced by the 2 methods on the x-axis is considered by many to be important to the body composition assessment field. They did not invent the method, but they advocated its application to the comparison of medical devices, laboratory tests, and other clinical techniques to ascertain bias in one method compared with another. One of the strengths of the Bland-Altman approach is that by calculating the mean of values measured by the 2 methods, there is recognition that there is random error associated with both methods (ie, the reference method is not infallible and thus is not a true “gold standard”). The Bland-Altman method can also be used to evaluate agreement between replicate measurements by one method, but this will not be discussed here. When using Bland-Altman plots to evaluate agreement between 2 methods, the mean values by the 2 methods are usually plotted on the x-axis. On the y-axis, the values for the difference between methods in the body composition parameter are typically plotted, but difference may also be expressed as a percentage of the mean body composition parameter values (either the mean value from both methods or the mean value by the reference method).

Figure 2 provides a series of Bland-Altman plots that might be constructed based on different hypothetical sets of FFM data generated by an SF-BIA equation and DXA. The first step in the process is to construct a scatterplot of the differences between methods (y) against the averages between methods (x). The
The next step is to take the mean difference (ie, bias) between methods and calculate the limits of agreement around the bias. This is done by calculating the mean ± 1.96 SD for the differences between the methods and drawing a horizontal line corresponding to the mean, to the value at the mean + 1.96 SD, and to the value at the mean – 1.96 SD. Assuming a normal distribution, the limits of agreement should encompass 95% of all measured values; the width of the limits of agreement has been described by the word “precision” based on the alternative definition of the word discussed earlier. Bland and Altman talked about how to calculate confidence intervals and SD for the limits of agreement as a way of evaluating the “precision” of those estimates, but that is beyond our discussion here.

A major criticism of the Bland-Altman technique is that the decision to accept the new technique, or, in our case, the bedside method as an acceptable alternative to the reference technique, is left entirely up to the subjective judgment of the evaluator. The ideal bedside method would demonstrate narrow limits of agreement, around a mean bias of zero. This is demonstrated in Figure 2A. Far more common, particularly when it comes to bioimpedance techniques compared with dilution reference methods, is to observe wide limits of agreement (see Figure 2B) when applying the methods in various clinical populations. Another important step is to determine if the differences are significantly correlated with the averages. If there is a significant correlation, then there is proportional bias. This can be seen in the example shown in Figure 2C. If, on the other hand, there is no significant correlation but the mean difference between methods is consistently and significantly less than or greater than zero, suggesting that the bedside method consistently produces underestimates or overestimates of the measured variable compared with the reference method, then fixed bias is evident (see Figure 2D).

The Bland-Altman method has recently been criticized, and suggestions have been made on how to optimize its application and interpretation. Ludbrook proposed that after plotting
the differences against the averages, the next step is to evaluate if the differences are correlated with the averages and, if the \( P \) value is not statistically significant for \( r \), then proceed with calculating the standard 95% limits of agreement. If, on the other hand, the averages and differences are significantly correlated, then proportional bias is evident and hyperbolic 95% prediction limits can be calculated. These recommendations hold if the scatter of differences is uniformly distributed across the range of averages. The reader is referred to the article by Ludbrook \(^{74} \) for additional information on what to do if the scatter of differences increases with the values of the averages and other contingencies.

One of the major problems in method comparisons is that both the bedside and reference techniques have a certain degree of random error associated with them. Even the most precise methods have some degree of error. The Bland-Altman approach is good, in that calculating the average between the 2 methods allows for relatively equal weighting between the 2 methods. However, as stated previously, it is up to the clinician to judge whether the limits of agreement are sufficiently narrow to allow for the bedside method to replace the reference method. To truly answer that question, it is important to consider the precision of both the reference and the bedside method. Bland and Altman \(^{57} \) described the importance of having at least 2 measurements by a particular method on each subject, so that the repeatability coefficient (ie, precision) for the method could be calculated from the SD of the differences between pairs of those repeated measurements, reflected by the equation \( 2 \times \text{SD} \). They described the repeatability coefficient as the “difference that will be exceeded by only 5% of pairs of measurements on the same subject.” \(^{57} \) This would allow for evaluating not only the limits of agreement between the 2 methods but also the repeatability (ie, precision) for each method separately. Repeated measures are not always obtained in studies; Ward \(^{75} \) recently provided a thought-provoking illustration of how a modified version of that approach might be used when repeated measures are not available to better define the validity of a method, building from a method described for techniques to measure cardiac output. \(^{76,77} \) These authors advocated for the calculation of the percentage error (PE) of the limits of agreement compared with the mean of the measurements as a way of defining a cutoff for acceptability of a method. In this approach, the “precision” of a method is defined to be 1.96 (or for simplicity, rounded to 2) times the CV for that method. Note that the use of the word “precision” here is not the traditional definition based on CV for repeatability and reproducibility but rather the alternate definition based on the spread of the individual data from the Bland-Altman analysis. Using the Bland-Altman plot shown in Figure 2B, we can see how we might reevaluate a hypothetical set of validation data. To do this, we will need to make an assumption that the precision of the DXA was actually determined from repeat measures to get a CV of 2%. We will also assume that we do not have repeat measures by our SF-BIA method, and thus, we can follow the recommended steps to calculate the “precision” of the SF-BIA equation as follows \(^{75} \):

\[
\text{Precision}_{\text{DXA}} = 2 \times \text{CV}_{\text{DXA}}, \text{ this we know and can calculate as } 2 \times 2 = 4.
\]

\[
\text{"Precision"}_{\text{SF-BIA}} = 2 \times \text{CV}_{\text{SF-BIA}}.; \text{ this we do not know and will try to estimate in a moment.}
\]

\[
\text{PE}_{\text{DXA} - \text{SF-BIA}} = 2 \times \text{CV}_{\text{DXA}} \times \text{100} ; \text{ this is the percent error of the difference between methods, which}
\]

\[
\text{we can calculate from Figure 2B using the limits of agreement of 5.8/mean of FFM by DXA, which}
\]

\[
\text{we will assume was determined to be 32 kg.}
\]

\[
\text{This means our PE}_{\text{DXA} - \text{SF-BIA}} = 5.8/32 \times 100 = 18\%.
\]

Next, we can think about the PE of DXA - SF-BIA in terms of each technique’s CV (with CV based on actual repeatability for the DXA method and CV based on methods comparison data from the Bland-Altman for the SF-BIA method) as outlined by Cecconi et al \(^{76} \):

\[
\text{CV}_{\text{DXA} - \text{SF-BIA}} = \sqrt{[\text{CV}_{\text{DXA}}^2 + \text{CV}_{\text{SF-BIA}}^2]},
\]

which can be rearranged to yield

\[
\text{PE}_{\text{DXA} - \text{SF-BIA}} = \sqrt{[\text{Precision}_{\text{DXA}}^2 + ("\text{Precision"}_{\text{SF-BIA}})^2]},
\]

which can be rearranged to yield

\[
\text{PE}_{\text{DXA} - \text{SF-BIA}}^2 = (\text{Precision}_{\text{DXA}}^2) + ("\text{Precision"}_{\text{SF-BIA}})^2
\]

\[
"\text{Precision"}_{\text{SF-BIA}} = \sqrt{[(\text{PE}_{\text{DXA} - \text{SF-BIA}})^2 - (\text{Precision}_{\text{DXA}})^2]},
\]

which can be solved as 

\[
"\text{Precision"}_{\text{SF-BIA}} = \sqrt{(18^2 - 4^2)}
\]

= 17%, which compared with the DXA precision of 2% is rather high and would be considered unacceptable.

Of course, had we not gone through these steps, we might have come to the same conclusion—namely, that the wide limits of agreement reflected too great a variability at the individual level; however, if our reference method had worse precision, say, a CV of 10%, then we may have deemed a 17% “precision” in our bedside method to be acceptable. If, on the other hand, we look at the data in Figure 2A, making the same assumptions about the precision of the DXA reference (still 4) and the mean of the FFM values by the 2 methods (still 32 kg), then using the limits of agreement of 1.0, we can calculate the PE to be 1.8/32 = 5.6%, and then “precision” of the SF-BIA in this case would be \( \sqrt{(5.6)^2 - (4)^2} = 3.9\% \); this would obviously be considered quite good. To reiterate Ward’s point, this is not to say that this approach should be adopted as standard but rather to suggest another way to look at limits of agreement in light of the reference method precision. \(^{75} \) Furthermore, the decision as to what degree of “precision” can be considered clinically acceptable is likely to vary depending on the application.
A 10% CV may be acceptable for a population-based study but may not be acceptable for individual bedside monitoring; there is not yet consensus regarding this point.

Hopefully, it is clear from these illustrations that it is important to keep in mind the precision of the reference method against which your bedside method is being compared. It is advantageous for the reference technique to be as accurate and precise as possible and for the precision of the reference technique to be measured within the same study, although logistical constraints may make that difficult to achieve, depending on the method. Although there is no consensus on this, it may be suggested that for clinical utility for bedside assessment and monitoring, a PE between reference and bedside methods of 10%, with an estimated “precision” for the bedside method that comes within 10% of the precision of the reference technique, could be considered acceptable for absolute whole-body estimates. If, on the other hand, we are attempting to measure longitudinal changes in whole-body estimates, to have confidence in our measures, we would want the percent error in our bedside (and reference) methods to be less than the expected percent change in the body composition variable.

Measuring Longitudinal Changes as a Way of Validating a Bedside Method

In many ways, for a method to be useful in the clinical setting, it is most critical to establish a method’s ability to detect changes over time. In fact, the evaluation of a bedside method’s ability to measure longitudinal changes in individuals who are undergoing a stimulus to induce anabolic or catabolic changes in lean tissue is an important way to validate it; furthermore, the measurement of changes can help to mitigate the problem of scaling differences (ie, fixed constant bias or systematic error) between methods. Two excellent examples of opportunities to do this include individuals undergoing cancer treatment who are likely to exhibit lean tissue wasting or individuals undergoing anabolic therapy to treat wasting in human immunodeficiency virus (HIV) infection. Obviously, the ability to measure changes in lean tissue is of fundamental importance to the clinician wishing to monitor a patient’s response to a nutrition intervention or simply to monitor nutrition status while the patient undergoes various medical therapies during a hospital stay. The clinician wants to have confidence that the bedside method is sufficiently sensitive to provide meaningful information about lean tissue changes. To achieve that, measures of change by the bedside method must be sufficiently close in value to those measured by the reference method to establish the method’s validity. If the reference method against which you are comparing your bedside method is sufficiently precise (ie, high reproducibility and repeatability), and the expected level of change in the body composition compartment is sufficiently high to be detected by that method, then one can evaluate the ability of a bedside method to measure that change. The minimal detectable change (MDC) is an important concept to keep in mind, as it is defined as the minimum change that must occur in a body compartment to achieve statistical significance at the group mean level; it is determined by the precision of the method, and it is expressed as a percentage (so it can be related to the size of the compartment being measured). The %MDC can be calculated as follows:

\[
\% \text{MDC} = 1.96 \times \frac{\text{precision}}{\sqrt{N}} \times 100
\]

When comparing a bedside method with a reference method for its ability to measure change, it is important to know the precision and MDC of the method. If a bedside method’s MDC is 6% for FFM, a measured change in FFM <6% could be erroneous (ie, it may or may not reflect a real change in body composition because it is below the threshold for detection). MDC will be discussed with regard to individual methods in subsequent sections. Another consideration for measuring body composition changes over time is the safety and efficacy of conducting repeat measurements by a method. Whereas multiple dilution, DXA, and CT would not be advisable to repeat over short-term follow-up due to logistical (dilution) and safety (DXA and CT) reasons, bioimpedance and ultrasound measures can be repeated frequently without concern. A review of the aforementioned reference methods will be presented next, followed by discussion of bedside techniques.

Reference Methods for Lean Tissue Assessment in Clinical Research: Multiple Dilution, DXA, and CT

Multiple Dilution

In the context of this discussion, multiple dilution techniques are commonly used to produce reference values for BCM and the fluid compartments to validate bioimpedance techniques. Multiple dilution typically involves the use of tritium (\(^{3}\)H), deuterium (\(^{2}\)H), or \(^{18}\)O as a tracer for TBW determination, in combination with a tracer such as bromide for ECW determination. By subtraction, ICW can be calculated, and from ICW, an estimate of BCM can be determined. The use of \(^{3}\)H is far less common, due to its radioactive nature, and \(^{18}\)O is typically only used when given in combination with \(^{2}\)H as doubly labeled water for total energy expenditure determination. In one approach, multiple dilution entails taking a predose blood or urine sample and then allowing a 3- to 5-hour equilibration period following dosing with deuterium oxide and sodium bromide solutions (with longer periods for individuals with expanded ECW, including those with extreme obesity, particularly for ECW determination by bromide dilution) before a final postdose blood or urine sample is collected. Deuterium enrichment of the biological sample may be determined...
by isotope ratio mass spectrometry,\textsuperscript{80–82} or Fourier transform infrared spectrometry,\textsuperscript{83} and bromide enrichment can be determined by high-performance liquid chromatography\textsuperscript{84,85} or non-destructive liquid X-ray fluorescence.\textsuperscript{86} The reader is referred to the International Atomic Energy Agency website resources\textsuperscript{87} and the excellent chapter by Schoeller\textsuperscript{88} on this topic.

Deuterium dilution for the determination of TBW has a reported precision of 1%–2%.\textsuperscript{78,80,81,88,89} The precision of ECW by bromide dilution is less well characterized but has been estimated to be attainable at 1%–3%.\textsuperscript{78,88} The determination of ICW by the multiple dilution technique has a somewhat lower precision, because it is derived by subtraction (TBW – ECW = ICW), and thus the errors in TBW and ECW lead to a propagation of error in the determination of ICW. Thus, it has been estimated that the ICW calculation by this approach has a relative precision of approximately 2.5\%\.\textsuperscript{88} It should be noted, however, that most research studies report precision estimates that are typically derived from repeated analytical measures of the final tracer concentration in biological samples, rather than repeated measures of TBW or ECW within a particular subject. This is because of the time required for clearance of tracer concentrations from the body, which makes it difficult to conduct repeated measures over the short term (ie, days apart), although separation by 2 or more weeks in a physiologically stable person and/or the use of 2 different tracers for the same compartment used simultaneously provide other ways of estimating precision.\textsuperscript{80,81,88,89} When precision can be determined by using 2 different tracers or by taking multiple measures to get a total CV, the contribution of biological variation can be calculated by subtracting the analytic error CV (from repeat analyses of the same sample or multiple samples drawn at the same time point for a subject) from the total precision CV as\textsuperscript{80}

\[
\text{Total CV}^2 - \text{Analytical CV}^2 = \text{Physiological Error CV}^2.
\]

It should be noted that Schoeller’s group\textsuperscript{80} applied this method to calculate the precision of total energy expenditure estimates using doubly labeled water; however, the concept can be applied to the determination of TBW by dilution methods as well. They reported a precision of 1.8\% in TBW determined by \textsuperscript{18}O and \textsuperscript{2}H within the same subject; thus, it is certainly possible to achieve excellent precision with dilution techniques (and to define it by using 2 different tracers for the same compartment).\textsuperscript{80} This raises another concern with dilution techniques, in that the use of different tracers (or the same tracer but applied using different protocols) can lead to differences in the dilution space that is used to estimate the fluid compartment; in addition, when a bedside method (eg, bioimpedance) developed using one tracer protocol is compared with a different one, large scaling differences may occur that could be interpreted as error in the bedside method when in fact it may be a systematic error in the reference. Schoeller et al\textsuperscript{81} have reported a systematic error of 3\% in the TBW measured by \textsuperscript{18}O compared with \textsuperscript{2}H. However, this is not likely to be a prominent concern in the validation literature; most studies of body composition will not use both tracers simultaneously (ie, doubly labeled water), unless they are also estimating total energy expenditure. These types of issues are not always easy to discern. If one is measuring fluid volume changes, at a minimum, it is important to keep in mind that the estimated %MDC\textsuperscript{78} based on the aforementioned precision estimates would be \textasciitilde3\%–5\% for TBW by deuterium dilution and \textasciitilde3\%–8.5\% for ECW by bromide dilution of the total volume. Similarly, the %MDC for ICW by multiple dilution would be at \textasciitilde7\%. True precision estimates based on repeated measures are not typically done in most research studies using dilution, so the %MDC for these compartments could be higher. To minimize error when a multiple dilution approach is applied in a clinical research study, it is advisable to adhere closely to standardized protocols\textsuperscript{88} and to obtain sample analyses of deuterium and bromide enrichment through collaboration and/or contract with established investigators and facilities. With that approach, multiple dilution can be a relatively affordable technique for the clinical researcher, compared with DXA or CT scanning; however, every effort should be made to minimize the risk of error as the study protocol is developed.

**DXA**

DXA scanners are becoming fairly commonplace in medical centers across the United States. DXA is commonly used as a reference method for the validation of bioimpedance and other bedside techniques to quantify FFM and LST. DXA is also used to generate a reference value for skeletal muscle mass; a value for appendicular skeletal muscle mass (ASM) is generated by summing the LST of the arms and legs. The ASM index can then be calculated by dividing the arm and leg LST by height squared. DXA-generated ASM and ASM index values have been used to define sarcopenia (see Table 1).\textsuperscript{100,100} Briefly, the DXA method requires an individual to lie down on a scanning table, and low-dose X-rays of 2 different energies pass through the body; a whole-body scan typically takes 10–20 minutes. An image is created as the photon detector measures the differential attenuation (or absorption) of the low and high X-ray energy by the soft tissue and bone. Soft tissue is further delineated into fat mass and LST, thus resulting in 3 body compartments: fat mass, LST, and bone mass (and density). Traditionally, weight and size limits constrained the use of DXA instruments in individuals with extreme obesity, but newer instruments allow for larger and heavier individuals to be scanned. DXA has a high degree of precision, with reports of 1%–2\% CV for LST.\textsuperscript{100,101,102} The MDC\textsuperscript{78} for DXA-measured LTM based on these precision estimates would be 3\%–6\%.

**CT**

CT is another imaging technique that has been used to estimate lean tissue in the clinical research setting. It involves exposure
to high-dose radiation and yields a highly accurate quantitative and qualitative image of skeletal muscle tissue, as well as total and regional adipose tissue, visceral organs, and bone from the detection of different X-ray attenuation.** Given the high-dose radiation exposure and the high cost, CT as a method for determining skeletal muscle is generally limited to those patients for whom a CT is ordered for diagnostic and other clinical reasons. Thus, one of the primary clinical populations in whom CT has been successfully used for lean tissue assessment and monitoring has been individuals with cancer.** The precision of CT has been reported to be 1% or less for the evaluation of skeletal muscle, which translates to an MDC of <3%. The reader is referred to the excellent tutorial by Prado and Heymsfield in the November 2014 issue of the Journal of Parenteral and Enteral Nutrition for a more complete discussion of the advantages and limitations of DXA and CT as reference techniques.

**Bedside Methods for Lean Tissue Assessment: Bioimpedance and Ultrasound**

**Bioimpedance**

Bioimpedance is the bedside approach that has been most widely investigated in clinical research and used by clinicians for assessment purposes in Europe and elsewhere around the globe, in large part due to the affordability, portability, and ease of use of bioimpedance devices. A bioimpedance measurement takes less than 15 minutes and is completely noninvasive, making it advantageous for repeat measurements. Like most body composition methods, bioimpedance devices do not directly measure body composition; they provide indirect estimates from the measurement of resistance of body tissues to an electric current. All bioimpedance techniques involve the application of a weak, alternating current at one or more frequencies, through leads attached to the body through current injection electrodes or through direct contact with electrodes in the case of stand-on scale devices. The differential flow of the current varies depending on the body composition; electrolyte-rich components, including blood and muscle, easily conduct the current, whereas fat and bone do not. The drop in voltage as the current passes through the body (ie, impedance) is detected whereas fat and bone do not. The drop in voltage as the current varies depending on the body composition; electrolyte-rich components, including blood and muscle, easily conduct the current, whereas fat and bone do not. The drop in voltage as the current passes through the body (ie, impedance) is detected through the voltage detection electrodes, and the impedance data (resistance, R; reactance, X; impedance, Z; and phase angle, PA) are recorded by the bioimpedance device. Some devices may only measure Z; it should be noted that R and Z are not interchangeable in BIA equations. Furthermore, different devices have different electronic circuitry, and the raw data generated cannot be considered interchangeable, although interdevice differences have not been well explored in the literature. Although no difference was detected between R and X measured by the RJL (RJL Systems, Clinton Township, MI) SF-BIA and Xitron 4000B and Hydra 4200 (Xitron Technologies, San Diego, CA) BIS devices, reported a difference of ~1%–1.5% in R values measured by the Valhalla (Valhalla Scientific, San Diego, CA) and RJL SF-BIA instruments that they corrected for in their FFM estimates from the National Health and Nutrition Examination Survey III (NHANES III) data. Although interdevice differences may not always be substantial, one should be wary of applying bioimpedance data from a particular device to an equation that was developed using a different device or a different approach (eg, whole body vs segmental).

The extrapolation from raw impedance data to volume or mass depends on several key assumptions, including that the body comprises 5 cylinders of uniform cross-sectional area, and height is an acceptable surrogate for the length of the conductor (ie, the distance between electrodes), among others. Body geometry assumptions are violated in individuals with obesity and in those with longer or shorter than average limbs. Basic concepts about bioimpedance are presented in an excellent online tutorial through the University of Vermont. There are many different ways that bioimpedance can be applied to estimate whole-body lean tissue (FFM or BCM) or fluid volume (TBW, ECW, or ICW) estimates using data generated by SF-BIA or multifrequency BIA (MF-BIA) or bioimpedance spectroscopy (BIS) devices. The reader is referred to several recent reviews that describe the different ways that bioimpedance may be used and the differences in technological applications and underlying assumptions between the 3 primary categories of bioimpedance techniques,** as well as currently available devices on the market, which range in cost from ~$500 to $20,000.

In general, the level of precision (ie, repeatability) produced by SF-BIA and MF-BIA devices is typically very good, with 1%–2% variability between repeat measures being reported in the literature. BIS devices may be somewhat more variable, due to the technical difficulties in producing stable measurements at the extremes of frequencies, but even most BIS devices have been reported to produce fairly precise measurements (2%–3%). Based on these precision estimates, the sensitivity of BIA/BIS techniques to measure whole-body composition changes can be calculated; for example, the MDC in FFM by SF- and MF-BIA has been estimated from 3%–6% and in TBW by BIS was observed to be 5%–8%. On the other hand, the level of accuracy in whole-body measures produced by various bioimpedance techniques, particularly in clinical populations with fluid overload, has been observed to be somewhat more variable.

Precision and accuracy of bioimpedance techniques are influenced by a number of factors, including patient (eg, degree of adiposity, fluid and electrolyte status, skin temperature) and environmental factors (ambient temperature, proximity to metal surfaces and electronic devices), the assumptions underlying prediction (SF-BIA or MF-BIA) or modeling (BIS) equations, instrumentation factors, and variations in measurement protocol. Close adherence to recommended measurement protocols for body weight (see Figure 3), height (see Figures 4 and 5),
and the bioimpedance\textsuperscript{96,108,109} measurement itself (see Table 2) is important to optimize the results of bioimpedance measurements and to minimize potential errors.

For example, how electrodes are placed is quite important. Depending on the device and whether the aim is to obtain whole-body or segmental measurements, adhesive electrodes may be placed in a standard wrist-ankle tetrapolar arrangement on the hand and the foot of one side of the body, on both sides of the body on both limbs using an 8-electrode arrangement, or on different segments of the body.\textsuperscript{97,110,111} Stand-on devices do not involve use of adhesive electrodes but rather direct contact with the electrodes at the feet and/or hands, depending on the device. Different devices and electrode placement techniques (eg, for whole-body and various segmental measurements) should not be considered interchangeable. For example, raw data obtained by a stand-on device should not be applied to an equation developed from data derived using a wrist-ankle tetrapolar electrode approach and vice versa, as to do so would introduce error into the estimate. The most important consideration is the consistent placement of the voltage detection (proximal) electrodes at standardized anatomical sites; for example, for wrist-ankle tetrapolar placement, the proximal electrodes are placed at specific sites on the dorsal surface of the wrist (eg, between the styloid processes of the ulna and
radius, ie, midline between the wrist bones) and anterior surface of the ankle (eg, between the medial malleolus formed by the lower end of the tibia and the lateral malleolus formed by the lower end of the fibula, ie, between the bony protuberances on the inner and outer sides of the ankle). Although it is common to have the electrode dissect the anatomical mark midline, different device manufacturers may suggest slight variations (eg, leading edge, midline, or trailing edge at the anatomical mark). Small variations in the placement of proximal electrodes can introduce error to impedance measurements (by ~2%). When applying data to an equation or making comparisons to reference data, one should follow the same procedure for electrode placement to optimize accuracy. Although it is often recommended that electrodes be placed at least 5 cm apart in adults, this may not always be possible; it has been suggested to move the distal electrode to create sufficient separation or to allow a 3-cm separation in children. The most important point is to keep the proximal electrodes at the anatomical marks and standardize the distance for follow-up measurements; fixed-distance (5-cm) electrodes have been shown to improve reproducibility in follow-up measurements, particularly in men. The placement of electrodes for segmental measurements to obtain consistency in the anatomical location and distance between them is also important to minimize potential error, particularly for longitudinal measurements, and to optimize accuracy when applying data to previously established equations and comparing data with reference data generated from those equations.

Protocol deviations (eg, failure to abduct the limbs) can contribute substantially to the error in measurements (eg, can affect
resistance measurements by 2%–3% for minor deviations\textsuperscript{109} and 18% and 43% for skin-to-skin contact with crossed legs and contact of hands to the waist, respectively\textsuperscript{108}. Body positioning is also quite important; after a person assumes a supine position, there is an immediate increase in impedance that occurs within the first 10 minutes and then a more progressive increase in impedance over the next several hours, reflecting the movement in fluid distribution toward the trunk.\textsuperscript{108} The greatest increase in impedance occurs at lower frequencies, reflecting the decrease in ECW. BIS-measured ECW decreased by approximately 3% and ICW increased by just less than 5% after moving from a semi-upright to a supine position for 30 minutes.\textsuperscript{112} Those who resume an upright position for 5 minutes return to presupine measurements.\textsuperscript{112} Because most bioimpedance reference data are measured within 5–10 minutes after patients assume the supine position, it is not recommended to apply such reference data to whole-body measurements taken from individuals who are restricted to bed rest, as errors in TBW and, in particular, ECW will be observed.\textsuperscript{108} Similarly, many investigators take measurements on the right side of the body, which may or may not be the dominant side; thus, much of the available reference data have been generated from right-sided measurements. It is estimated that impedance measurements can vary by approximately 2% between the dominant and nondominant side of the body;\textsuperscript{109} thus, it is recommended to follow the same protocol that was used to generate the equation and/or reference data that you intend to use and to maintain consistency in the side of body for longitudinal measurements.\textsuperscript{96} Increasing skin temperature by 6.5°C (11.7°F) also has a significant inverse effect on impedance, increasing 50-kHz estimates of TBW by approximately 13%; it is recommended that ambient temperature be maintained between 22.3 and 27.7°C (72.1–81.9°F) to minimize these effects.\textsuperscript{109,113} Other deviations (eg, failure to clean skin with alcohol or failure to remove metallic objects from the body prior to measurement) are associated with smaller errors (<1% effect on resistance).\textsuperscript{109} The reader is referred to several excellent sources that describe errors associated with protocol violations.\textsuperscript{106,108,109,113}

In general, the application of all varieties of bioimpedance techniques for the estimation of whole-body compartments has been accepted as reasonably accurate for the assessment of healthy nonobese individuals and for large-scale epidemiologic studies (ie, mean-level accuracy), but they have typically been questioned in terms of their applicability to the clinical setting due to variability at the individual level. That variability has been attributed to random and/or systematic errors that are sufficiently problematic to raise doubts about the capacity of bioimpedance methods to accurately quantify and monitor changes in whole-body volumes and masses, and the question has been raised if bioimpedance techniques can offer any advantages beyond simple anthropometrics to the assessment of nutrition status.\textsuperscript{114} On the other hand, bioimpedance data are typically highly correlated with lean tissue and fluid volumes measured by reference techniques, as well as with other indices of nutrition status, and so the primary problem seems to be the degree of variability in individual measures when comparing reference and bioimpedance methods. It has been proposed that the failure to account for precision error in the reference techniques (particularly multiple dilution) used to validate bioimpedance methods may have contributed (at least to some degree) to erroneous conclusions about the capacity of bioimpedance techniques to estimate body composition.\textsuperscript{75,115} Indeed, one recent study reported that bioimpedance techniques had errors of a similar magnitude to those observed in the reference (dilution) techniques and underscored the importance of reference method precision when interpreting validation data.\textsuperscript{115} Although bioimpedance techniques have been criticized for their limitations, particularly when applied to individuals with acute and chronic disease as well as individuals with obesity due to the potential violation of underlying assumptions,\textsuperscript{97} bioimpedance remains one of the few inexpensive, noninvasive, and portable bedside options currently available to clinicians.

It is not unreasonable to suggest that bioimpedance techniques might be able to contribute important objective information that may help the clinician to identify sarcopenia and malnutrition, whether or not a truly accurate measure of lean tissue or a fluid compartment volume can be obtained. Where bioimpedance techniques (particularly MF-BIA and BIS) may play an important role is in those individuals who have muscle loss that is not evident upon visual examination (including individuals with sarcopenic obesity) and those with excess ECW. Furthermore, bioimpedance techniques may be useful in monitoring the early changes in lean tissue that can happen with inflammation associated with chronic disease conditions (eg, cancer and HIV infection).\textsuperscript{116,117} An objective method for monitoring anabolic changes in response to nutrition and other interventions would also be highly advantageous to the clinician. Of note, bioimpedance is being used in dialysis centers around the world, based on a growing body of literature demonstrating that bioimpedance techniques (segmental MF-BIA and, to an even larger extent, wrist-ankle BIS) can be used to monitor fluid status and assess dry weight in individuals on dialysis.\textsuperscript{118–124} Finally, although much focus has been paid to evaluating the accuracy of bioimpedance techniques to quantify lean tissue and/or fluid volumes in absolute terms with variable success, it has been suggested by many that its true benefit may be in different applications (eg, by getting away from whole-body volume or mass estimates and using the data as markers to reflect lean tissue changes, nutrition status, and/or clinical outcomes).\textsuperscript{125,126} These and other novel applications of bioimpedance techniques are described in several recent reviews\textsuperscript{97,100,127} and are an active area of research.

It is important to understand how the different bioimpedance approaches (ie, SF-BIA, MF-BIA, and BIS) use the bioimpedance data to generate whole-body lean tissue mass (eg, FFM, LBM, or BCM) or fluid volumes (eg, TBW, ECW, or ICW), because they vary in terms of underlying assumptions.
and thus their theoretical applicability to different clinical conditions. For example, applying an SF-BIA equation to estimate ICW in an edematous patient is likely to produce substantial error because of the limitations of using just one frequency, whereas an MF-BIA or BIS device would offer at least the theoretical potential of differentiating ICW from ECW by measuring impedance at low frequencies. Furthermore, assumptions regarding normal body geometry and consistent, predictable hydration of FFM, as well as fluid distribution between the intra- and extracellular compartments, are frequently violated in the setting of acute and chronic illness. The generation of whole-body lean tissue estimates by SF- and MF-BIA requires the utilization of empirically derived prediction formulas that are often erroneous when applied to populations that differ fundamentally from those in which the formula was developed. The reader is referred to several recent reviews (from the abundant literature on the topic) that ideally matches the patient that he or she wishes to assess. The difficulties of making such a selection cannot be overemphasized. This is one of several limitations to the application of SF-BIA in the clinical setting, regardless of questions about potential accuracy for whole-body lean tissue estimates.

SF-BIA. SF-BIA devices measure impedance variables (ie, R, X, Z, and PA) at a single frequency, typically 50 kHz. To obtain an estimate of a body compartment, it is necessary to apply one or more of the impedance variables (most commonly Ht²/R, which relates the length of the conductor [ie, height of the body] to volume [eg, TBW]) to a prediction equation. Because components of the body (eg, TBW, ICW, FFM, BCM, and LST) are all intercorrelated, impedance can be used to predict any body composition compartment, if calibrated sufficiently to a reference method. Many prediction equations have been published in the literature, almost as many as there are studies that have been conducted to evaluate this approach. SF-BIA equations are typically developed by regressing impedance data against a reference method for some body composition variable (eg, FFM), perhaps using DXA or a 4-compartment model as the reference, in a homogeneous population sample. Ideally, the population-specific equation that gets developed is subsequently cross-validated in a separate sample (and if not cross-validated, the equation should not be used). It is important to realize that the measurement of impedance at only one frequency (ie, 50 kHz) is theoretically unable to differentiate between ECW and ICW (or BCM); in fact, in the clinical setting, it is not likely to be sufficiently high to quantify the entire TBW. This will become more clear once we have discussed BIS. Indeed, many SF-BIA devices produce elaborate output consisting of ECW, ICW, BCM, FFM, and more; however, they are simply predicting these compartments based on assumptions of static relationships between the body compartments from normative data; these assumptions may hold sufficiently true in healthy individuals. SEE for predictions of FFM by various SF-BIA equations have been observed to range from 1.8–4 kg in healthy normal-weight adults and from 1.6–3.4 kg in healthy elderly individuals and as high as 8.8 kg in overweight women, compared with DXA, 4-compartment models, and/or densitometry. Although they reported %body fat (%BF), rather than FFM, it is interesting to note that in women with extreme obesity (mean %BF 51.4% ± 3.6%), Das et al observed limits of agreement of ±5.1% and ±5.8% for 2 SF-BIA equations, as well as ±2.2% for 2H dilution for %BF measures, compared with %BF by the reference 3-compartment model. After gastric bypass-induced weight loss (and −16.8 ± 8.5 %BF), the methods agreement worsened considerably, with limits of agreement for the measures of change in %BF compared with a reference of ±8.5% (estimated PE calculated as per Ward of 51%) and ±9.6% (estimated PE 57%) for 2 SF-BIA approaches, compared with ±5.7% (estimated PE of 33.9%) for 2H dilution, underscoring the particular difficulties (by any method) when measuring changes in individuals with extreme obesity undergoing massive weight loss. In clinical populations, similar or higher errors have typically been observed for FFM and SM estimates. For a clinician who wants to evaluate lean tissue using an SF-BIA device, he or she is faced with the dilemma of choosing a prediction equation (from the abundant literature on the topic) that ideally matches the patient that he or she wishes to assess. The difficulties of making such a selection cannot be overemphasized. This is one of several limitations to the application of SF-BIA in the clinical setting, regardless of questions about potential accuracy for whole-body lean tissue estimates.

It has been suggested by a growing number of researchers that a potentially more useful approach to the use of SF-BIA might be to apply the raw data to a published equation that has established normative reference values or to generate the PA and compare that with published normative reference values Many BIA equations have been published, several different applications of 50-kHz data for estimating lean tissue have been primarily advocated, based on the concept of normalizing lean tissue data to height to create a standardized measure that might be used to indicate nutrition status. One is to apply the 50-kHz data to a validated multietnic prediction equation for skeletal muscle mass (SM) and then to calculate the skeletal muscle index (SMI); this has been incorporated into the definition of sarcopenia (relabeled as “whole-body fat-free mass index without bone” by Fearon et al) as part of the international consensus definition of cachexia, based on evidence that it correlated well with MRI and predicted functional status in older adults. This SF-BIA SM equation is as follows:

\[
\text{Skeletal muscle mass (kg)} = [(\text{Height}^2 / \text{R} \times 0.401) + (\text{Sex} \times 3.825) + (\text{Age} \times -0.071)] + 5.102, \\
\text{where height is in cm, R is resistance in ohms,}
\]

and for sex, male = 1 and female = 0.

This equation, first developed from bioimpedance data generated by an SF-BIA device (RJL Systems, Detroit, MI) in 269
white men and women (mean age, 42.5 years [range, 18–86 years]; mean BMI, 28.9 kg/m² [range, 16–48 kg/m²]), was subsequently cross-validated against MRI measurements in Hispanic (n = 26), African American (n = 53), and Asian (n = 40) men and women. The equation underpredicted skeletal muscle mass in Asian individuals (mean BMI, 22 kg/m²) but performed reasonably well in the other groups who had an average BMI similar to the white group, with an overall $R^2$ and SEE of 0.86 and 2.7 kg (9%), respectively. Individual-level agreement was not evaluated, and thus, it is not known how well this equation can predict SMI for bedside assessment. The first version of the SMI using this equation was calculated by dividing the skeletal muscle mass by body weight in kilograms and multiplying by 100. Subsequently, these investigators calculated SMI by dividing the skeletal muscle mass by height in meters squared and identified cutpoints to define sarcopenia based on physical disability; these cutpoints have been published as potential defining cutpoints (among others from the literature) in the European consensus paper on the definition of sarcopenia. Reference values for sarcopenia cutpoints based on the above BIA-derived SMI estimates are presented in Table 1.

Another proposed way to apply 50-kHz data is to calculate FFM using a validated SF-BIA equation developed from healthy reference population data. Kyle et al have taken the lead in investigating this concept for its application in the clinical setting; they recommend calculating FFM by their validated equation (coined the Geneva equation) generated from a Swiss population sample using a BIS (Xitron 4000B, then owned by Xitron Technologies) device and converting it to an index (fat-free mass index [FFMI]) by dividing it by the squared height; values can then be compared with reference data to evaluate the individual compared with healthy reference norms. The Geneva FFM calculation is as follows:

$$\text{FFM (kg)} = 4.104 + (0.518 \times \text{Height}^2 / R) + (0.231 \times \text{Weight}) + (0.130 \times \text{Reactance}) + (4.229 \times \text{Sex})$$

where height is in cm, $R$ is resistance in ohms, and weight is kg, and for sex, male = 1 and female = 0; convert to FFMI by dividing by height in meters squared.

FFM by this equation has been validated against DXA and has been shown to have acceptable error for healthy individuals (CV of 3.6%, SEE of 1.72 kg, and estimated PE of 6.3% in a large sample of 343 Swiss whites spanning ages 20–94 years and BMI from 17–33.8 kg/m²). Note that the authors did not calculate PE, but it was estimated from their data based on the calculation of 2 CV for the reported difference between methods, which would be 1.7 SD $\times$ 2/mean FFM of 54 = 6.3% error.

Chumlea et al also developed and cross-validated FFM SF-BIA prediction equations for men and women (ethnicity combined) based on the NHANES III data; however, FMMI values were not calculated in their study, and only the mean and SD for SF-BIA–predicted FFM in various age groups were presented rather than percentile data. Means and SD for FFM are inherently limited and should be interpreted with caution. It would be advantageous to have reference data by percentiles for FFM (and FFMI) from the NHANES III equations, if they are to be useful as a way to identify individuals with low muscle mass or at increased nutrition risk. Schutz et al have proposed that FFMI values below the fifth percentile cutpoint from healthy reference norms could be interpreted to suggest compromised nutrition status, although this is not well established, and fifth percentile cutpoints for FFMI are only available based on Swiss white population data, and thus, their application to other ethnicities and population samples may be limited. Nevertheless, the fifth percentile reference cutpoints defining low FFM calculated by the Geneva equation were based on a Swiss white population sample, range from <12.9 and <16.6 kg/m² for women and men, respectively, older than 75 years, to <14.4 and <17.2 kg/m² for women and men, respectively, 35–54 years in age.

Some studies reported that a low FFM in individuals (calculated by the Geneva equation but defined by slightly higher cutpoints than previously reported) was significantly associated with increased length of stay compared with those with normal or better values. Another recent study in a Brazilian population sample reported that low FFM (defined by the Geneva equation and Kyle cutpoints) was correlated with malnutrition diagnosed by subjective global assessment. Low FFM (defined by the Kyle cutpoints) but with FFM estimated by a BIS device with the device software (i.e., nor the Geneva equation) before surgery was shown to be independently associated with a higher prevalence of postoperative infections and increased length of stay in a Dutch ICU study sample. Another group recently reported that discrepancies between 2 different bioimpedance devices (an SF-BIA device and a BIS device) were evident for FFM values that resulted in significant differences in the number of individuals identified at nutrition risk; however, it should be clarified that the authors did not simply use the raw bioimpedance data generated by the 2 devices to calculate the Geneva equation FFMI (which would be a more appropriate way to make that comparison). Rather, they took the FFM values generated by the proprietary software in the SF-BIA device and the FFM value generated by the modeling and mixture-based equations programmed into the BIS device. This important detail was overlooked in a recent review by this author. Thus, it is not clear that important differences in actual impedance values were evident in that study. Furthermore, it should be emphasized that it is not recommended to use the cutpoints generated by one method (i.e., using the Geneva SF-BIA FFM equation) to interpret a measurement made by a completely different method (i.e., FFM generated by device software). The only way that it could be advocated to use the Kyle reference cutpoints for FFM would be to use the raw bioimpedance data to calculate FFM by the Geneva equation and then
convert to FFMI. That said, significant differences between devices might yield different values for the raw bioimpedance data that would affect the end FFMI or SMI values, as has been observed with reference data for PA, which will be discussed next.

PA is a simple bioimpedance parameter generated from the arctangent of the ratio of reactance to resistance at 50 kHz (see Figure 6). It can be calculated from the following equation:

$$\text{PA} = \text{arctangent} \left( \frac{X}{R} \right) \times 180^\circ / \pi.$$ 

Although it has been purported to relate to cell biology—namely, membrane integrity, permeability, size, hydration, and capacitance—it is not entirely clear that there is a physiologic foundation for the observed statistical associations. Nevertheless, a low PA has been reported to be an independent prognostic indicator of disease and/or nutrition risk/poor nutrition status in a variety of clinical populations, including HIV infection, cancer, chronic heart failure, cirrhosis, stage 5 chronic kidney disease, and hospitalized elderly. Normative values for PA measured at 50 kHz have been published based on healthy German, Swiss, and American populations; however, fifth percentile cutpoints are available only from German and American reference data. These fifth percentile cutpoints are presented in Table 3. The mean reference data (see Table 4) can be used to calculate the standardized PA (SPA), which involves creating a z score by using the following equation:

$$\text{SPA} = \frac{\text{Observed PA} - \text{Mean PA}}{\text{SD}},$$

where the mean and SD are from reference values.

The SPA was reported to be a significant independent predictor of malnutrition and decreased functional status, as well as 6-month survival. A standardized PA below −1.65 has been suggested as a cutoff to indicate malnutrition in individuals with cancer, representing the 5th percentile. Age, sex, and BMI are major factors that may affect the PA; Bosy-Westphal et al have published age-, sex-, and BMI-specific reference cutpoints. Others have presented age- and sex-specific reference data without delineating BMI categories. Observed differences between published reference cutpoints may be attributed to device-, BMI-, and/or population-specific differences. These issues are addressed quite well in the recent review by Norman et al; in addition to discussing the clinical use of 50-kHz PA and SPA, they also discuss the applications of 50-kHz bioelectrical impedance vector analysis (BIVA) originally developed by Piccoli et al for the assessment of hydration status. Whether or not PA or SPA can be used to identify individuals with muscle loss (compared with reference methods) or to uniquely contribute to the diagnosis of malnutrition beyond other diagnostic criteria has yet to be established.

**Figure 6.** Cole plot. The R, term is mostly representative of the intracellular water (ICW) resistance, but ICW resistance is also affected by cell membrane capacitance. Exponent α accounts for the distribution effects (including cell size and shape) that cause suppression of the semi-circle below the x-axis. Adapted from Kyle UG, Bosaeus I, De Lorenzo AD, et al. Bioelectrical impedance analysis—part I: review of principles and methods. *Clin Nutr*. 2004;23(5):1226–1243. Reprinted with permission from Elsevier.

**MF-BIA.** MF-BIA approaches may be taken using data collected by an MF-BIA device or by a BIS device, although the data are handled quite differently by the BIS technique, as will be discussed later. MF-BIA is quite similar to SF-BIA, but instead of just one frequency, MF-BIA devices measure impedance at 2 or more frequencies, typically at 4 or 5 frequencies, including at least 1 low (most commonly 5 kHz) and several higher ones (typically 50, 100, 200, and 500 kHz). As with SF-BIA, the MF-BIA approach to estimate whole-body volumes or lean tissue masses typically involves the use of linear regression–derived, population-specific equations. However, it has the advantage of taking measurements at a very low frequency and one or more high frequencies and thus theoretically allows for the estimation of both ECW and ICW. Separate equations are typically developed to predict ECW and TBW based on reference values derived from bromide dilution and deuterium (or tritium) dilution, respectively, in a particular sample population. Subsequently, ICW can be estimated through subtraction (ie, TBW − ECW = ICW). Alternatively, total body potassium reference values may be used to develop prediction equations for ICW and BCM. DXA, 4-compartment models, densitometry, or deuterium dilution data may be used to develop prediction equations for FFM or other lean compartments. Another less conventional approach to the use of MF-BIA data, if measurements are taken at 3 or more (and ideally 6 or 7) frequencies, is to model the data through nonlinear least squares curve-fitting techniques similar to those applied in BIS, although this approach has not been well studied and is likely to be less robust than the BIS approach of using spectral data measured at 50 or more frequencies, which will be discussed later.

Despite the added advantage of measuring at multiple frequencies, the more typical MF-BIA prediction equation approach
is limited for whole-body estimates by the same complaints that may be made about the SF-BIA prediction equation approach, particularly in the clinical setting and in individuals with obesity, where underlying assumptions may be violated. Far fewer MF-BIA equations than SF-BIA equations have been published in the literature, but some have been specifically developed for lean tissue and/or fluid volumes (TBW, ECW, and/or ICW) in healthy individuals, and some have been developed for clinical assessment of fluid volumes (eg, in surgical patients and individuals with HIV infection). Although MF-BIA equations may do reasonably well in healthy individuals, reports of errors are quite variable in clinical populations, with reported SEE of 2.0 for ECW and 4.0 for TBW measures in individuals with cirrhosis, and limits of agreement of ±18% for ICW estimates in individuals with HIV infection. Although clinicians may search the literature and identify appropriate equations with which to apply MF-BIA impedance data to estimate whole-body volumes or masses with no indication to the clinician which equations and/or specific variables were used to generate them. Any bioimpedance device that does not provide the raw data and/or uses “black box” proprietary equations is not as useful as it could be. These issues are major barriers facing the further development of bioimpedance for clinical use.

Given the concerns about the accuracy of whole-body measures in the clinic setting, there has been significant interest in the utilization of a simple ratio of impedance measured at 200 kHz to impedance measured at 5 kHz. Although it is possible to create ratios from impedance measured at other frequencies, the 200/5-kHz impedance ratio has been the most commonly investigated as a potential marker for nutrition status and/or disease severity, with worse outcomes the closer to 1.0 the ratio is. For example, higher 200/5-kHz impedance ratios have been observed in individuals with poor nutrition status, postoperative edema, and impaired renal and cardiac function. Values >1.0 suggest device error.

In a study that included 151 healthy volunteers, normal values for this impedance ratio were ≤0.78 in men and ≤0.82 in women. Values above these cutpoints were associated with more than a 4-fold higher risk of malnutrition, identified by low total body protein measured by neutron activation analysis. Importantly, a high impedance ratio was reported to have a sensitivity of 79% and a specificity of 53% for identifying malnutrition. Similar to the concept of using PA as a predictive variable, the simplicity of using a ratio to evaluate nutrition status has appeal, given that it requires no population-specific regression equations and
gets away from the potential errors associated with whole-body estimates; however, for it to be clinically useful, reference cutpoints need to be established. Furthermore, just as is true with the PA, the use of the impedance ratio as a marker relies on a statistical relationship, such that other variables that are equally good at predicting nutrition status could be easier to measure. The question remains to be answered regarding whether the measurement of the 200/5-kHz or any other high- to low-frequency impedance ratio can improve upon the diagnosis of malnutrition beyond physical examination and other clinical data.

**BIS.** In contrast to SF-BIA and MF-BIA approaches, which require the use of a regression-derived equation to predict whole-body volumes or masses from raw data assuming a linear relationship, BIS devices use a more scientific approach based on biophysical modeling. Impedance spectroscopy is a well-known analytical technique used in materials and biological research; technological advances in the early 1990s facilitated the application of BIS for the assessment of in vivo body composition for the first time. In contrast to SF-BIA and MF-BIA devices, BIS devices apply the electrical current (typically ≤800 µA) over a range of frequencies, from very low (eg, 1 or 5 kHz) to very high (eg, 1000–1200 kHz), measuring impedance data (ie, R, X, Z, and PA) at 50 or more frequencies. The application of a variable-frequency current to biological tissue yields a response that gives rise to a semicircular graphical relationship between R and X due to the cell membrane capacitance (Cm); this relationship is best represented by the standard and classic Cole model of biological tissue. Nonlinear least squares curve-fitting techniques are applied to the spectral data, generating a semicircular Cole plot and Cole model terms RE, RI, Cm, and exponent α (see Figure 6). These terms can then be used to provide estimates of ECW and ICW through algorithms that will be discussed later.

**Table 4.** Mean Values for Phase Angle* From German, Swiss, and American Reference Data.

<table>
<thead>
<tr>
<th>BMI</th>
<th>20–29</th>
<th>30–39</th>
<th>40–49</th>
<th>50–59</th>
<th>60–69</th>
<th>≥70</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bosy-Westphal et al<em>138</em>: Males (n = 30,572)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.5–25</td>
<td>6.9 ± 0.72</td>
<td>6.7 ± 0.70</td>
<td>6.5 ± 0.70</td>
<td>6.2 ± 0.66</td>
<td>5.8 ± 0.82</td>
<td>5.1 ± 0.86</td>
</tr>
<tr>
<td>25.1–30</td>
<td>7.0 ± 0.72</td>
<td>6.9 ± 0.69</td>
<td>6.7 ± 0.70</td>
<td>6.4 ± 0.72</td>
<td>6.0 ± 0.75</td>
<td>5.4 ± 0.77</td>
</tr>
<tr>
<td>30.1–35</td>
<td>7.0 ± 0.71</td>
<td>6.9 ± 0.72</td>
<td>6.8 ± 0.68</td>
<td>6.4 ± 0.70</td>
<td>6.0 ± 0.76</td>
<td>5.5 ± 0.76</td>
</tr>
<tr>
<td>35.1–40</td>
<td>6.9 ± 0.74</td>
<td>6.9 ± 0.69</td>
<td>6.6 ± 0.74</td>
<td>6.4 ± 0.76</td>
<td>6.0 ± 0.85</td>
<td>5.4 ± 0.73</td>
</tr>
<tr>
<td>40.1–50</td>
<td>6.7 ± 0.69</td>
<td>6.7 ± 0.76</td>
<td>6.4 ± 0.77</td>
<td>6.2 ± 0.77</td>
<td>5.8 ± 0.86</td>
<td>5.0 ± 0.87</td>
</tr>
<tr>
<td>Barbosa-Silva et al<em>139</em>: Males (n = 832)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean BMI 25.6 ± 4.2</td>
<td>8.0 ± 0.75</td>
<td>8.0 ± 0.85</td>
<td>7.8 ± 0.85</td>
<td>7.3 ± 0.89</td>
<td>7.0 ± 1.1</td>
<td>6.2 ± 0.97</td>
</tr>
<tr>
<td>Bosy-Westphal et al<em>138</em>: Females (n = 183,176)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.5–25</td>
<td>6.0 ± 0.68</td>
<td>6.0 ± 0.67</td>
<td>6.0 ± 0.68</td>
<td>5.7 ± 0.68</td>
<td>5.5 ± 0.78</td>
<td>5.1 ± 0.84</td>
</tr>
<tr>
<td>25.1–30</td>
<td>6.1 ± 0.68</td>
<td>6.2 ± 0.67</td>
<td>6.1 ± 0.67</td>
<td>5.9 ± 0.70</td>
<td>5.6 ± 0.72</td>
<td>5.3 ± 0.78</td>
</tr>
<tr>
<td>30.1–35</td>
<td>6.2 ± 0.68</td>
<td>6.3 ± 0.67</td>
<td>6.2 ± 0.69</td>
<td>5.9 ± 0.70</td>
<td>5.6 ± 0.73</td>
<td>5.3 ± 0.75</td>
</tr>
<tr>
<td>35.1–40</td>
<td>6.2 ± 0.68</td>
<td>6.2 ± 0.66</td>
<td>6.2 ± 0.70</td>
<td>5.9 ± 0.72</td>
<td>5.6 ± 0.75</td>
<td>5.3 ± 0.84</td>
</tr>
<tr>
<td>40.1–50</td>
<td>6.2 ± 0.66</td>
<td>6.2 ± 0.71</td>
<td>6.1 ± 0.72</td>
<td>5.8 ± 0.70</td>
<td>5.5 ± 0.77</td>
<td>5.1 ± 0.72</td>
</tr>
<tr>
<td>Barbosa-Silva et al<em>139</em>: Females (n = 1135)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean BMI 26.0 ± 6.4</td>
<td>7.0 ± 0.92</td>
<td>6.9 ± 0.84</td>
<td>6.9 ± 0.85</td>
<td>6.6 ± 0.87</td>
<td>6.0 ± 0.83</td>
<td>5.6 ± 1.02</td>
</tr>
<tr>
<td>Kyle et al<em>157,b</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (n = 2735)</td>
<td>7.3 ± 0.8</td>
<td>7.5 ± 0.8</td>
<td>7.2 ± 0.9</td>
<td>7.2 ± 0.9</td>
<td>6.6 ± 0.9</td>
<td>6.1 ± 0.9</td>
</tr>
<tr>
<td>Mean BMI</td>
<td>22.2</td>
<td>23.5</td>
<td>24.1</td>
<td>24.4</td>
<td>25.0</td>
<td>25.7</td>
</tr>
<tr>
<td>Females (n = 2490)</td>
<td>6.6 ± 0.8</td>
<td>6.6 ± 0.9</td>
<td>6.7 ± 0.8</td>
<td>6.5 ± 0.9</td>
<td>6.0 ± 0.8</td>
<td>5.4 ± 0.9</td>
</tr>
<tr>
<td>Mean BMI</td>
<td>21.0</td>
<td>21.4</td>
<td>21.9</td>
<td>22.7</td>
<td>24.3</td>
<td>25.9</td>
</tr>
</tbody>
</table>

The Bosy-Westphal et al*138* study used the BIA 2000 Data Input (Frankfurt, Germany) single-frequency bioelectrical impedance analysis (SF-BIA) device; the Barbosa-Silva et al*139* study used the Model 101 RJL Systems (Mt Clemens, MI) SF-BIA device; the Kyle et al*157* study used a combination of Model 109 and 101 RJL Systems SF-BIA devices and the Xitron 4000B (Xitron Technologies, San Diego, CA) bioimpedance spectroscopy device. All values presented as means rounded to the nearest tenth to simplify the presentation. BMI, body mass index.

*BMI listed for each age group represents the mean group BMI for that category.
semicircle and extrapolated to the exact point where there is only ECW and no ICW conduction (ie, $R_\infty$ or $R_m$) and where there is complete conduction through both ECW and ICW (ie, $R_c$). From $R_\infty$ and $R_m$, $R_c$ and $C_m$ can be derived; the exponent $\alpha$ term that is also a Cole model parameter has been discussed elsewhere.\textsuperscript{176} To best solve for the primary components of the Cole model, sufficient data points are required; thus, impedance data are typically measured at 50 frequencies.

The characteristic frequency ($f_c$) is also computed from $R_\infty$, $R_c$, and $C_m$; it is defined as the point of maximal reactance or frequency where the measurement is halfway between zero and infinity.\textsuperscript{102} As you can see from Figure 6, it is graphically represented as the top of the semicircular Cole plot. Theoretically speaking, at any frequency other than zero and infinity, the proportion of ICW measured varies with increasing frequency and when the tissue changes, causing a change in $f_c$.\textsuperscript{102} The BIS modeling approach is the only bioimpedance approach that actually calculates the $f_c$. Although $f_c$ may in time have its own important clinical applications, for this discussion, it is useful to understand $f_c$ because it underscores an important advantage of the BIS modeling approach.\textsuperscript{102} For example, at 10 kHz, there is more ICW measured than at 5 kHz, and at 50 kHz, there is less ICW measured if $f_c$ is 100 kHz than when $f_c$ is 30 kHz. Although 50 kHz is used to predict TBW, it is typically >20:1 away from where ICW is fully measured (ie, at infinitely high frequency). One of the assumptions made by SF-BIA is that 50 kHz is sufficiently high to predict TBW, but that is likely to only hold true in healthy people (for whom the $f_c$ is typically around 40 kHz and underlying relationships between compartments are not altered).\textsuperscript{102} That will almost certainly not hold true in many individuals with acute and chronic disease, who may have significantly greater $f_c$ (eg, >200 kHz in individuals on dialysis\textsuperscript{177}). For these individuals, much higher frequencies would likely be needed. For more in-depth insight into this topic, the reader is referred to the excellent review by Matthie.\textsuperscript{102}

Interestingly, it has been suggested that the reason 50 kHz PA has been shown to predict nutrition status is because, for most individuals, 50 kHz exceeds their $f_c$;\textsuperscript{102} it is quite possible that the application of PA as a marker of malnutrition might be improved by using the PA measured at the $f_c$ measured by a BIS device, but this remains to be investigated. Nevertheless, it is important to realize that no single frequency (5, 50, $f_c$, 100, 200, 500, or even infinite kHz) is a measure of TBW but rather a mixture of ICW and ECW, which have substantially different (ie, 5:1) resistivities.\textsuperscript{176} The only way to accurately estimate ECW and ICW is to independently evaluate them, and BIS is theoretically ideal to do this.

Typically, BIS devices are programmed with software that first models the data, using nonlinear least squares curve fitting and assuming that the fit is sufficiently good, then applies parameters ($R_\infty$, $R_c$, $C_m$) generated by the modeling procedure to calculate fluid volumes and other body composition compartments using a particular algorithm. The Cole model parameters can be applied to equations that were originally developed based on Hanai’s mixture theory,\textsuperscript{178} which, as it applies to humans, assumes that the electrical properties of tissues are affected by the mixture of conducting (eg, body water, electrolytes, muscle and other lean tissue) and nonconducting (eg, bone, fat) body components. The BIS approach involves several other key assumptions, including constants for estimating the specific resistivity of the ICW and ECW compartments (termed apparent resistivities or resistivity coefficients) and constants for body density and shape; these are likely sources of error in individuals with obesity and edema, who may vary from normal-weight shape and density. ICW and ECW resistivity coefficients used in device software algorithms may differ, depending on the manufacturer’s decision to use published values or to generate its own from reference (eg, multiple dilution) data. Most commonly applied are the coefficients published by De Lorenzo et al,\textsuperscript{176} which were derived from TBW and ECW determined by deuterium and bromide dilution, respectively, and by the Xitron 4000B (Xitron Technologies) BIS device in healthy normal-weight individuals, but coefficients have been generated in other populations, including individuals with extreme obesity.\textsuperscript{47} The use of different dilution techniques and different BIS devices will yield different constants; this is a potential source of error in BIS measurements, although it has not been well studied. The original mixture theory-based equations for determining fluid volumes from BIS spectral data were first developed by Xitron Technologies in the early 1990s. Improvements were made after recognizing how the resistive properties of the ICW compartment were affected by adiposity, and those second-generation equations were published.\textsuperscript{176,179}

Much of the published literature on BIS has evaluated the original Xitron-Hanai–based mixture equation approach using resistivity coefficients generated by the device manufacturer or those published by De Lorenzo et al\textsuperscript{176} to estimate fluid volumes and/or FFM or BCM in various populations. In general, a similar pattern of errors in whole-body estimates has been observed with BIS as with other bioimpedance approaches—namely, good mean-level agreement but sometimes large variability at the individual level. For example, focusing on ICW given its relevance to lean tissue, relatively acceptable mean-level errors (SEE of ~1.5–2.5 L in ICW, approximately 8%–14% error) have been observed in healthy individuals.\textsuperscript{180} BIS using the Xitron-Hanai–based mixture equation approach was also shown in individuals with HIV infection undergoing treatment with oxandrolone to produce similar SEE estimates for ICW in absolute terms before (SEE 9%) and after (SEE 6%) treatment and to track changes in ICW (BCM) better than an HIV-specific SF-BIA equation,\textsuperscript{116} but variability in individual measurements was in some cases quite large.\textsuperscript{46} Wide limits of agreement (±5 L, compared with the mean change of ~4.4 L) were observed in measures of ECW changes in ICU patients with major trauma and sepsis.\textsuperscript{181} Substantially greater error (with fixed and proportional bias, as well as wide limits of agreement) was observed...
for all fluid volume measures in individuals with extreme obesity.\textsuperscript{47–49} Since then, it was determined that a significant portion of the error observed in BIS measurements came from the impact that adiposity had on the ICW estimates, and new equations that incorporated a correction for BMI were advocated and tested.\textsuperscript{182} These revised equations, with the correction for BMI, are as follows\textsuperscript{182}:

\[
\begin{align*}
\text{ECW} &= k_{\text{ECW}} \times \left[ \left( \text{Height}^2 \times \sqrt{\text{Weight}} \right) / R_i \right]^{2/3}, \\
\text{ICW} &= k_{\text{ICW}} \times \left[ \left( \text{Height}^2 \times \sqrt{\text{Weight}} \right) / R_i \right]^{2/3},
\end{align*}
\]

where height is in cm, weight is in kg, \( R_i \) is the intracellular resistance (value generated from the Cole modeling procedure), and \( k_{\text{ECW}} \) and \( k_{\text{ICW}} \) are functions of BMI, calculated as follows:

\[
k_{\text{ECW}} = (a / \text{BMI}) + b, \text{ and } k_{\text{ICW}} = (c / \text{BMI}) + d,
\]

where \( a = 0.188, b = 0.2883, c = 5.8758, \) and \( d = 0.4194, \) as determined from reference data and later cross-validated.

The BMI correction reduced, but did not eliminate, the individual errors associated with BIS ICW measurements in individuals on dialysis and in healthy individuals, particularly in those at the extremes of BMI.\textsuperscript{182}

This same group has incorporated a new conceptual model (which will be referred to henceforth as the Chamney model) regarding the hydration of adipose and lean tissue based on cadaver data\textsuperscript{183} into the BIS approach and is applying the technology to the problem of fluid management in individuals on dialysis.\textsuperscript{184} In the most recently published version of the model, it is conceptually possible to quantify excess fluid and thus more accurately assess the lean tissue compartment, which they term "normally hydrated lean tissue." The ECW and ICW volumes calculated from the BMI-corrected mixture based equations could be applied to the Chamney model equations\textsuperscript{183}:

\[
\begin{align*}
M_{\text{EF}} &= (1.136 \times \text{ECW}) - (0.430 \times \text{ICW}) - (0.114 \times \text{Weight}), \\
M_{\text{NH}, \text{LT}} &= (2.725 \times \text{ICW}) + (0.191 \times M_{\text{EF}}) - (0.191 \times \text{Weight}), \\
M_{\text{NH}, \text{AT}} &= Wt - M_{\text{EF}} - M_{\text{NH}, \text{LT}},
\end{align*}
\]

\[
M_{\text{Fat}} = 0.753 \times M_{\text{NH}, \text{AT}}, \quad \text{where } M_{\text{EF}} \text{ is the excess fluid mass (kg), } M_{\text{NH}, \text{LT}} \text{ is the normally hydrated lean tissue mass (kg), } M_{\text{NH}, \text{AT}} \text{ is the normally hydrated adipose tissue mass, } M_{\text{Fat}} \text{ is the fat mass (kg), ECW is the extracellular water mass (kg), ICW is the intracellular water mass (kg), and weight is in kg.}
\]

The validation of this approach to quantify lean tissue is limited by the fact that there is no clear reference method for specifically measuring the "normally hydrated lean tissue" compartment. However, it does appear that their approach, or some derivation, is being successfully applied to evaluate dry weight and the adequacy of dialysis in individuals with renal failure.\textsuperscript{183,184} It may be surmised that the better their model-derived estimates of excess fluid become, the more refined will be the estimates of lean tissue. The refinement of the apparent resistivity constants, among other factors, will almost certainly be a necessary part of that process. A recent report provides compelling evidence of the efficacy of this approach for lean tissue assessment.\textsuperscript{185} Marcelli et al\textsuperscript{185} reported that individuals on hemodialysis who had NH_LT values below the 10th percentile from healthy reference data had significantly higher mortality than those with higher values, suggesting that it could be indicative of malnutrition and poor clinical status. This is an exciting and promising area of bioimpedance application, and additional research is warranted to see if BIS-measured "normally hydrated lean tissue" can be used to evaluate muscle loss for diagnosing malnutrition and to track changes in response to nutrition interventions.

**Ultrasound Measurements of the Quadriceps Muscle**

Ultrasonography is a readily available, portable technique used for diagnostic and monitoring purposes in clinical settings. Good-quality ultrasound devices may cost $30,000–$50,000, making them slightly more expensive than the top-end BIS devices but far less expensive than a DXA or CT scanner. The application of ultrasound to measuring body composition is not entirely new; it has historically been used for assessing the abdominal adipose compartment to distinguish visceral and subcutaneous fat depots.\textsuperscript{186} Much of the research that has been done on ultrasound applications in muscle tissue has come from the sport injury and neuromuscular disease fields.\textsuperscript{187,188} More recently, there has been growing interest in the use of ultrasound to quantify lean tissue at the bedside, particularly in the intensive care setting.\textsuperscript{189–192} For example, one recent study tracked loss of muscle mass in critically ill patients using ultrasound measurements of the quadriceps muscle thickness.\textsuperscript{192} Muscle thickness can be measured at one or more anatomical sites, and these values may be extrapolated to whole-body lean tissue estimates,\textsuperscript{193} although there are only a few equations that have been proposed to predict FFM\textsuperscript{194,195} or skeletal muscle mass\textsuperscript{196,197} using the sum of thicknesses from 3–9 anatomical sites in healthy individuals. For a more complete discussion and review of the available validation literature on ultrasound for lean tissue assessment, the reader is referred elsewhere.\textsuperscript{198}

Generally speaking, ultrasound involves the transduction of high-frequency sound waves through the skin; the ultrasound beam is partially reflected back (as an echo) to the
transducer from the interface of different underlying tissues (e.g., subcutaneous adipose and skeletal muscle). How much the tissue reflects is expressed as acoustic impedance, and this ranges from the lowest acoustic impedance levels associated with air to the highest acoustic impedance values associated with bone tissue. The best image quality is observed with the highest acoustic impedance values. As the transducer receives the reflected beam, the echo is converted into electric signals and a 2-dimensional image is formed. The scanning procedure itself is deceptively simple; the difficulties arise when it comes to interpreting the image to determine muscle thickness at the site(s) of measurement. The interface between adipose tissue and muscle can be difficult to differentiate, because the acoustic impedance of these tissues is somewhat similar, and they produce a weaker echo compared with one that would include interface with bone tissue. One must identify the tissue boundaries and then measure the thickness of muscle using the electronic calipers that are standard with most systems; boundary identification is somewhat subjective and can lead to interrater errors.

In addition, ultrasound measurements of muscle may be made with the muscle in the contracted or relaxed state, in the standing or supine position. For most clinical settings, the individual will likely be in the supine position and the muscle will be relaxed; in the relaxed state, muscle tissue is compressible. There is not yet consensus on the optimal protocol to follow in terms of the degree of pressure to apply when taking a measurement in the supine, relaxed position. Most published studies have used no or minimal tissue compression, but one recent report used maximal compression. Although maximal pressure has been advocated as a way to counter the problem of edema interfering with measurements in clinical populations, it has also been suggested that minimal or no compression may better allow for the identification and perhaps even quantification of edema, as has been observed in measures of the anterior thigh muscle thickness. However, the application of no/minimal transducer compression on the muscle tissue may actually reduce the ability to detect small lean tissue changes in the presence of edema and/or adipose stranding. Furthermore, the no/minimal pressure approach in individuals with obesity will quite likely be ineffective in visualizing full tissue thickness and reference bone because of shear limb thickness exceeding ultrasound depth capabilities. On the other hand, with maximal compression, the degree of muscle compressibility and presence of edema may contribute to error, and the degree of pressure placed on the transducer against the skin is difficult to control to standardize measurements, making reproducibility difficult to achieve for longitudinal measurements. Maximal vs no/minimal transducer force will yield quite different muscle thickness measures, although this has not yet been well studied. In the adipose ultrasound literature, it was observed that the use of maximal transducer force reduced the measurements of subcutaneous adipose thickness by 24%–37% compared with minimal transducer force. These are vitally important issues that need to be resolved so that consensus on protocol can be reached and appropriate reference data can be generated for use in clinical populations.

A limited number of studies that have investigated the validity of ultrasound for quantifying muscle mass are relevant to the clinical setting. Ultrasound measurements of the biceps, forearm, and mid thigh muscles (degree of transducer force was not described) were shown to correlate reasonably well ($r = 0.872$; $P < .001$) with LST measured by DXA in individuals with multiple-organ failure. The intraobserver variability for mid thigh measurements was observed to be 2.3% (CV), and the interobserver CV for the mid thigh was observed to be 4.3%. In another study, the thickness of the quadriceps femoris muscle was negatively correlated with length of stay in individuals in the ICU; interrater CV was observed to be 1.3% using no/minimal compression. Ultrasound of the quadriceps muscle using no/minimal compression has been shown to correlate well with quadriceps muscle strength in individuals with chronic obstructive pulmonary disease and has also been shown to correlate reasonably well ($r = 0.78$; no $P$ value provided) with MRI measurements in healthy individuals undergoing a bedrest study.

Two recent studies have evaluated the reliability and reproducibility of ultrasound measurements of the quadriceps muscle thickness in the setting of critical illness. Following a maximal compression protocol (see Figures 7 and 8), Tillquist et al reported good intrarater (average ICC = 0.98 with between-subject variance of 0.26 and within-subject variance of 0.05) and interrater (average ICC = 0.95 with between-subject variance of 0.26 and within-subject variance of 0.015) reliability, with average muscle thicknesses for the left and right leg of 2.01 ± 0.52 cm for trainer 1, 2.0 ± 0.50 cm for trainer 2, and 2.09 ± 0.52 cm for the trainee. More recently, following a minimal compression protocol, another group reported excellent intrarater CVs of 1.91% and 1.32% with absolute variability expressed as the limits of agreement of 0.33 and 0.12 cm for operators 1 and 2, respectively. The absolute median interrater variability was 0.05 cm, with an ICC of 0.99 for single measurements; the mean muscle thickness over 86 duplicate measurements for the 14 participants was 2.10 ± 0.85 cm.

Ultrasound holds great promise, given its portability, relatively low cost, noninvasiveness, and clinical availability; ultrasound devices are prolifically abundant in ICUs and other units of most hospitals, and it takes less than 15 minutes to complete a measurement. Further research is needed to develop consensus on the optimal way to conduct ultrasound measurements to assess lean tissue with minimal error in clinical populations where edema, adipose stranding within muscle, and obesity in general are important considerations. The optimal site(s) for ultrasound measurement, the level of compression force to minimize error (i.e., maximal vs no or minimal compression), and the optimal application of the muscle thickness measures for lean tissue assessment have yet to be established; furthermore, the
use of simple muscle thickness measurements alone or the application of these measurements to equations to generate whole-body lean tissue values will certainly require the generation of population-specific reference data for their interpretation. Finally, once consensus on optimal protocol is reached and reference data are generated, additional research is needed to evaluate the role of ultrasound in bedside assessment of lean tissue and diagnosis of malnutrition in the hospital setting.

Applications to Clinical Case Scenarios

Scenario 1

A 67-year-old African American man is admitted to the hospital after experiencing a fall at home. He is diagnosed with a hip fracture. He admits to not eating very well over the past 6 months but is not sure if he has lost any weight. He has noticed that he is having difficulty taking care of chores around the house and is slower getting out of his chair than usual. No visible signs of muscle or fat loss are evident upon examination.

Height: 188 cm; current weight: 128.6 kg; BMI: 39.0 kg/m²

How might a bioimpedance and/or ultrasound measurement help us to evaluate this patient’s nutrition status? Although it is not yet standard practice, and there is a need for additional research to clearly establish how bioimpedance and/or ultrasound measurements might uniquely contribute to the assessment of nutrition status in the hospital setting, particularly in terms of what they tell us about muscle mass, we can...
Figure 8. Ultrasound of the quadriceps muscle. (A) Quadriceps muscle layer thickness measurements. Image illustrates the location for ultrasound readings. Reprinted with permission from Tillquist M, Katsogiannis DJ, Wischmeyer PE, et al. Bedside ultrasound is a practical and reliable measurement tool for assessing quadriceps muscle layer thickness. JPEN J Parenter Enteral Nutr. 2013;38(7):886-890. (B) Correct and incorrect positioning of the calipers for a quadriceps muscle layer thickness measurement. Image reprinted from the VALIDUM study’s ultrasound protocol with permission from Daren Heyland’s research group at the Clinical Evaluation Research Unit.
hypothetically consider how these data might be used in a particular clinical case.

Using bioimpedance

A. With any type of bioimpedance device (SF-BIA, MF-BIA, BIS), we could use the 50-kHz impedance (R and X) data to calculate Janssen’s SMI equation to evaluate for sarcopenia, although it is important to understand that the use of this equation for bedside assessment in this way has not been validated. In addition, we should recognize that the equation was developed from a fairly small population sample, using a particular SF-BIA device (ie, RJL), and if we use a different device, some error may result. Although the equation and its reference cutpoints have been published as diagnostic for sarcopenia, its use for bedside assessment has not been established and would need to be further evaluated before it can be recommended as part of standard nutrition assessment practice. As an exercise, we can consider how this might be used as one component of a comprehensive evaluation of nutrition status. In addition, we could use our 50-kHz data to evaluate his PA and SPA.

From the 50-kHz measurement, we have the following:

R = 412 ohms
X = 29.6 ohms
Z = 413 ohms
PA = 4.1°

1. Calculate SMI using the Janssen equation:\(^{12}\):

\[
SMI = [(188^2 / 412 \times 0.401) + (1 \times 3.825) + (67 \times -0.071)] + 5.102
\]

\[
SMI = 38.6 / 3.5 = 11 \text{ kg/m}^2.
\]

\[
SMI = 38.6 / 128.6 \times 100 = 30%.
\]

Interpretation: Referring to Table 1, we can compare our calculated values with the SMI by published SF-BIA cutpoints, recognizing the limitations of applying it in this way. Assuming for the moment that this has been validated for assessing muscle loss at the bedside, based on the SMI (height) cutpoint, his value of 11 kg/m\(^2\) would not suggest that he is sarcopenic, although he is not far from the 10.75 cutpoint for moderate risk of disability based on that method. Based on the SMI (%weight) cutpoint, his value of 30% does suggest that he may be sarcopenic, given that values <31% indicate class II sarcopenia. Given that he has been having difficulty with activities of daily living, this information could be considered corroborative of some degree of muscle loss and/or nutrition compromise.

2. Evaluate his PA:

Interpretation: To interpret PA, it is most appropriate to use reference data generated from a population that most closely resembles your patient in terms of ethnic background and other attributes. Ideally, we would like to have PA reference cutpoints generated by the same device that we are using and from an American reference population such as NHANES. Given that we do not currently have NHANES PA data available, the best option we have is the reference data generated in American individuals published by Barbosa-Silva et al.\(^{139}\) Referring to Table 3, we can compare his value of 4.1 with the non-BMI-specific cutpoint of 5.4, corresponding to the fifth percentile value for 60- to 69-year-olds (who were overweight on average). By these criteria, he has a low PA, which could indicate malnutrition. For the purposes of this hypothetical exercise, given the very high BMI of our patient, we might also look at the Bosy-Westphal et al.\(^{138}\) BMI-specific reference cutpoint representing the fifth percentile value for 60- to 69-year-olds, with BMI between 35.1 and 40, which is 4.7. Based on this cutpoint, we would also conclude that he has a low PA and may have some degree of malnutrition, although it may not be appropriate to apply the German white population cutpoints to our African American patient. We can also calculate his SPA.

3. Calculate SPA using the following equation:

\[
SPA = (\text{Observed PA} - \text{Mean PA}) / \text{SD},
\]

where the mean and SD are from reference values.

Referring to Table 4, we can see from the Barbosa-Silva et al.\(^{139}\) reference data for 60- to 69-year-olds that

\[
SPA = (4.1 - 7.0) / 1.1 = -2.45.
\]

If we calculate the SPA using the Bosy-Westphal et al.\(^{138}\) reference data for 60- to 69-year-old men with a BMI between 35.1 and 40 (mean PA is 6.0 and the SD is 0.85),

\[
SPA = (4.1 - 6.0) / 0.85 = -2.24.
\]

Interpretation: Given that SPA values below –1.65 are suggested to indicate malnutrition, both of our calculated SPA values, regardless of the minor differences between them, suggest that he may be malnourished.

B. If we have an MF-BIA or BIS device, we can then also evaluate the impedance ratio at 200/5 kHz.

From the MF-BIA or BIS measurement, we have the following:

Z at 200 kHz = 383
Z at 5 kHz = 453
Impedance ratio = 0.85

Interpretation: We have less to go on in terms of reference data for the 200/5-kHz impedance ratio; however, values \(\leq 0.78\) have been observed in healthy men.\(^{171}\) Furthermore, in men from various research studies, mean values of 0.81 have been associated with the development of postoperative edema,\(^{208}\) values >0.85 with worsening renal function in individuals with heart failure,\(^{174}\) and values \(>0.78\) with increasing odds of having malnutrition.\(^{171}\) In that light, we might reasonably interpret this value as corroborative of the other bioimpedance data that we have, in that it could indicate nutrition compromise and/or altered fluid status that could occur with lean tissue loss.

Putting all the bioimpedance data together, along with this patient’s history of altered functional status, we might conclude that he has at least some degree of nutrition compromise and may in fact have sarcopenic obesity.

Using ultrasound

Quadriceps muscle thickness measured with maximal compression: 2.05 cm
Quadriceps muscle thickness measured with minimal compression: 4.2 cm

Given that there is currently no consensus on the optimal protocol to follow to take an ultrasound measurement, and thus we do not yet have any reference data to draw from, at this point in time it is difficult to know how we might best use an ultrasound measurement in our patient. In the interest of this hypothetical case, however, we might consider using ultrasound measurements to track changes in lean tissue (through the measurement of quadriceps muscle thickness, following a consistent protocol) over time, after taking a baseline measurement upon admission. For the purposes of illustration, you can see that when we used a maximal compression protocol, his quadriceps muscle thickness was 2.05 cm, compared with 4.2 cm when we applied minimal compression. Thus, it is vitally important that we use the same protocol for follow-up measurements if we want to track muscle changes. At this point, our measurements cannot be easily interpreted because we have no reference data for either method, but we will remeasure the quadriceps muscle thickness every week while he is in the hospital to monitor, following consistent protocol.

Scenario 2

A 51-year-old white woman of European descent is receiving outpatient chemotherapy in combination with radiation therapy for the treatment of throat cancer. She is in her fourth week of treatment. Her intake has been quite poor, and she complains of sores in her mouth. She is generally fatigued and tires easily on exertion. She has stopped her daily walks (since the first week of treatment) and now tries to keep activity to the minimum necessary to get by. She has lost weight since starting treatment, although she is not certain how much. She appears to have some loss of fat from the orbital area but does not exhibit any other notable signs of muscle or fat loss upon physical examination.

Height: 160 cm; current weight: 70.8 kg; BMI: 28 kg/m\(^2\)

How might a bioimpedance and/or ultrasound measurement help us to evaluate this patient’s nutrition status?

Using bioimpedance

A. With any type of bioimpedance device (SF-BIA, MF-BIA, BIS), we could use the 50-kHz impedance (R and X) data to calculate Janssen’s SMI equation (and the aforementioned caveats apply), and we might also look at her PA and SPA.

From the 50-kHz measurement, we have the following:
R = 447 ohms
X = 17 ohms
Z = 447 ohms
PA = 2.2°

1. Calculate SMI using the Janssen equation\(^{12}\):

\[
\text{Skeletal muscle mass (kg)} = \frac{[\text{Height}^2 / R 	imes 0.401] + (\text{Sex} 	imes 3.825) + (\text{Age} \times -0.071)]}{\text{Height}^2 	imes 0.401} + 5.102.
\]

\[
\text{SMI} = \frac{(51 \times -0.071)]}{24.5} + 5.102 = 24.5 \text{ kg.}
\]

\[
\text{SMI (height)} = 24.5 / 2.6 = 9.4 \text{ kg/m}^2.
\]

\[
\text{SMI (%weight)} = 24.5 / 70.8 \times 100 = 34.6%.
\]

Interpretation: Referring to Table 1, we can compare our calculated values with the SMI by SF-BIA cutpoints, with the same reservations as previously discussed. For the purposes of this exercise, based on the SMI (height) cutpoint, her value of 9.4 kg/m\(^2\) does not suggest that she is sarcopenic. Based on the SMI (%weight) cut-point, her value of 34.6% also does not suggest sarcopenia.

2. Evaluate her PA:

Interpretation: Referring to Table 3, we can compare her value of 2.2 with the non-BMI-specific cutpoint of 5.5, corresponding to the fifth percentile value for 50- to 59-year-olds (who were overweight on average). By these criteria, she has a low PA, which could indicate malnutrition. For the purposes of this hypothetical exercise, given the BMI and white ethnicity of our patient, we might also look at the Bosy-Westphal et al\(^{138}\) BMI-specific reference cutpoint representing the fifth percentile.
value for 50- to 59-year-olds, with BMI between 25.1 and 30, which is 4.9. Based on this cutpoint, we would also conclude that she has a low PA and may have some degree of malnutrition. We can also calculate her SPA.

3. Calculate SPA using the following equation:

\[
SPA = (\text{Observed PA} - \text{Mean PA}) / \text{SD},
\]

where the mean and SD are from reference values.

She is 51 years old, with a BMI of 28.7, so referring to Table 4, we can see from the Barbosa-Silva et al\(^{139}\) reference data for 50- to 59-year-olds that

\[
SPA (\text{Barbosa-Silva et al}^{139}) = (2.2 - 6.6) / 0.87 = -5.06.
\]

If we calculate the SPA using the Bosy-Westphal et al\(^{138}\) reference data for 50- to 59-year-old women with a BMI between 25.1 and 30 (mean PA is 5.9 and the SD is 0.70):

\[
SPA (\text{Bosy-Westphal et al}^{138}) = (2.2 - 5.9) / 0.70 = -5.29.
\]

**Interpretation:** Given that SPA values below –1.65 have been suggested to indicate malnutrition, both of our calculated SPA values, regardless of the minor differences between them, suggest that she may be malnourished.

B. If we have an MF-BIA or BIS device, we can then also evaluate the impedance ratio at 200/5 kHz.

From the MF-BIA or BIS measurement, we have the following:

- \(Z\) at 200 kHz = 431
- \(Z\) at 5 kHz = 468
- Impedance ratio = 0.92

As stated earlier, we do not have large sample healthy reference data for the 200/5-kHz impedance ratio; however, values ≤0.82 have been observed in healthy women.\(^{171}\) Given what has been observed in the literature with impedance ratios above this cutpoint, our patient’s value of 0.92 could be interpreted as suggesting altered nutrition status and/or altered hydration status.

Taken together, the bioimpedance data could be interpreted to indicate that this patient has at least some degree of muscle loss and/or nutrition compromise.

**Using ultrasound**

- Quadriceps muscle thickness measured with maximal compression: 1.7 cm
- Quadriceps muscle thickness measured with minimal compression: 2.8 cm

In this individual, we measured a quadriceps muscle thickness using maximal compression of 1.7 cm and 2.8 cm using minimal compression. As mentioned in the first case scenario, we do not currently have appropriate reference data to interpret this cross-sectional measurement at this point in time, but we could monitor our patient weekly to observe for any changes, following a consistent protocol.

**What about using bioimpedance or ultrasound to generate whole-body lean tissue estimates at the bedside?** Generally speaking, it is not easy to know how best to obtain an accurate measure of whole-body FFM, BCM, or LST. At present, we do not have validated equations for predicting whole-body lean mass from ultrasound-measured quadriceps thickness. With regard to bioimpedance techniques, if we have access to a BIS device, we could consider generating the “normally hydrated lean tissue” (assuming we have a device that has appropriate software to model the data and calculate it from ECW and ICW values generated from the BMI-modified Xitron-Hanai–based mixture equations). Absolute accuracy of whole-body lean tissue measurements is less critical if we are interested in monitoring changes over time, although small changes may not be detectable given conservative estimates of 3%–8% MDC for bioimpedance techniques. It is important to recognize that all body composition measurement involves some degree of error due to biological variation and measurement error caused by protocol violations and human error, as well as device- and equation-specific error. If we wish to use our SF-BIA or MF-BIA device to evaluate whole-body lean tissue, we can refer to published sources\(^{78,96,104}\) and, if needed, search the literature to find a population-specific, validated equation appropriate for the individual we wish to assess, to calculate FFM, BCM, or LST; however, this is not always an easy task, and whole-body estimates are difficult to interpret, given potential violations of underlying assumptions, particularly in acute illness. If we choose to evaluate whole-body measures, our best option is to use the same device with the same predictive equation and standardize the testing conditions as much as possible for longitudinal measures (see Table 2) to minimize error.

**Caution regarding interpretation**

- We must exercise extreme caution when interpreting whole-body values coming from a device that does not provide the source of the equation used to predict the values or from a known equation that does not well match the characteristics of our patient.
- Note that we do not apply reference cutpoints generated by one bioimpedance or reference method to interpret a body composition compartment or bioimpedance parameter measured by a completely different method (eg, do not use the Geneva equation-generated FFMI reference cutpoints to evaluate your patient’s segmental MF-BIA-generated FFM value).
- We should understand that the interpretation of these various bioimpedance parameters, including PA, impedance ratio, SMI, and FFMI, is subject to the limitations of available reference data, the potential error...
caused by device-specific and population variable-specific differences between the various research laboratories producing them, and the lack of clarity over exactly what is being measured by these parameters. We must interpret findings in light of other clinical data at our disposal to build the most complete picture that we can of our patient’s nutrition status.

Summary and Call to Action: Future Research Needs

Because of the key role that the body’s lean tissue reserves play in the response to injury and illness, it is central to the maintenance of good nutrition status; this concept has been incorporated into the diagnostic criteria for malnutrition developed by AND and A.S.P.E.N.\textsuperscript{15} Clearly, given the implications that the loss of lean tissue has on clinical outcomes in acute and chronic illness, its assessment is of critical importance to the nutrition field. Our ability to measure lean tissue at the bedside has been limited by the lack of available methods that can be considered sufficiently accurate for whole-body lean tissue estimates in the clinical setting, although absolute accuracy may not be the only consideration in the application of bedside techniques as monitoring tools. Many have advocated the use of subjective physical examination techniques to evaluate muscle loss in the clinical setting. The application of these more subjective techniques is likely to be limited by lack of precision and sensitivity, although research is needed to validate nutrition-focused physical examination techniques (applied using standardized training protocols) against more objective measures. To identify muscle loss in a severely cachectic individual through nutrition-focused physical examination techniques\textsuperscript{16} is likely to be far less challenging than it would be in an overweight individual dealing with acute or chronic illness. Indeed, the detection of sarcopenia in the presence of overweight or obesity presents a significant and important challenge to the assessment of malnutrition today and is likely to require objective measurement techniques (eg, CT, as observed in the recent study by Sheean et al\textsuperscript{19}). Whether or not existing bedside techniques (eg, bioimpedance and ultrasound) can be used to routinely evaluate lean tissue loss (and thus nutrition status) in the clinical setting remains to be determined.

It is hoped that through this tutorial, the reader has gained an appreciation for the complexities involved in the validation of bedside body composition assessment methods and for the different ways that bioimpedance data are generated and applied. There is substantial and growing interest worldwide in the potential utility of simple bioimpedance variables to identify malnutrition, apart from the estimation of whole-body lean tissue compartments that have sometimes yielded wide variation in individual measures compared with reference methods, particularly in individuals with extreme obesity and other populations where assumptions may be violated. Although these violations are likely to yield some degree of error in the bioimpedance measurements, the precision (ie, from repeated measurements) of the reference method is not always evaluated in research studies, particularly those using dilution methods; lower precision in the reference method could contribute error to the methods comparison that may be inadvertently interpreted as solely coming from the bioimpedance method. A consensus on how best to evaluate the validity of a bedside method in terms of limits of agreement PE and measurement error in light of reference method precision (among other possible criteria) needs to be reached. Further refinements of the BIS technique as applied in the dialysis population\textsuperscript{184} to quantify “excess fluid” may yield improvements in the “normally hydrated lean tissue” whole-body estimates; in that event, reference data would need to be generated and made available. SMI by the Janssen equation,\textsuperscript{137} FFMI by the Geneva equation,\textsuperscript{135} and FFM by the Chumlea NHANES III equation\textsuperscript{98} are perhaps the best candidates for equations that could be cross-validated from the most recent NHANES data; reference data could be generated from NHANES for those equations proven to be most valid to provide useful cutoffpoints for sarcopenia based on American population data, ideally by BMI, age, and ethnicity. Similarly, reference data for the 50-kHz PA and 200/5-kHz (and also $Z_∞/Z_0$ if available) impedance ratio could be generated from the most recent NHANES data set. For raw bioimpedance parameters to be useful at the bedside, a consensus on normal reference values and cutpoints to define sarcopenia and/or malnutrition is needed to facilitate interpretation; this may require bioimpedance device companies to come together to somehow cross-calibrate and/or streamline technologies. In addition, recent developments in the ultrasound field are making that method an attractive option for bedside nutrition assessment. It will need additional refinement in terms of standardizing measurement protocols to minimize inter- and intraobserver errors so that good precision can be consistently achieved. In addition, reference values for normal muscle thickness as well as cutoffpoints to define sarcopenia and/or malnutrition will need to be established. How well ultrasound and/or any of the available bioimpedance techniques can contribute to nutrition assessment at the bedside (to produce whole-body estimates of lean tissue, to identify muscle loss [ie, sarcopenia], and/or to produce useful markers of malnutrition) is a vitally important question that remains to be answered.

Call to Action

The global clinical nutrition community needs to work together to come to consensus on the optimal tool(s) to use to assess nutrition status at the bedside. Clinician researchers who are interested in validating the AND/A.S.P.E.N. consensus malnutrition diagnostic criteria should strive to include at least one reference technique for lean tissue (eg, CT and/or DXA) to validate the subjective nutrition-focused physical examination method of evaluating muscle loss. In addition, at least one and ideally both of the currently available bedside techniques, ultrasound and bioimpedance (eg, an
MF-BIA and/or BIS device can produce PA and an impedance ratio; a BIS device appropriate to provide whole-body “normally hydrated lean tissue” and other whole-body estimates in addition to the simpler raw BIA parameters would have added benefit should be used in the same study to evaluate their ability to contribute uniquely to the identification of lean tissue loss and malnutrition. Collaboration and open sharing of data among researchers and cooperation between device manufacturers to streamline technologies and facilitate the production of normal reference data would allow for greater progress.

Two key questions that must be answered are:

1. Can a standardized subjective nutrition-focused physical examination method accurately detect muscle loss in individuals across the acute/chronic illness and weight spectrum compared with objective reference measures (eg, CT, DXA)?

2. In these same clinical populations, can ultrasound or a particular bioimpedance approach help us to identify and monitor subtle changes in lean tissue (and thus nutrition status) compared with reference measures that would not have otherwise been possible through other, more subjective assessment methods?

Acknowledgments

The author thanks David Frankenfield for sharing his insights on precision, accuracy, and bias as they relate to the comparison of energy expenditure methodologies, as well as how to discuss the interpretation of Bland-Altman and linear regression plots in terms of precision and bias. Figures 1 and 2 were developed from his unpublished figures, with his permission. The author also recognizes Daren Heyland, Marina Mourtzakis, and Michael Paris for their input on ultrasound techniques and Adam Kuchnia, Urvashi Mulasi, Levi Teigen, and Abigail Cole for their input and reflection on the presentation of key concepts. In addition, the author thanks the reviewers who gave their time and care in providing feedback and input on this manuscript.

Statement of Authorship

C. P. Earthman was responsible for the conception/design of the research; was responsible for the acquisition, analysis, and interpretation of the data; drafted the manuscript; critically revised the manuscript; agrees to be fully accountable for ensuring the integrity and accuracy of the work; and read and approved the final manuscript.

Glossary

Accuracy: refers to the closeness of agreement in a particular variable between 2 assessment methods.

BCM: body cell mass.

Bias: refers to the systematic error, determined by the average differences in a particular measurement variable between the bedside and reference method.

BIS: bioimpedance spectroscopy.

Cachexia: among many definitions: significant weight loss, protein catabolism, and muscle and fat tissue loss that occur due to underlying disease processes.

ECW: extracellular water.

FFM: fat-free mass.

FFMI: fat-free mass index.

Fixed bias: refers to the type of systematic error that occurs when a method yields measurements of a particular variable that are consistently higher or lower than those taken by the reference method.

ICW: intracellular water.

Impedance ratio: most commonly studied is the ratio of impedance measured at 200 kHz to 5 kHz; other high- to low-frequency impedance ratios may be useful, including $Z_\infty/Z_0$ from BIS measurements.

LST: lean soft tissue (term used in the DXA literature for the bone-free FFM, which is also called lean body mass).

MF-BIA: multiple-frequency bioelectrical impedance analysis.

PA: phase angle (most commonly studied is the PA measured at 50 kHz).

Precision: refers to the degree of agreement among repeated measurements by a method, determined by repeatability and reproducibility.

Proportional bias: refers to the type of systematic error that is proportional to the value of the variable being measured.

Repeatability: refers to the degree of agreement between repeated, independent measurements of particular variable made by the same operator using a single method in the same individual under the same conditions; is affected by intraobserver variability, among other factors.

Reproducibility: refers to the degree of agreement in measurements of a particular variable taken by a single method but by different operators; is affected by interobserver variability, among other factors.

Sarcopenia: loss of muscle mass with loss of muscle strength and/or function.

SF-BIA: single-frequency bioelectrical impedance analysis.

SMI: skeletal muscle mass index.

TBW: total body water.

Reading List


References


Kusheer RF, de Vries PM, Gudivaka R. Use of bioelectrical impedance analysis measurements in the clinical management of patients undergoing dialysis. Am J Clin Nutr. 1996;64(3(suppl)):503S-509S.


