The validity of ultrasound-derived equation models to predict whole-body muscle mass

Van den Broeck, Jona; Buzzatti, Luca; Jager-Wittenaar, Harriët; Perkisas, Stany; Scafoglieri, Aldo

Published in:
Clinical Nutrition ESPEN

DOI:
10.1016/j.clnesp.2021.08.012

Publication date:
2021

Clinical Nutrition ESPEN 46 (2021) 133–141

Contents lists available at ScienceDirect

Clinical Nutrition ESPEN

journal homepage: http://www.clinicalnutritionespen.com

Meta-analysis

The validity of ultrasound-derived equation models to predict whole-body muscle mass: A systematic review

Jona Van den Broeck [a, *], Luca Buzzatti [a, b], Harriët Jager-Wittenaar [a, c, d], Stany Perkisas [e], Aldo Scafoglieri [a]

[a] Vrije Universiteit Brussel, Faculty of Physical Education and Physiotherapy, Department Physiotherapy and Human Anatomy, Research Unit Experimental Anatomy, Laarbeeklaan 103, 1090, Brussels, Belgium
[b] School of Allied Health, Anglia Ruskin University (ARU), Cambridge, UK
[c] Research Group Healthy Ageing, Allied Health Care and Nursing, Hanze University of Applied Sciences, Groningen, the Netherlands
[d] Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands
[e] University of Antwerp, University Geriatric Center, Leopoldstraat 26, 2000, Antwerp, Belgium

ARTICLE INFO

Article history:
Received 11 February 2021
Accepted 17 August 2021

Keywords:
Ultrasound
Equation
Prediction
Muscle mass

SUMMARY

Background & aims: Sarcopenia is defined as the age-related loss in muscle quantity and quality which is associated with physical disability. The assessment of muscle quantity plays a role in the diagnosis of sarcopenia. However, the methods used for this assessment have many disadvantages in daily practice and research, like high costs, exposure to radiation, not being portable, or doubtful reliability. Ultrasound has been suggested for the estimation of muscle quantity by estimating muscle mass, using a prediction equation based on muscle thickness. In this systematic review, we aimed to summarize the available evidence on existing prediction equations to estimate muscle mass and to assess whether these are applicable in various adult populations.

Methods: The databases PubMed, PsycINFO, and Web of Science were used to search for studies predicting total or appendicular muscle mass using ultrasound. The methodological quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies, version 2 (QUADAS-2) and the quality assessment checklist (QA) designed by Pretorius and Keating (2008).

Results: Twelve studies were included in this systematic review. The participants were between 18 and 79 years old. Magnetic Resonance Imaging and dual-energy X-ray absorptiometry were used as reference methods. The studies generally had low risk of bias and there were low concerns regarding the applicability (QUADAS-2). Nine out of eleven studies reached high quality on the QA. All equations were developed in healthy adults.

Conclusions: The ultrasound-derived equations in the included articles are valid and applicable in a healthy population. For a Caucasian population we recommend to use the equation of Abe et al., 2015. While for an Asian population, we recommend to use the equation of Abe et al., 2018, for the South American population, the use of the equation of Barbosa-Silva et al., 2021 is the most appropriate.

1. Introduction

Loss of muscle mass has been associated with poor quality of life, physical disability, increased health care costs, and mortality [1–5]. According to the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), sarcopenia is defined as a progressive and generalised skeletal muscle disorder that is associated with an increased likelihood of adverse outcomes including falls, fractures, physical disability, and mortality [1]. Sarcopenia can be confirmed when low muscle strength in combination with either low muscle quantity or quality is detected [1]. Although sarcopenia is a well-studied disorder, it is still difficult to diagnose because muscle quantity and quality are technically difficult to measure accurately in a clinical setting [6]. Muscle quantity can be defined as muscle

* Corresponding author. Department of Physiotherapy, Human Physiology, and Anatomy (KIMA), Vrije Universiteit Brussel, Laarbeeklaan 103, 1090, Brussels, Belgium.
E-mail address: jona.van.den.broeck@vub.be (J. Van den Broeck).

https://doi.org/10.1016/j.clnesp.2021.08.012

mass [1]. Muscle quality can be defined as the micro- and macroscopic change in muscle architecture and composition, and is associated to muscle functions delivered per unit of muscle mass [1].

Muscle quantity and quality can be measured with magnetic resonance imaging (MRI), computed tomography (CT), dual-energy X-ray absorptiometry (DXA), and bioelectrical impedance analysis (BIA) [1]. MRI and CT are considered gold standards to measure muscle quantity and quality [1,6–8]. Both methods showed to be precise, reliable, and safe, but the disadvantages are that MRI and CT are expensive, involve extensive measurements, not portable, and cannot measure oversized people [1,6,7]. In addition, participants are exposed to radiation during a CT-scan [9]. DXA can be used to measure lean mass [10]. Lean mass includes skeletal muscle mass (SMM) as well as non-skeletal components such as skin, connective tissue, and the fat-free component of adipose tissue cells (FFAT) [11]. DXA is safe, requires minimal participation of the participant, is fast, and easy to use [12]. Disadvantages of DXA are the high costs, exposure to radiation, limited availability of equipment, and not being portable [6,13]. Although BIA is considered an indirect method, it also can be used to assess muscle quantity [1,14]. BIA is a cheap and portable field method that measures the electrical impedance of an electric current passing through the body, by which muscle mass, lean mass, or fat-free mass can be estimated using a prediction equation [15]. However, the reliability of these equations can be negatively influenced by various factors related to the instrument itself, the electrodes, the operator, the participant, and the environment [9]. Similarly, since prediction equations for BIA to assess muscle quantity rely on the assumption that body hydration status remains constant, the validity of BIA to assess muscle quantity is limited to an individual level, and the accuracy has been demonstrated to be less compared to DXA and MRI [1,9].

Ultrasound, especially Brightness mode (B-mode) ultrasound, has been suggested as a method for the estimation of muscle quantity and quality [14,16]. Ultrasound is a non-invasive, relatively cheap, portable, readily available, and safe method to estimate muscle thickness, cross-sectional area (CSA), and echo intensity (EI) [6,14]. Ultrasound can detect changes in muscle thickness and CSA, which means this technique could potentially be used in clinical practice [14]. EI is defined as the brightness of the image acquired and expresses the fat infiltration and fibrotic changes into a muscle in greyscales [17]. Muscle thickness and CSA are parameters that have been used to estimate muscle quantity [1]. The measurement of muscle thickness has been shown to be highly reliable in young and older populations and could, therefore, be used in equations to predict muscle mass [17–19]. However, knowledge is lacking about the validity of these equations to determine how useful they are in practice.

To the best of our knowledge, no systematic review assessing the validity and applicability of these ultrasound-derived equations that use muscle thickness to estimate muscle quantity is available. With the applicability, we wanted to assess in what kind of population the equation can be used, which sites should be used in the equation, and how many sites should be included. Therefore, we aimed to assess the validity and applicability of these ultrasound-derived equations that use muscle thickness to estimate muscle mass.

2. Methods

2.1. Eligibility criteria and search strategy

Studies predicting total or appendicular muscle mass in adults (18 years or older) were included in this systematic review. The ultrasound-derived equations had to be based on ultrasound measurements validated against a criterion method for muscle mass described by the EWGSOP2 (i.e., MRI, CT, DXA, or BIA) [1]. Articles written in English, Dutch, or French were included when published after 1990. Excluded articles were reviews, animal studies, and studies using cadaver specimens. Studies that described an equation to predict body density or fat mass were also excluded.

A combination of terms related to muscle mass, ultrasound, and equation was used in the search strategy: ("muscle mass" OR "muscle thickness" OR "muscle volume" OR "muscle weight" OR "muscle quantity" OR "cross-sectional area" OR "lean mass" OR "lean body mass" OR "lean soft tissue" OR "appendicular muscle mass" OR "appendicular lean mass" OR "fat-free mass" OR "adipose tissue free mass" OR "non-fat tissue") AND ("ultrasound" OR "ultraso* imaging" OR "medical sonography" OR "sonography" OR "echography" OR "ultrasonography" OR "ultrasonographic technique") AND ("equation" OR "validity" OR "formula" OR "prediction"). The databases PubMed, PsycINFO, and Web of Science were used to search for articles up until January 14, 2021. Rayyan [20] was used to import retrieved articles from the databases. After deleting duplicates, three authors (J.V.d.B., L.B., and A.S.) screened the titles, abstracts, and full texts independently. Depending on the inclusion and exclusion criteria, the abstracts were independently scored as relevant or not relevant. In case of disagreement, discussions were held between the three authors (J.V.d.B., L.B., and A.S.) to reach a consensus after reading the full text.

2.2. Study appraisal and synthesis methods

The methodological quality of the included studies was assessed using the Quality Assessment of Diagnostic Accuracy Studies, version 2 (QUADAS-2) [21]. The quality of the ultrasound measuring method was assessed using the quality assessment checklist (QA) designed by Pretorius and Keating (2008) [7]. The QUADAS-2 is a tool for systematic reviews to examine the diagnostic accuracy of the included studies [21]. QUADAS-2 comprises four domains; patient selection, index test, reference standard, and flow and timing [21]. The domains are assessed in terms of risk of bias and concerns regarding applicability [21]. For each domain, the result can be 'low risk of bias', 'high risk of bias', or 'unclear risk of bias' [21]. The QA checklist of Pretorius and Keating (2008) [7] is designed to score the quality of a study by assessing the methodology of ultrasound muscle measurement. The checklist includes ten questions that can be answered with 'yes', 'no', or 'not stated/not clear'. For every 'yes', the study gets 1 point [7]. In case of 'no' or 'not stated/not clear', the study gets 0 points [7]. A score of 7 points or higher indicates that the method is of high quality [7]. Two authors (J.V.d.B. and L.B.) scored the studies independently and discussed in case of disagreement with a third author (A.S.).

3. Results

The flow chart in Fig. 1 shows that out of 2994 selected studies, twelve studies were included in this systematic review.

3.1. Study quality

The diagnostic accuracy of the studies is shown in Table 1. For patient selection, all studies had low risk of bias and low concerns regarding the applicability [8,10–13,16,22–27]. All studies had a low risk of bias and low concern regarding the applicability of the index test [8,10–13,16,22–27]. Most studies had an unclear risk regarding the reference standard [10–13,16,23–26] because, in literature, there is no consensus that DXA can be used as a gold

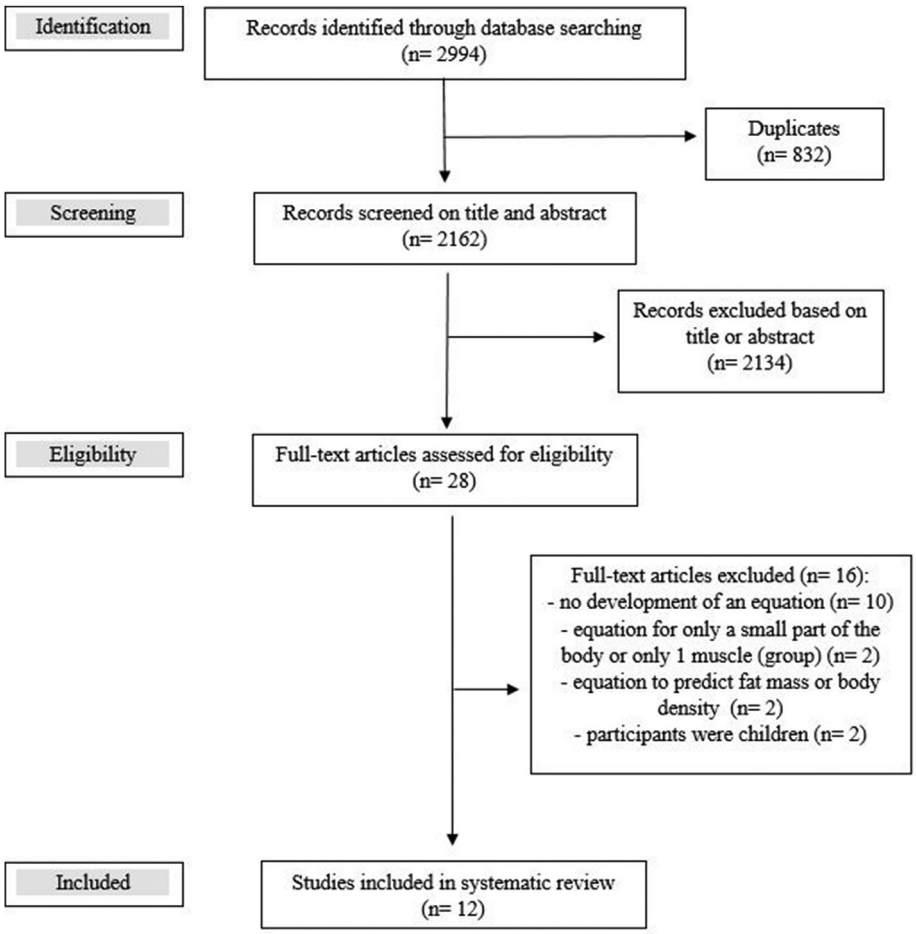

Fig. 1. Flow chart of the selection procedure for including studies.

standard reference for muscle mass [28]. For the flow and timing, there was only low risk of bias [8,10–13,16,22–27].

The score for quality of the included articles, using the QA checklist, ranged between 6 and 8 points. Nine out of eleven studies used a methodology of high quality (i.e. score 7 or higher) [10–13,16–26]. These studies did not reach the maximum score because the assessors were not blinded [10–13,22–26], the muscles were not relaxed [10–13,16,22–25], the timeframe was not stated [12,13,16,22,23], or the position of the transducer was not described [10,11,16,25,26]. Although the method of the studies often stated that the muscles were relaxed, we still chose to indicate that this was not the case. The reason for this decision is that we do not associate the standing position, in which the participants were measured, with an actual relaxed posture. When the position of the transducer is not described, it is unknown in which plane the scan was taken and how large the angle between the skin and the transducer was. The studies that scored 6 or lower did not take into account the aforementioned items and did not describe the contact pressure of the transducer [8,27] and/or did not describe the transducers' position [27].

3.2. Measuring method

Study methods are shown in Table 2. All studies used a portable ultrasound and a linear transducer [8,10–13,16,22–27]. Eight out of twelve studies performed the measurements with the participants standing [10–13,22–25]. In two other studies, the participants were in a lying position [16,26] and in the remaining two studies, the position of the participants was not described [8,27]. The muscle thickness was measured with the probe perpendicular to the tissue interface at the marked sites [8,10–13,16,22,23,25,27]. The muscle thickness was determined by the distance from the adipose tissue–muscle interface to the muscle–bone interface [8,10–13,16,22–25,27]. For muscles without a bony surface, the upper and lower boundaries of the muscle fascia were used [8,11,13]. One study also used adipose tissue thickness in the equation [11]. The adipose tissue thickness was measured using the same measuring method as described for the muscle thickness and was determined by the distance between the skin surface and the subcutaneous adipose tissue–muscle interface [11].

3.3. Population

The characteristics of the population are shown in Table 2. The participants in the included studies were healthy adults [8,10–13,23–25], athletes [22,27], or community-dwelling adults [16,26] with a Caucasian [13,16,24], Asian [8,10–12,23,25], South American [26] or mixed [22,25] ethnicity. Most studies had a mixed sample of men and women [8,10–13,16,22–24,26]. In only two studies only men [27] or only women [25] were included. The age of the participants ranged between 18 and 79 years old [8,10–13,16,22–27].

Table 1
Quality assessment.

	Abe et al. (2018) [10]	Abe et al. (2015) [13]	Abe et al. (2018) [11]	Abe et al. (2015) [24]	Abe et al. (2015) [25]	Paris et al. (2017) [16]	Abe et al. (2016) [23]	Takai et al. (2014) [12]	Sanada et al. (2006) [8]	Toda et al. (2016) [27]	Abe et al. (2019) [22]	Barbosa-Silva et al. (2021) [26]
Result QUADAS-2												
Patient selection												
Risk of bias	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Concerns applicability	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Index test												
Risk of bias	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Concerns applicability	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Reference standard												
Risk of bias	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Low	Unclear	Low	Unclear
Concerns applicability	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Flow and timing												
Risk of bias	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Result QA	7	7	7	8	7	7	7	7	6	6	7	7

QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies, version 2, QA = Quality Assessment (Pretorius et al., 2008 [7]).

3.4. Reference methods and outcome measures

The ultrasound-derived equations were validated against MRI [8,22,27] or DXA [10–13,16,23–26].

The outcomes of the prediction equations, shown in Table 3, were appendicular lean mass (aLM) [10,13,16,23,26], aLM minus fat-free adipose tissue (aLM-FFAT) [10,11], fat-free mass (FFM) [12,24], SMM [8,22,27], and total muscle mass (TMM) [24,25].

3.5. Variables in the equations and applicability

The variables that were included in the prediction equations are shown in Table 3. The variables most used in the equations were muscle thickness of various sites or muscles of the body, sex, and height [8,10–13,16,22–27]. For the studies that described the standard error of estimate, it ranged between 0.19 kg and 2.9 kg [8,11–13,16]. The mean absolute deviation, described in another study, ranged from 1.0 kg to 1.1 kg [10]. The root mean square error in the study of Barbosa-Silva et al., 2021 [26] was 1.3 kg. The coefficient of determination was in some studies described as R^2, while in other studies it was described as the adjusted R^2. The R^2 ranged from 0.87 to 0.96 [8,10–13] and the adjusted R^2 ranged from 0.72 to 0.97 [10,11,13,16,26]. In the study of Abe et al., 2015 [13], 15 different equation models were developed. Table 3 shows the standard error of estimate (SEE) of the best equation (lowest SEE), the most practical equation (least number of parameters), and a cross-validated equation (cross-validated in another study). In Table 3 only the studies that developed an equation are described. Table 4 shows the results of the studies that performed cross-validation. Seven studies performed cross-validation [8,13,22–25,27], using the equations of previously described equations [8,10,12,13,27]. The SEE in the cross-validated studies

Table 2
Methods of the studies included in this systematic review.

Author	Type of ultrasound (brand)	Type of transducer (frequency)	Participant position	Participants					Reference method
				Number	Population	Ethnicity	Sex	Age range	
Abe et al., 2018 [10]	Portable (SSD-500, Aloka)	Linear (5 MHz)	Standing	389	Healthy adults	Asian	M and F	60-79y	DXA
Abe et al., 2015 [13]	Portable (SSD-500, Aloka)	Linear (5 MHz)	Standing	Development: 71 Cross-validation: 31	Healthy adults	Caucasian	M and F	50-76y	DXA
Abe et al., 2018 [11]	Portable (SSD-500, Aloka)	Linear (7.5 MHz)	Standing	Development: 215 Cross-validation: 96	Heathy adults	Asian	M and F	60-79y	DXA
Abe et al., 2015 [24]	Portable (SSD-2000, Aloka)	Linear (5 MHz)	Standing	79	Healthy adults	Caucasian	M and F	50-78y	DXA
Abe et al., 2015 [25]	Portable (SSD-500, Aloka)	Linear (5 MHz)	Standing	41	Healthy adults	Mixed	F	50-78y	DXA
Paris et al., 2017 [16]	Portable (M-Turbo, SonoSite, Markham, ON)	Linear (5–10 MHz)	Lying prone or supine	96	Community dwelling adults	Caucasian	M and F	24-72y	DXA
Abe et al., 2016 [23]	Portable (Pro-Sound 2, Aloka)	Linear (7.5 MHz)	Standing	158	Healthy adults	Asian	M and F	50-79y	DXA
Takai et al., 2014 [12]	Portable (SSD-900, Aloka)	ND	Standing	77	Healthy adults	Asian	M and F	52-78y	DXA
Sanada et al., 2006 [8]	Portable (SSD-500, Aloka)	Linear (5 MHz)	ND	Development: 48 Cross-validation: 24	Healthy adults	Asian	M and F	18-61y	MRI
Toda et al., 2016 [27]	Portable (ND)	ND	ND	Development: 40 Cross-validation: 21	Healthy athletes	Asian	M	average 20y	MRI
Abe et al., 2019 [22]	Portable (SD-500, Aloka)	Linear (5 MHz)	Standing	23	Healthy athletes	Mixed	M and F	21y	MRI
Barbosa-Silva et al., 2021 [26]	Portable (Xario SSA-660A)	Linear (5–12 MHz)	Supine	192	Community dwelling adults	South American	M and F	>60y	DXA

MHz = Megahertz, M = male, F = female, DXA = dual-energy X-ray absorptiometry, ND = not described, MRI = magnetic resonance imaging.

Table 3
Ultrasound-derived equations to predict muscle mass, developed by the included studies.

Author	Equation outcome	Number of parameters	Equation	Abbreviations	SEE (kg)	MAD (kg)	RSME (kg)	Coefficient of determination (R^2)	Coefficient of determination (Adj. R^2)
Abe et al., 2018 [10]	aLM - FFAT	5	Mixed: - 7.9116 (sex × 5.1693) + (age × 0.0345) + (MTFA × height × 2.2752) + (MTUA × height × 0.0743) + (MTTA × height × 0.4927) + (MTLA × height × 1.4892) − (sex × age × 0.0380) − (sex × MTFA × height × 0.3379) - (sex × MTUA × height × 0.1263) − (sex × MTTA × height 0.1754) − (sex × MTLA × height × 0.3083)	MTFA = MT Forearm Anterior MTUA = MT Upper arm Anterior MTTA = MT Thigh Anterior MTLA = MT Lower leg Anterior Sex: M = 0, F = 1		1.0		0.90	0.90
		2	Mixed: - 2.0940 + (sex × 4.1273) − (age × 0.0094) + (MTFA × height × 3.5599) − (sex × age × 0.0307) − (sex × MTFA × height × 0.8349)			1.1		0.87	0.87
Abe et al., 2015 [13]	aLM	7	Mixed best model: (0.058 × age) + (0.98 × MTFA × height) + (1.13 × MTUA × height) + (0.45 × MTP × height) + (0.48 × MTTA × height) + (0.89 × MTLA × height) + (0.77 × MTLP × height) + (0.44 × MTTP × height) − 16.11	MTFA = MT Forearm Anterior MTUA = MT Upper arm Anterior MTP = MT trunk Posterior MTTA = MT Thigh Anterior	1.1			0.97	0.97
		1	Mixed most practical model: (10.90 × MTFA) − 18.83	MTLA = MT Lower leg Anterior	2.3			0.88	0.88
		2	Mixed other model: (4.89 × MTFA × height) − 9.15	MTLP = MT Lower leg Posterior MTTP = MT Thigh Posterior	2.0			0.91	0.91
Abe 2018 et al. [11]	aLM - aFFAT	4	Mixed: [(4.89 × MTFA × height) − 9.15] − [(0.472 × sex) + (0.867 × height) + (0.392 × ATFA) + (0.77 × BMI) − 3.01]	MTFA = MT Forearm Anterior ATFA = Adipose tissue Thickness Forearm Anterior BMI= Body Mass Index	0.19			0.76	0.75
Paris et al., 2017 [16]	aLM	4	Mixed:3.895 + (0.100 × (($MTRF_R$ + $MTRF_L$ + $MTVI_R$ + $MTVI_L$)/4))	$MTRF_R$ = MT Rectus Femoris right	2.9				0.72
		6	Mixed: − (1.985 × sex) + (0.0247 × age) + (height × (MTUA + (1.555 × (($MTRF_R$ + $MTRF_L$ + $MTVI_R$ + $MTVI_L$)/4))))) + 2.929	$MTRF_L$ = MT Rectus Femoris left $MTVI_R$ = MT Vastus Intermedius right $MTVI_L$ = MT Vastus Intermedius left Sex: M = 0, F = 1 MTUA = MT Upper Arm	1.6				0.91
Takai et al., 2014 [12]	FFM	4	Mixed: (sex × 7.217) + (MTTA × 1.985) + (MTTP × 2.355) + (MTLA × 3.633) + (MTLP × 2.670) − 6.759	Sex: M = 1, F = 2 MTTA = MT Thigh Anterior	2.5			0.93	
		8	Mixed: (sex × 5.233) + ((MTUA × LLUA) × 0.006630) + ((MTTA × LLTA) × 0.05153) + ((MTTP × LLTP) × 0.05579) + ((MTLP × LLLP) × 0.07097) + 1.774	MTTP = MT Thigh Posterior MTLA = MT Lower leg Anterior MTLP = MT Lower leg Posterior MTUA = MT Upper Arm LLUA = LL Upper Arm LLTA = LL Thigh Anterior LLTP = LL Thigh Posterior LLLP = LL Lower leg Posterior	2.0			0.96	
Sanada et al., 2006 [8]	SMM	10	Males: 0.641 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) × height − 12.087	MTFL = MT Forearm Lateral MTUA = MT Upper arm Anterior MTUP = MT Upper arm Posterior	2.2				
		7	Males: 0.809 × MT6 × height − 4.834	MTA = MT trunk Anterior	1.8				
		10	Females: 0.594 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) × height − 11.320	MTP = MT trunk Posterior MTTA = MT Thigh Anterior MTTP = MT Thigh Posterior	2.8				
		7	Females: 0.831 × MT6 × height − 7.992	MTLA = MT Lower leg Anterior MTLP = MT Lower leg Posterior MT6 = sum of muscle thicknesses of 6 sites (not defined which)	2.9				

(*continued on next page*)

Table 3 (*continued*)

Author	Equation outcome	Number of parameters	Equation	Abbreviations	SEE (kg)	MAD (kg)	RSME (kg)	Coefficient of determination (R^2)	Coefficient of determination (Adj, R^2)
Toda et al., 2016 [27]	SMM	10	Mixed: 0.645 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) × height − 7.821	MTFL = MT Forearm Lateral MTUA = MT Upper arm Anterior MTUP = MT Upper arm Posterior MTA = MT trunk Anterior MTP = MT trunk Posterior MTTA = MT Thigh Anterior MTTP = MT Thigh Posterior MTLA = MT Lower leg Anterior MTLP = MT Lower leg Posterior					
Barbosa-Silva et al., 2021 [26]	aLM	6	Mixed: 3.24 × sex +16 × height + 0.2 × arm length + 0.09 × dominant arm circumference + 0.04 × dominant thigh circumference + 1.25 × dominant MTUA + 0.72 × dominant MTTA − 24.9	Sex: M = 1, F = 0 MTUA = MT Upper arm Anterior MTTA = MT Thigh Anterior			1.3		0.89

SEE = standard error of the estimate, MAD = Mean absolute deviation, RSME = Root Mean Square Error, aLM = appendicular lean mass, FFAT = fat-free adipose tissue, MT = muscle thickness, M = male, F = female, aFFAT = appendicular fat-free adipose tissue, FFM = fat-free mass, LL = limb length.

anged from 0.3 kg to 2.8 kg [23,24]. The total error ranged from 1.5 kg to 2.0 kg [13,22]. The coefficient of determination R^2 ranged from 0.83 to 0.96 [24].

4. Discussion

In this systematic review, all included studies that developed equations for the prediction of muscle mass were of high quality. In most of the included studies, the ultrasound-derived equations to estimate muscle mass were validated against DXA as a reference method [10–13,16,23–26]. All included studies used multiple muscle thicknesses of the arm and leg in the equations to predict muscle mass [8,10–13,16,22–27].

4.1. Measuring method

In most studies, the measurement procedure was well described and similar. However, the measurement procedure used in the studies [10–13,22–25] differs from an ultrasound measurement method, developed later by Perkisas et al., 2021 [17]. When measuring muscle quantity, using MRI or DXA, the participants are in a supine position with the muscles relaxed [8,16]. However, in the included studies, assessment of muscle thickness with ultrasound was performed in a standing position [10–13,22–25]. According to Thoirs et al., 2009 [29], it makes no difference to the intra-rater reliability if the participants are tested lying down or standing. However, the measurement results taken with the participants lying down were smaller compared to the measurement results with the participants standing [29]. The reasons for the differences in measurements may be due to postural or positional forces acting on muscle shape, physiological changes, and the effect of gravitational forces, joint position, and degree of muscle activation [29]. Because of these differences in muscle measurement between the two positions, it is important that the participants are measured in the standing position if the equation was originally developed in the standing position to minimize errors [29]. But considering their higher reliability, minimal variability, and easier measurement with recumbent participants, equations developed in supine position should be preferred [29]. Muscles can't be completely relaxed when a participant is standing [30]. There is always a minimum of muscle activity to stand upright [30]. In addition, during the measurement with MRI or DXA, the participant is always lying supine. During these examinations, the muscles need to be completely relaxed to avoid measurement artefacts.

In the included articles, muscle thickness was measured at nine sites of the body: at the lower leg anterior and posterior, upper leg anterior and posterior, trunk anterior and posterior, upper arm anterior and posterior, and forearm lateral. For the applicability of the ultrasound-derived equations, it is important that the interrater reliability is high, which has been shown previously [31–34]. According to Perkisas et al., 2021 [17], muscle thickness should be measured of a single muscle instead of a site of the body. Measuring a site of the body means that when measuring for example the thigh anterior, the rectus femoris muscle and the intermedius muscle are measured simultaneously. When measuring only a single muscle, only the rectus femoris is measured. Measuring a single muscle is more difficult than measuring a site because when measuring a single muscle it is essential to measure the maximal muscle bulk [17]. For every muscle the point of the maximal muscle bulk is different and no consensus about these points for all different muscles is available. Thus, for whole body composition purposes it is easier to measure muscle sites instead of individual muscles.

Table 4
Results of the cross-validation of the ultrasound-derived equations to predict muscle mass.

Author	Equation outcome	Equation	SEE (kg)	TE (kg)	RMSE (kg)	Coefficient of determination (R^2)
Abe et al., 2015 [13]	aLM	Equations from Abe et al. 2015 [13]		1.5		
		Mixed best model: (0.058 × age) + (0.98 × MTFA x height) + (1.13 × MTUA x height) + (0.45 × MTP x height) + (0.48 × MTTA x height) + (0.89 × MTLA x height) + (0.77 × MTLP x height) + (0.44 × MTTP x height) − 16.11		2.5		
		Mixed most practical model: (10.90 × MTFA) − 18.83		2.2		
		Mixed other model: (4.89 × MTFA x height) − 9.15				
Abe et al., 2015 [24]	TMM	Equations from Sanada et al. 2006 [8]	2.2			0.92
	TMM	Males: 0.641 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) x height − 12.087	2.8			0.83
	FFM	Females: 0.594 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) x height − 11.320	0.3			0.96
		Equation from Takai et al. 2014 [12]				
		Mixed: (sex x 5.233) + ((MTUA x LLUA) x 0.006630) + ((MTTA x LLTA) x 0.05153) + ((MTTP x LLTP) x 0.05579) + ((MTLP x LLLP) x 0.07097) + 1.774				
Abe et al., 2015 [25]	TMM	Equation from Sanada et al. 2006 [8]	1.1			
		Females: 0.594 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) x height − 11.320				
Abe et al., 2016 [23]	aLM	Equations from Abe et al. 2015 [13]		2.6		0.78
		Mixed most practical model: (10.90 × MTFA) − 18.83		1.4		0.89
		Mixed other model: (4.89 × MTFA x height) − 9.15				
Sanada et al., 2006 [8]	SMM	Equation from Sanada et al. 2006 [8]				0.94
		Not defined which				
Toda et al., 2016 [27]	SMM	Equation from Toda et al. 2016 [27]				
		Mixed: 0.645 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) x height − 7.821				
Abe et al., 2019 [24]	SMM	Equation from Sanada et al. 2006 [8]		1.5		
		Males: 0.641 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) x height − 12.087		1.5		
		Females: 0.594 × (MTFL + MTUA + MTUP + MTA + MTP + MTTA + MTTP + MTLA + MTLP) x height − 11.320				

SEE = standard error of estimate, TE = Total Error, RMSE = Root Mean Square Error, aLM = appendicular lean mass, TMM = total muscle mass, FFM = fat-free adipose tissue SMM = skeletal muscle mass, ND = not described.

4.2. Population

In all included articles, only healthy participants were used to create an equation. It can therefore be said that these equations can be used in this population, but their use may not be automatically applicable in other populations. However, in the studies of Abe et al., 2018 and 2015 [10,13], 7% [13] and 20% [10] of the sample were participants who had a muscle mass score below the cut-off value for low muscle mass (7.25 kg/m^2 for men and 5.67 kg/m for women) [35]. This means that these equations may be used to support the diagnosis of low muscle mass [36]. Further research is needed for equations in non-healthy populations. Based on the results of the cross-validation of the equations in samples with different ethnicities, it seems questionable whether ethnicity plays a role in the different ultrasound-derived equations or not. In these studies, equations [8,12,24] were applied to a sample of a different ethnicity [10,11,22,23] and still showed a good coefficient of determination and low measurement errors. However, we think that the use of an equation in a population with a different ethnicity should be interpreted with caution since ethnicity could influence the accuracy of the equation due to differences in body tissue distribution [24].

Furthermore, it is unclear whether age is an important variable in the equations. Most of the equations were developed for a specific age group (e.g. > 50 years old), but some equations were developed for participants with a large variability in age. Ageing affects muscles [37,38] and thus could affect the standard error of estimate or total error. Unfortunately, there is no consensus on this discussion and so we encourage future researchers to investigate whether age plays a role in the use of these equations. In this context, the question can be asked whether the equations should be created per age category or whether one equation is sufficient if the variable 'age' is included in the equation itself.

4.3. Reference methods and outcome measures

Clinicians should be aware of the limitations of the body composition technique DXA, as lean mass includes skeletal components and non-skeletal components [10,11,28]. These non-skeletal components could cause an overestimation of the muscle mass because of the included FFAT, especially in clinical populations [10]. The explanation for choosing DXA lies in the fact that DXA is more feasible, safer, and cheaper compared to a gold standard [9]. The outcomes of the equations using DXA were aLM and aLM minus FFAT. The aLM is the sum of the SMM and the non-skeletal components (i.e. skin, connective tissue, and FFAT) in the limbs [10,13,23]. Since the main goal is to estimate SMM, some researchers have attempted to subtract the FFAT from the aLM to make better estimations of the SMM [10,13,23]. FFAT was calculated based on the methods of Heymsfield et al., 2002 [39]. Heymsfield et al., 2002 [39] reported that 85% of adipose tissue is fat and 15% of adipose tissue is the remaining calculated fat-free component. To calculate adipose tissue mass, Abe et al., 2018 [10] used DXA-determined fat mass (adipose tissue = fat mass ÷ 0.85). Then, FFAT can be calculated as FFAT = adipose tissue × 0.15 [10]. However, these equations only make an estimation, which compromises the validity of the final ultrasound-derived equation. Because only aLM has been predicted, the muscle mass of the trunk was not included in these equations. The absence of trunk muscle variables in the prediction should not be a problem, according to Zhao et al., 2013 [40],

since the total body SMM, measured with MRI, correlates well with the aLM, measured with DXA. But according multiple other studies, this might be an important limitation because, the greater rates of age-related losses of muscle mass occur in the lower leg, thigh, and lower trunk [8,12,41]. As such, the loss of muscle mass is less in the upper-trunk and arm regions [8,12]. Moreover, it has been suggested that loss of muscle mass in the lower body regions is physical activity-related, while the loss of muscle mass in upper body regions is more diet-related [42].

4.4. Variables in the equations and applicability

The amount of included items in the ultrasound-derived equations differ from one study to the other. The most frequently used muscle sites for the measurement of muscle thickness were: the upper arm anterior [8,10–12,26,27], thigh anterior [8,10–12,26,27], lower leg anterior [8,10,11,27], and lower leg posterior [8,11,12,27].

In addition to muscle thickness, some equations incorporated height, sex, and age. According to Abe et al., 2014 [36], height should be included in the equations to make sure the equation can be used in different ethnicities because, for instance, Caucasian adults are generally taller than other ethnicities [27, 36]. Height is indeed incorporated in the equations with the lowest SEE in the development studies [10, 11]. The most applicable equation in clinical practice is the equation that yields the least measurements as possible, the highest possible coefficient of determination and the smallest possible prediction error. Abe et al., 2018 [10], for example, developed an equation based on only one site of the body. In practice, this is a very easy and fast method to predict muscle mass, compared to the equations requiring nine body sites to be measured. Based on these requirements, both the equations of Abe et al., 2018 [10], Abe et al., 2015 [13] (only the first and third equation in Table 3), Paris et al., 2017 [16] (only the second equation in Table 3) and Barbosa-Silva et al., 2021 [26] are the most reliable and applicable equations. The equations of Abe et al., 2018 [10] can be used in an Asian population. We recommend to use the second equation, from Table 3, because to use this equation there is only need for two measurements: muscle thickness of the forearm lateral and height. The equations of Abe et al., 2015 [13] and Paris et al., 2017 [16] can be used in a Caucasian population. Therefore, in a Caucasian population we recommend to use the third equation from Abe et al., 2015 [13]. In this equation, only two measurements are needed: muscle thickness of the forearm lateral and height. This equation has been cross-validated in the same population and showed good results [13]. The equation of Barbosa-Silva et al., 2020 [26] can be used in a South American population. Therefore, six measurements are needed: muscle thickness of the tight anterior and upper arm anterior, circumference of the dominant arm and leg, length of the arm, and height.

4.5. Recommendations

The overview of ultrasound-derived equations provided in this systematic review may be helpful for researchers and clinicians to choose the most appropriate equation for their target population. For future research, we recommend developing ultrasound-based equations, compared to a gold standard, like MRI or CT. In addition, the measurement procedure with ultrasound should be accorded to an evidence-based standardization, which may also include the measurement of muscle quality. The addition, muscle quality could be an aspect of the diagnosis of sarcopenia [1]. Moreover, measuring muscle quality is also an additional opportunity for the clinician that uses ultrasound, since muscle quality cannot be measured with DXA or BIA [1]. When echo intensity of a muscle can also be included in the assessment, it will increase the accurate diagnosis of low muscle quantity and support the diagnosis of sarcopenia. Ideally, a combination of the muscle thickness and echo intensity of one particular muscle at the same measured site might be a good way to determine muscle quantity and quality simultaneously.

Most important is the development of an equation to predict muscle mass in clinical populations to support the diagnosis of sarcopenia and other malnutrition-related disorders.

4.6. Strengths and limitations

To the knowledge of the authors, all relevant research concerning muscle mass equations has been identified and appraised. Valid equations for a Caucasian, Asian, and South American population have been identified.

The study also has some limitations. First, the accuracy of the equations was reported in various ways. This increased the difficulty to compare the applicability of the equations. In addition, no consensus regarding the cut off values for the applicability of the equations has been established yet (e.g.: SEE, R2 etc.). Second, the equations were validated against the body composition reference method DXA, which is not a gold standard for measuring muscle mass [28]. And third, we were not able to conclude whether there is an equation applicable in non-healthy populations.

5. Conclusion

The results of this systematic review highlight that the ultrasound-derived equations described in the included articles are valid and applicable in a healthy population, based on the low risk of bias research on ultrasound-derived equations. The clinician should choose an equation that best matches the population the equation was developed for. Furthermore, the protocol used to create the equation must be followed correctly. For a Caucasian population we recommend to use the equation of Abe et al. [13], where two measurements are needed. While for an Asian population, we recommend to use the equation of Abe et al., 2018 [11], with also two measurements needed. And for the South American population, the use of the equation of Barbosa-Silva et al., 2021 [26] is the most appropriate. For this equation, six measurements are needed. There is need for more research in clinical populations of different ethnicities, in which the equations are developed based on the results of gold standards for muscle mass, like MRI or CT.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contribution

CRediT roles:

- Jona Van den Broeck: Data curation, Formal analysis, Writing original draft
- Luca Buzzatti: Formal analysis, Review and editing
- Harriët Jager-Wittenaar: Review and editing
- Stany Perkisas: Review and editing
- Aldo Scafoglieri: Conceptualization, Methodology, Supervision, Review and editing

Declaration of competing interest

Jona Van den Broeck, Luca Buzzatti, Harriët Jager-Wittenaar, Stany Perkisas, and Aldo Scafoglieri declare that they have no conflict of interest.

References

[1] Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019;48(4):601.

[2] Filippin LI, Teixeira VN, da Silva MP, Miraglia F, da Silva FS. Sarcopenia: a predictor of mortality and the need for early diagnosis and intervention. Aging Clin Exp Res 2015;27(3):249–54.

[3] Janssen I, Shepard DS, Katzmarzyk PT, Roubenoff R. The healthcare costs of sarcopenia in the United States. J Am Geriatr Soc 2004;52(1):80–5.

[4] Moisey LL, Mourtzakis M, Cotton BA, Premji T, Heyland DK, Wade CE, et al. Skeletal muscle predicts ventilator-free days, ICU-free days, and mortality in elderly ICU patients. Crit Care 2013;17(5):R206.

[5] Rizzoli R, Reginster JY, Arnal JF, Bautmans I, Beaudart C, Bischoff-Ferrari H, et al. Quality of life in sarcopenia and frailty. Calcif Tissue Int 2013;93(2):101–20.

[6] Nijholt W, Scafoglieri A, Jager-Wittenaar H, Hobbelen JSM, van der Schans CP. The reliability and validity of ultrasound to quantify muscles in older adults: a systematic review. J Cachexia Sarcopenia Muscle 2017;8(5):702–12.

[7] Pretorius A, Keating JL. Validity of real time ultrasound for measuring skeletal muscle size. Phys Ther Rev 2008;13(6):415–26.

[8] Sanada K, Kearns CF, Midorikawa T, Abe T. Prediction and validation of total and regional skeletal muscle mass by ultrasound in Japanese adults. Eur J Appl Physiol 2006;96(1):24–31.

[9] Beaudart C, Bruyère O, Geerinck A, Hajaoui M, Scafoglieri A, Perkisas S, et al. Equation models developed with bioelectric impedance analysis tools to assess muscle mass: a systematic review. e-SPEN, Clin Nutr 2020;35:47–62.

[10] Abe T, Thiebaud RS, Loenneke JP, Fujita E, Akamine T. DXA-rectified appendicular lean mass: development of ultrasound prediction models in older adults. J Nutr Health Aging 2018;22(9):1080–5.

[11] Abe T, Loenneke JP, Thiebaud RS, Fujita E, Akamine T, Loftin M. Prediction and validation of DXA-derived appendicular fat-free adipose tissue by a single ultrasound image of the forearm in Japanese older adults. J Ultrasound Med 2018;37(2):347–53.

[12] Takai Y, Ohta M, Akagi R, Kato E, Wakahara T, Kawakami Y, et al. Applicability of ultrasound muscle thickness measurements for predicting fat-free mass in elderly population. J Nutr Health Aging 2014;18(6):579–85.

[13] Abe T, Thiebaud RS, Loenneke JP, Young KC. Prediction and validation of DXA-derived appendicular lean soft tissue mass by ultrasound in older adults. Age 2015;37(6):114.

[14] Cruz-Jentoft AJ, Baeyens JP, Bauer JM, Boirie Y, Cederholm T, Landi F, et al. Sarcopenia: European consensus on definition and diagnosis: report of the European working group on sarcopenia in older people. Age Ageing 2010;39(4):412–23.

[15] Kyle UG, Bosaeus I, De Lorenzo AD, Deurenberg P, Elia M, Gómez JM, et al. Bioelectrical impedance analysis–part I: review of principles and methods. Clin Nutr 2004;23(5):1226–43.

[16] Paris MT, Lafleur B, Dubin JA, Mourtzakis M. Development of a bedside viable ultrasound protocol to quantify appendicular lean tissue mass. J Cachexia Sarcopenia Muscle 2017;8(5):713–26.

[17] Perkisas S, Bastijns S, Sanchez-Rodriguez D, Piotrowicz K, De Cock AM. Application of ultrasound for muscle assessment in sarcopenia: 2020 SARCUS update: reply to the letter to the editor : SARCUS working group on behalf of the Sarcopenia Special Interest Group of the European Geriatric Medicine Society. Eur Geriatr Med 2021;12(2):427–8.

[18] Mendis MD, Wilson SJ, Stanton W, Hides JA. Validity of real-time ultrasound imaging to measure anterior hip muscle size: a comparison with magnetic resonance imaging. J Orthop Sports Phys Ther 2010;40(9):577–81.

[19] Reeves ND, Maganaris CN, Narici MV. Ultrasonographic assessment of human skeletal muscle size. Eur J Appl Physiol 2004;91(1):116–8.

[20] https://www.rayyan.ai/.

[21] Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155(8):529–36.

[22] Abe T, Dankel SJ, Buckner SL, Jessee MB, Mattocks KT, Mouser JG, et al. Magnetic resonance imaging-measured skeletal muscle mass to fat-free mass ratio increases with increasing levels of fat-free mass. J Sports Med Phys Fit 2019;59(4):619–23.

[23] Abe T, Fujita E, Thiebaud RS, Loenneke JP, Akamine T. Ultrasound-derived forearm muscle thickness is a powerful predictor for estimating DXA-derived appendicular lean mass in Japanese older adults. Ultrasound Med Biol 2016;42(9):2341–4.

[24] Abe T, Loenneke JP, Young KC, Thiebaud RS, Nahar VK, Hollaway KM, et al. Validity of ultrasound prediction equations for total and regional muscularity in middle-aged and older men and women. Ultrasound Med Biol 2015;41(2):557–64.

[25] Abe T, Patterson KM, Stover CD, Young KC. Influence of adipose tissue mass on DXA-derived lean soft tissue mass in middle-aged and older women. Age 2015;37(1):9741.

[26] Barbosa-Silva TG, Gonzalez MC, Bielemann RM, Santos LP, Costa CDS, Menezes AMB. $2 + 2 (+ 2) = 4$: a new approach for appendicular muscle mass assessment by ultrasound. Nutrition 2021;83:111056.

[27] Toda Y, Kimura T, Taki C, Kurihara T, Homma T, Hamaoka T, et al. New ultrasonography-based method for predicting total skeletal muscle mass in male athletes. J Phys Ther Sci 2016;28(5):1556–9.

[28] Scafoglieri A, Clarys JP. Dual energy X-ray absorptiometry: gold standard for muscle mass? J Cachexia Sarcopenia Muscle 2018;9(4):786–7.

[29] Thoirs K, English C. Ultrasound measures of muscle thickness: intra-examiner reliability and influence of body position. Clin Physiol Funct Imag 2009;29(6):440–6.

[30] Muraoka H, Suzuki T. Effects of trunk anterior tilt and knee joint flexion angle changes on muscle activity in the lower limb muscles. J Phys Ther Sci 2021;33(6):472–9.

[31] Fischer A, Hahn R, Anwar M, Hertwig A, Pesta M, Timmermann I, et al. How reliably can ultrasound help determine muscle and adipose tissue thickness in clinical settings? An assessment of intra- and inter-examiner reliability in the USVALID study. Eur J Clin Nutr 2021. https://pubmed.ncbi.nlm.nih.gov/34226675/. online ahead of print.

[32] Pardo E, El Behi H, Boizeau P, Verdonk F, Alberti C, Lescot T. Reliability of ultrasound measurements of quadriceps muscle thickness in critically ill patients. BMC Anesthesiol 2018;18(1):205.

[33] Cho KH, Lee HJ, Lee WH. Intra- and inter-rater reliabilities of measurement of ultrasound imaging for muscle thickness and pennation angle of tibialis anterior muscle in stroke patients. Top Stroke Rehabil 2017;24(5):368–73.

[34] Ishida Y, Carroll JF, Pollock ML, Graves JE, Leggett SH. Reliability of B-mode ultrasound for the measurement of body fat and muscle thickness. Am J Hum Biol 1992;4(4):511–20.

[35] Delmonico MJ, Harris TB, Lee JS, Visser M, Nevitt M, Kritchevsky SB, et al. Alternative definitions of sarcopenia, lower extremity performance, and functional impairment with aging in older men and women. J Am Geriatr Soc 2007;55(5):769–74.

[36] Abe T, Thiebaud RS, Loenneke JP, Loftin M, Fukunaga T. Prevalence of site-specific thigh sarcopenia in Japanese men and women. Age 2014;36(1):417–26.

[37] Rantanen T. Muscle strength, disability and mortality. Scand J Med Sci Sports 2003;13(1):3–8.

[38] Sowers MR, Crutchfield M, Richards K, Wilkin MK, Furniss A, Jannausch M, et al. Sarcopenia is related to physical functioning and leg strength in middle-aged women. J Gerontol Biol Med Sci 2005;60(4):486–90.

[39] Heymsfield SB, Gallagher D, Kotler DP, Wang Z, Allison DB, Heshka S. Body-size dependence of resting energy expenditure can be attributed to non-energetic homogeneity of fat-free mass. Am J Physiol Endocrinol Metab 2002;282(1):E132–8.

[40] Zhao X, Wang Z, Zhang J, Hua J, He W, Zhu S. Estimation of total body skeletal muscle mass in Chinese adults: prediction model by dual-energy X-ray absorptiometry. PloS One 2013;8(1):e53561.

[41] Scafoglieri A, Clarys JP, Cattrysse E, Bautmans I. Use of anthropometry for the prediction of regional body tissue distribution in adults: benefits and limitations in clinical practice. Aging Dis 2014;5(6):373–93.

[42] Janssen I, Ross R. Effects of sex on the change in visceral, subcutaneous adipose tissue and skeletal muscle in response to weight loss. Int J Obes Relat Metab Disord 1999;23(10):1035–46.