Validity of ultrasonography derived predictions for estimating skeletal muscle volume- A systematic literature review

Rasmus Liegnell (rasmus.liegnell@gmail.com)
Karolinska Institutet https://orcid.org/0000-0002-7773-1501
Fredrik Wessman
Karolinska Institutet
Adel Shalabi
Akademiska sjukhuset
Marita Harringe
Karolinska Institutet

Research article

Keywords: Bland-Altman analysis, magnetic resonance imaging, muscle thickness, muscle volume, prediction equation, ultrasonography.

Posted Date: September 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-65429/v1

Validity of ultrasonography derived predictions for estimating skeletal muscle volume - A systematic literature review

Rasmus Liegnell[1], Fredrik Wessman[1], Adel Shalabi[2] and Marita Harringe[1]

[1]Stockholm Sports Trauma Research Center, Dept of Molecular Medicine and Surgery Karolinska Institutet, Stockholm, Sweden. [2]Centre for Medical Imaging, Akademiska University Hospital, Uppsala University, Uppsala, Sweden.

Corresponding author:

Rasmus Liegnell

E-mail: rasmus.liegnell@gmail.com

ABSTRACT

Background: The amount of muscle mass or muscle volume (MV) varies between individuals and is important for health, wellbeing, and performance. Imaging is a useful tool to monitor MV, magnetic resonance imaging (MRI) is considered gold standard. MRI are not always easily accessible, and the measurements are expensive, therefore ultrasonography (US) has become a more accessible method for estimating MV.

Methods: A systematic literature review was conducted in the electronic databases PubMed, CINHAL and Web of Science with the purpose of collecting the current published equations to estimate MV with US and answering the following question: How well does US derived equations based on muscle thickness (MT) predict MV based on MRI?

Results: The literature search resulted in 363 citations. Twelve articles met the eligibility criteria and were included. Ten articles scored eight out of eleven on the QUADAS score and two scored nine out of eleven. 36 different prediction equations were identified. Correlations were good, r values ranged between 0.53-0.961 and the standard error of estimates (SEE) ranged between 6-25.6%. Eight studies did further analysis with a Bland-Altman plot and found no systematic errors. The overall strength and quality of the evidence was rated as “low quality” as defined by the GRADE system.

Conclusions: We conclude that the validity of US derived equations based on MT is specific to the populations from where it is developed. The agreement with MV based on MRI is moderate with SEE ranging between 6-12% in healthy populations. Suggestions for future research are to investigate if testing positions or increasing the number of measuring points could improve the validity for prediction equations.

KEY WORDS

Bland-Altman analysis, magnetic resonance imaging, muscle thickness, muscle volume, prediction equation, ultrasonography.

BACKGROUND

The skeletal muscle accounts for about 40% of the total bodyweight. Its primary function is to generate force and create physical movement essential for everyday living, health and performance (1). The skeletal muscle is also an endocrine organ, secreting a collection of factors called myokines that seems to have positive health effects on a variety of organs throughout the body (2). The amount of muscle mass or muscle volume (MV) varies between individuals and is influenced by a complex interaction between nutrition, physical load, hormones, age, injuries and diseases (1). MV gradually declines with age, which eventually may lead to Sarcopenia, affecting 10-50% of individuals above 65 years of age (3). Sarcopenia is associated with an increased risk of being hospitalised and all-cause mortality (4, 5). On the other hand, an increased or high amount of total MV seems to be protective and reduce the likelihood of common diseases and disabilities like cardiovascular disease, diabetes and immobility (6-8).

MV is strongly correlated with the ability to produce force and therefore, a good predictor for strength, thru the ability to create joint torque (force x moment arm) (9, 10). It is common with a decreased MV and reduced strength after an injury, surgery, or immobilization. Meier et al. 2009 reported that after knee arthroplasty an inability to activate quadriceps was contributing more to the loss in strength the first months and that quadriceps MV was a strong predictor of strength after more than 1 year (11). Quadriceps MV is also predictive of patient reported function and persistent strength deficit after ACL reconstruction (12).

A focus to improve strength in performance, rehabilitation or activities of daily living is often hypertrophy. Growth of contractile proteins within the skeletal muscle is a main outcome after repeated sessions of loading, through exercise or heavy daily activities, leading to hypertrophy and an increase in MV (13). The skeletal muscle is a plastic tissue that constantly adapts to the exposure and requirements in life and valid measures of MV and the changes in mass over time is therefore of great interest. To ensure that the intervention cause hypertrophy and muscle growth, valid methods of measures is needed.

Direct measure of the changes in protein synthesis is possible but requires muscle biopsies and expensive tracers (14, 15). Measurement of MV is achievable with high validity via the water displacement method (16) but it requires that the muscle be removed from its owner making it impossible to measure living beings and changes between different occasions.

Imaging is a useful tool to reduce suffering and enable non-invasive measurements of MV. Magnetic resonance imaging (MRI) or computed tomography (CT) are considered gold standard (17). MRI is preferable since CT involves radiation. The method for estimating MV measured with MRI (MV_{MRI}) is determined by measuring muscles single axial anatomical cross-sectional area (ACSA), in multiple sections along the entire length of the muscle and then multiplying ACSA with the length of each section (18). MRI are not always easily accessible, and the measurements are expensive, therefore ultrasonography (US) has become a widely used method to measure changes in muscle thickness (MT). Several studies measure the acute and long-term difference in MT with US, before and after a period of exercise (19-21). MT dimensions are measured as the distance from the subcutaneous adipose tissue muscle interface to the muscle bone interface (22). MT is well correlated to the MRI cross-sectional area (CSA) in both the lower (23) and upper extremity (24).

The estimation of MV with US (MV_{US}) is commonly based on measurement of MT and achieved by developing prediction equations through multiple regression analysis including limb length or other anthropometric variables (9, 25). The true value of MV is unknown but since MRI is considered gold standard, preferably the results from MV_{US} and the results from MV_{MRI} would be the same. When comparing MV_{US} to the water displacement method a standard errors of estimates (SEE) between 10-13% have been reported (16), similar SEE percentage are reported when MV_{MRI} and MV_{US} are compared (25). Even though SEE varies, the correlation in a population should be good, since both methods aim to measure the same thing (26). What´s interesting is if the more accessible US can estimate MV in a satisfying manner?

The aim of this study was to perform a systematic literature review with the purpose of collecting the current published equations to estimate MV_{US} and answer the following question: How well does US derived equations based on muscle thickness predict MV_{MRI}?

METHODS

Search strategy

A systematic search took place on the 30th of January 2020 in the electronic databases PubMed, CINHAL and Web of Science. MeSH terms were identified and used when possible. MeSH terms “ultrasonography” and “magnetic resonance imaging” were used as a concept and combined with AND Boolean operator. Search terms “muscle thickness” and “muscle volume” were used as a concept and combined with OR Boolean operator. Both concepts were combined with AND Boolean operator. Investigators (RL and FW) screened the titles of all identified articles and if eligible the abstracts were read and discussed. Unless both investigators agreed that the study did not meet the eligibility criteria the study was included

for full text review. Reference lists of the included studies were screened for eligible literature.

Eligibility criteria

To be included, the studies needed to meet the following criteria: 1: Measure muscle thickness with a B-mode ultrasound. 2: Use ultrasound derived equations based on muscle thickness to predict MV. 3: Use magnetic resonance imaging as the reference method for MV. 4: Published in the English language. Criteria for exclusion were the following: 1: Published before the year of 2000. 2: Animal studies. 3: Cadaver studies. 4: Reviews.

Quality assessment

To assess the quality of the included studies a translated version of QUADAS (27) published by the *Swedish Agency for Health Technology Assessment and Assessment of Social Services* was used. Both investigators did first assess each study independently. Then met to review and discuss each study until consensus was reached. GRADE was used to assess the overall strength and quality of the evidence (28).

Ethical considerations

All included studies declared that written or informed consent was given from study participants. In three studies the participants were children or adolescents below the age of 18 years, these studies did also obtain consent from their parents (29-31). Most included studies declared that they had approval from an independent ethical committee with the exception for two studies (32, 33) were there was no such declaration.

Statistical analysis

Two Bland-Altman plots were created from the mean values identified in the included studies, with the purpose to examine the agreement between the two methods in a descriptive manner. The values reported in cm^3 and kg were separated in different plots. Both plots were plotted against the mean value of MV_{MRI} and MV_{US} for every segment, the BIAS, standard deviation, upper and lower limits of agreement were calculated and reported as a percentage, according to the method described by Bland and Altman (26).

RESULTS

The literature search resulted in 299 citations in the PubMed database, 23 in Cinahl and 41 in Web of Science. After abstracts had been analysed and discussed 21 articles were selected for full text review. In the end 12 articles met the eligibility criteria and were included in the systematic literature review (Figure 1). Ten articles scored eight out of eleven on the QUADAS score and two scored nine out of eleven (Table 1). All articles lacked the same items on the QUADAS score, stated that it was unclear if the ones analysing the index test were blinded to the results of the reference test, and vice-versa.

All together the studies included 591 subjects. Five studies included only men (9, 32-35). Four studies included both men and women (25, 36-38). Two studies included prepubertal children (29, 30). One study also included adolescents (29) and one study included children with cerebral palsy (31). Descriptive data is presented in Table 1.

A total of 12 different body parts or muscle groups were measured, and 36 different prediction equations were identified. Correlations between MV_{US} and MV_{MRI} were good, r values ranged between 0.53-0.961 and the SEE ranged between 6-25.6%. Regressions and measured segments are presented in Table 2. Eight studies did further analysis with a Bland-Altman plot (25, 29, 30, 32, 33, 35, 36, 38) and they found no systematic errors.

A total of 13 segment reported in cm^3 from five studies (Table 1) (25, 30, 32, 34, 38) were included in the first plot and plotted against the average (Figure 2 A). Showing an even spread in percentage when differences between methods were plotted against the average mean. One measure crossed the lower limit of agreement, which was the anterior upper arm data reported from Miyatani et al. 2000. Three studies reported values in kg (Table 1) (29, 35, 36) although when total body estimates were excluded from the Bland-Altman, two studies remained (29, 36). Midorikawa et al. 2009 tested the equation derived from Sanada et al. 2006 and a total of eleven segments were plotted against the average (Figure 2 B). In this plot, the data show a bigger spread, illustrated by the Y-axis in the plots in Figure 2. Mainly two data points are the reason for this, the segment arm (-44 %) and lower leg (-13 %) calculated from measures on prepubertal children reported by Midorikawa et al. 2009. The segment arm in prepubertal children crossed the lower limit of agreement.

The overall strength and quality of the evidence was rated as “low quality” as defined by the GRADE system.

DISCUSSION

The most important finding of the present investigation was that the validity of US derived equations based on MT is specific to the populations from where it is developed. This systematic literature review included twelve studies that investigated the validity of US derived equations based on MT to predict MV_{MRI}. Previous reviews have looked at the association between MT and MV for the upper extremity (39) and the lower extremity (40). These reviews also included other reference methods like CT and cadavers, although had some similar studies included as our review (25, 32, 33, 37). Nijholt et al. 2017 conducted a systematic review investigating the validity of US derived prediction equations to estimate MV (41). However, they looked at solely elderly populations aged >60 years and included

only two articles that used DEXA as a reference. To our knowledge, no other systematic review has investigated the validity of MV_{US} with MV_{MRI} as a reference.

The quality assessment showed that all included studies were of high quality and that they had similar scores. Our eligibility criteria were narrow and therefore all included studies had more or less the same design, which can be regarded as a strength since it makes it easier to comprehend the results. Unfortunately, this is also a weakness since 11 out 12 studies were conducted in the same country and many of those studies came from the same research group. This fact affected the strength of evidence synthesis according to GRADE along with a lack of some descriptive data that were not published. Most commonly we lacked means for MV (9, 31, 33, 37) and there was no individual data published in any of the studies.

The results of the present review showed good correlations generally across the included studies which are in line with previous reviews by Abe et al. (39, 40). The SEE varied between 6-25.6% (Table 2) which can seem like quite large variations across the studies. It should be noted that one study conducted by Park et al. 2014 on children with cerebral palsy deviated from the rest and explained the high range of SEE (31). This was the only study among the included that had subjects with a diagnosis, they reported SEE of 20.6% for the medial gastrocnemius and 25.6% of the lateral gastrocnemius. If Park et al. 2014 would have been excluded, the range of SEE would be 6-12% and thus, less variation across the included studies.

Eight studies did further analysis by using a Bland-Altman plot and all of those plotted against the average (25, 29, 30, 32, 33, 35, 36, 38). Whether to plot against the average or against the reference is debatable (43, 44). If MRI is considered gold standard and the purpose is to develop another method to reach agreement with MRI, plotting against the reference seems to be more appropriate. All of the included studies used manual slice-by-slice segmentation technique to measure MV_{MRI}, however it should be noticed that, despite being

widely adopted as gold standard, this method has only been validated against the water displacement method in one study according to a recent systematic review (17). With an unknown true value for MV, plotting against the average mean is most likely correct. Bland and Altman discuss that plotting difference against standard method might be misleading and to plot against the average is more useful in almost all applications to medical measurements (43).

The prediction equations seem to be specific for the populations it´s developed for. Midorikawa et al. 2009 tested the validity of MV_{US} for adolescents and prepubertal children based on equations previously derived from adults and found inferior validity for prepubertal children (29). There was no significant difference for adolescents, however their Bland-Altman analysis showed a relatively high level of variability for both adolescents and prepubertal children. Nakatani et al. 2016 found that prediction equations developed for young adults were not valid for middle-aged and older men and women (38). Toda et al. 2016 investigated if prediction equations derived from a sedentary population was applicable for young male athletes, reporting it not to be valid in this population (35). It is important to recognize which population the prediction equations have been developed from, if the equations are to be used in the clinic. Our review includes prediction equations developed from a wide range of age and both men and women, but it should be noted that all studies except one (31) are conducted in Japan.

Prediction equations being specific for the population, is illustrated in our Bland-Altman analysis (Figure 2). With the wide range in cm^3 and kg, our Bland-Altman analysis was conducted in percentage. Figure 2B illustrates data from only two studies, where one Midorikawa et al. 2009 tested the equation derived from Sanada et al. 2006 on different populations. The strength of Figure 2B is consequently that the same equation was used, however the downside is that the equation wasn´t derived for the prepubertal children and

adolescents, resulting in a larger BIAS (-6 %) compared with the data in Figure 2A (-1 %). Figure 2A is the exact opposite from 2B where different equations are mixed, but they are all derived for a specific population, resulting in a better outcome. Since the data has a large spread, scattered between small and bigger segments, a Bland-Altman plot with absolute values would have been misleading, favouring the smaller segments and thus the method was chosen to plot differences as percentage (42). The reason for excluding the total body data were due to the large values, since the total body means would have displaced values on the X-axis, the total body data would have been unrepresentative for the segment data.

Many different prediction equations were identified, and the studies used different variables in their regression analysis in addition to MT (Table 2). Miyatani et al. 2000 did the first prediction equations with the formula for calculating a cylinder, including limb length as a variable (34). The same group later reported that the prediction improved when MT is combined with limb length compared to MT alone (33). Different variables have then been tested including sex, limb length, body height, circumference, and age. Still the SEE does not vary a lot as seen in Table 2. Akagi et al. 2010 found that a decrease in MV did not correspond to a decrease in MT with ageing (25). Highlighting the complexity of developing a highly accurate prediction equation based on MT and the need for additional variables.

Factors that contribute to this complexity is that the measurement of MT with US does not differentiate between contractile and noncontractile tissue and that the changes in MV does not only depend on MT, but also muscle width (25). Muscles also varies in shape and are not cylindrical which add to the complexity to predict MV_{US} based on MT. All included studies did in some way include muscle length and MT in their regressions. MT is measured at only one site for each segment. Potentially it would be interesting to study if more measured sites of MT along the muscle could improve the validity of the prediction. For

example, if MT at 30%, 50% and 70% along the quadriceps were included in a multiple regression. Ogawa et al. 2012 did measure these spots at the medial anterior aspect of the thigh but they only made simple regressions separately for each spot (37).

When conducting an MRI scan, the subject is commonly placed in a supine position, even though there are MRI scans that can scan subjects in an upright position (45). Not surprisingly almost all studies had their subjects in supine position when measuring MV_{MRI} (Table 1). More noticeable is that almost all studies had their subjects in standing position when measuring MT (Table 1). We do not know the reason for this. It is also unclear if this has any significance for the validity of MV_{US}. One could speculate that the muscle changes slightly in shape in different positions and that US derived MT measured in the same position as the reference method would make the predictions better, thereby increasing the validity of MV_{US}.

In the field of sport medicine, quadriceps MV partly explains persistent weakness after ACL injury (46) and atrophy of the quadriceps muscles negatively impacts knee extension strength (12). A valid measure of MV could be applied to monitor progress after ACL reconstruction and assist in return-to-play decisions by giving clinicians a quick and simple prediction of the athletes MV. The present study has listed all the segments, genders and derived equations and compiled them into Table 2, helping clinicians with a user-friendly reference card to estimate MV with the help of US.

CONCLUSIONS

In summary this systematic review identified twelve studies of high quality assessed with QUADAS that investigated the validity of ultrasound derived equations based on MT to predict MV_{MRI}. The studies were homogeneous in design and almost all the included studies were conducted in the same country. We conclude that the validity of US derived equations

based on MT is specific to the populations from where it is developed. The agreement with MV_{MRI} is moderate with SEE ranging between 6-12% in healthy populations. We have designed a user-friendly reference card for clinicians in Table 2. The strength of the synthesized evidence is rated as low quality according to GRADE. Suggestions for future research are to investigate if testing positions or increasing the number of measuring points with MT could improve the prediction equations.

LIST OF ABBREVIATIONS

Axial anatomical cross-sectional area (ACSA)

Computed tomography (CT)

Magnetic resonance imaging (MRI)

Muscle thickness (MT)

Muscle volume (MV)

The method for estimating MV measured with MRI based on ACSA (MV_{MRI})

Standard errors of estimates (SEE)

Ultrasonography (US)

The method for estimating MV measured with US based om MT (MV_{US})

DECLARATIONS

Ethics approval and consent to participate

All studies included in the present investigation declared that written or informed consent was given from study participants.

Consent for publication

Not applicable

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

RL and FW designed the study, conducted the literature review, analysed the data, and drafted the manuscript. RL made the figures. AS participated in drafting the manuscript. MH participated in designing the study and drafting the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable

REFERENCES

1. Frontera WR, Ochala J. Skeletal muscle: a brief review of structure and function. Calcif Tissue Int. 2015;96(3):183-95.

2. Giudice J, Taylor JM. Muscle as a paracrine and endocrine organ. Curr Opin Pharmacol. 2017;34:49-55.

3. Sergi G, Trevisan C, Veronese N, Lucato P, Manzato E. Imaging of sarcopenia. Eur J Radiol. 2016;85(8):1519-24.

4. Liu P, Hao Q, Hai S, Wang H, Cao L, Dong B. Sarcopenia as a predictor of all-cause mortality among community-dwelling older people: A systematic review and meta-analysis. Maturitas. 2017;103:16-22.

5. Zhang X, Zhang W, Wang C, Tao W, Dou Q, Yang Y. Sarcopenia as a predictor of hospitalization among older people: a systematic review and meta-analysis. BMC Geriatr. 2018;18(1):188.

6. Westcott WL. Resistance training is medicine: effects of strength training on health. Curr Sports Med Rep. 2012;11(4):209-16.

7. McLeod JC, Stokes T, Phillips SM. Resistance Exercise Training as a Primary Countermeasure to Age-Related Chronic Disease. Front Physiol. 2019;10:645.

8. Tyrovolas S, Panagiotakos D, Georgousopoulou E, Chrysohoou C, Tousoulis D, Haro JM, et al. Skeletal muscle mass in relation to 10 year cardiovascular disease incidence among middle aged and older adults: the ATTICA study. J Epidemiol Community Health. 2020;74(1):26-31.

9. Fukunaga T, Miyatani M, Tachi M, Kouzaki M, Kawakami Y, Kanehisa H. Muscle volume is a major determinant of joint torque in humans. Acta Physiol Scand. 2001;172(4):249-55.

10. Baxter JR, Piazza SJ. Plantar flexor moment arm and muscle volume predict torque-generating capacity in young men. J Appl Physiol (1985). 2014;116(5):538-44.

11. Meier WA, Marcus RL, Dibble LE, Foreman KB, Peters CL, Mizner RL, et al. The long-term contribution of muscle activation and muscle size to quadriceps weakness following total knee arthroplasty. J Geriatr Phys Ther. 2009;32(2):79-82.

12. Kuenze CM, Blemker SS, Hart JM. Quadriceps function relates to muscle size following ACL reconstruction. J Orthop Res. 2016;34(9):1656-62.

13. Goodman CA, Frey JW, Mabrey DM, Jacobs BL, Lincoln HC, You JS, et al. The role of skeletal muscle mTOR in the regulation of mechanical load-induced growth. J Physiol. 2011;589(Pt 22):5485-501.

14. Foster DM, Barrett PH, Toffolo G, Beltz WF, Cobelli C. Estimating the fractional synthetic rate of plasma apolipoproteins and lipids from stable isotope data. J Lipid Res. 1993;34(12):2193-205.

15. Gasier HG, Riechman SE, Wiggs MP, Previs SF, Fluckey JD. A comparison of 2H2O and phenylalanine flooding dose to investigate muscle protein synthesis with acute exercise in rats. Am J Physiol Endocrinol Metab. 2009;297(1):E252-9.

16. Bandholm T, Sonne-Holm S, Thomsen C, Bencke J, Pedersen SA, Jensen BR. Calf muscle volume estimates: implications for botulinum toxin treatment? Pediatr Neurol. 2007;37(4):263-9.

17. Pons C, Borotikar B, Garetier M, Burdin V, Ben Salem D, Lempereur M, et al. Quantifying skeletal muscle volume and shape in humans using MRI: A systematic review of validity and reliability. PLoS One. 2018;13(11):e0207847.

18. Tracy BL, Ivey FM, Jeffrey Metter E, Fleg JL, Siegel EL, Hurley BF. A more efficient magnetic resonance imaging-based strategy for measuring quadriceps muscle volume. Med Sci Sports Exerc. 2003;35(3):425-33.

19. Dankel SJ, Buckner SL, Jessee MB, Mattocks KT, Mouser JG, Counts BR, et al. Post-exercise blood flow restriction attenuates muscle hypertrophy. Eur J Appl Physiol. 2016;116(10):1955-63.
20. Lowery RP, Joy JM, Loenneke JP, de Souza EO, Machado M, Dudeck JE, et al. Practical blood flow restriction training increases muscle hypertrophy during a periodized resistance training programme. Clin Physiol Funct Imaging. 2014;34(4):317-21.
21. Loenneke JP, Kim D, Fahs CA, Thiebaud RS, Abe T, Larson RD, et al. The influence of exercise load with and without different levels of blood flow restriction on acute changes in muscle thickness and lactate. Clin Physiol Funct Imaging. 2017;37(6):734-40.
22. Schoenfeld BJ, Grgic J, Contreras B, Delcastillo K, Alto A, Haun C, et al. To Flex or Rest: Does Adding No-Load Isometric Actions to the Inter-Set Rest Period in Resistance Training Enhance Muscular Adaptations? A Randomized-Controlled Trial. Front Physiol. 2019;10:1571.
23. Giles LS, Webster KE, McClelland JA, Cook J. Can ultrasound measurements of muscle thickness be used to measure the size of individual quadriceps muscles in people with patellofemoral pain? Phys Ther Sport. 2015;16(1):45-52.
24. Abe T, Nakatani M, Loenneke JP. Relationship between ultrasound muscle thickness and MRI-measured muscle cross-sectional area in the forearm: a pilot study. Clin Physiol Funct Imaging. 2018;38(4):652-5.
25. Akagi R, Takai Y, Kato E, Wakahara T, Ohta M, Kanehisa H, et al. Development of an equation to predict muscle volume of elbow flexors for men and women with a wide range of age. Eur J Appl Physiol. 2010;108(4):689-94.
26. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307-10.

421 27. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of
422 QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in
423 systematic reviews. BMC Med Res Methodol. 2003;3:25.
424 28. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE:
425 an emerging consensus on rating quality of evidence and strength of recommendations. Bmj.
426 2008;336(7650):924-6.
427 29. Midorikawa T, Sanada K, Yoshitomi A, Abe T. Is the use of ultrasound-derived prediction
428 equations for adults useful for estimating total and regional skeletal muscle mass in Japanese
429 children? Br J Nutr. 2009;101(1):72-8.
430 30. Midorikawa T, Ohta M, Hikihara Y, Torii S, Sakamoto S. Prediction and validation of total
431 and regional skeletal muscle volume using B-mode ultrasonography in Japanese prepubertal
432 children. Br J Nutr. 2015;114(8):1209-17.
433 31. Park ES, Sim E, Rha DW, Jung S. Estimation of gastrocnemius muscle volume using
434 ultrasonography in children with spastic cerebral palsy. Yonsei Med J. 2014;55(4):1115-22.
435 32. Miyatani M, Kanehisa H, Kuno S, Nishijima T, Fukunaga T. Validity of ultrasonograph
436 muscle thickness measurements for estimating muscle volume of knee extensors in humans.
437 Eur J Appl Physiol. 2002;86(3):203-8.
438 33. Miyatani M, Kanehisa H, Ito M, Kawakami Y, Fukunaga T. The accuracy of volume
439 estimates using ultrasound muscle thickness measurements in different muscle groups. Eur J
440 Appl Physiol. 2004;91(2-3):264-72.
441 34. Miyatani M, Kanehisa H, Fukunaga T. Validity of bioelectrical impedance and
442 ultrasonographic methods for estimating the muscle volume of the upper arm. Eur J Appl
443 Physiol. 2000;82(5-6):391-6.

444 35. Toda Y, Kimura T, Taki C, Kurihara T, Homma T, Hamaoka T, et al. New
445 ultrasonography-based method for predicting total skeletal muscle mass in male athletes. J
446 Phys Ther Sci. 2016;28(5):1556-9.
447 36. Sanada K, Kearns CF, Midorikawa T, Abe T. Prediction and validation of total and
448 regional skeletal muscle mass by ultrasound in Japanese adults. Eur J Appl Physiol.
449 2006;96(1):24-31.
450 37. Ogawa M, Mitsukawa N, Bemben MG, Abe T. Ultrasound assessment of adductor muscle
451 size using muscle thickness of the thigh. J Sport Rehabil. 2012;21(3):244-8.
452 38. Nakatani M, Takai Y, Akagi R, Wakahara T, Sugisaki N, Ohta M, et al. Validity of muscle
453 thickness-based prediction equation for quadriceps femoris volume in middle-aged and older
454 men and women. Eur J Appl Physiol. 2016;116(11-12):2125-33.
455 39. Abe T, Loenneke JP, Thiebaud RS, Loftin M. Morphological and functional relationships
456 with ultrasound measured muscle thickness of the upper extremity and trunk. Ultrasound.
457 2014;22(4):229-35.
458 40. Abe T, Loenneke JP, Thiebaud RS. Morphological and functional relationships with
459 ultrasound measured muscle thickness of the lower extremity: a brief review. Ultrasound.
460 2015;23(3):166-73.
461 41. Nijholt W, Scafoglieri A, Jager-Wittenaar H, Hobbelen JSM, van der Schans CP. The
462 reliability and validity of ultrasound to quantify muscles in older adults: a systematic review. J
463 Cachexia Sarcopenia Muscle. 2017;8(5):702-12.
464 42. Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb).
465 2015;25(2):141-51.
466 43. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference
467 against standard method is misleading. Lancet. 1995;346(8982):1085-7.

44. Krouwer JS. Why Bland-Altman plots should use X, not (Y+X)/2 when X is a reference method. Stat Med. 2008;27(5):778-80.

45. Botchu R, Bharath A, Davies AM, Butt S, James SL. Current concept in upright spinal MRI. Eur Spine J. 2018;27(5):987-93.

46. Birchmeier T, Lisee C, Kane K, Brazier B, Triplett A, Kuenze C. Quadriceps Muscle Size Following ACL Injury and Reconstruction: A Systematic Review. J Orthop Res. 2020;38(3):598-608.

47. Moher D, Liberati A, Tetzlaff J, Altman DG, Group atP. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. Annals of Internal Medicine. 2009;151(4):264-9.

TABLES

Table 1: Descriptive data

Subjects	Body part	Mean MV_{MRI} [±SEM] (±SD)	Mean MV_{UL} [±SEM] (±SD)	Position MRI	Position UL	QUADAS	Reference
26 ♂ (23-34y)	Anterior upper arm Posterior upper arm	297.5 [14.9] cm^3 405.4 [22.6] cm^3	273.6 [15.4] cm^3 387.9 [30.9] cm^3	Supine	Standing	8/11	(34)
26 ♂ ($\bar{X}$ 25.7y)	Anterior upper arm Posterior upper arm			Supine	Standing	8/11	(9)
46 ♂ (20-70y)	Anterior upper thigh	1637.5 (383.51) cm^3	1660.7 (386.2) cm^3	Supine	Standing	8/11	(32)
27 ♂ ($\bar{X}$ 25.3y)	Anterior upper arm Posterior upper arm Anterior upper thigh Posterior lower leg	311.5 (68.5) cm^3 427.0 (112.2) cm^3 1825.3 (428) cm^3 1072.7 (226.3) cm^3		Supine	Standing	8/11	(33)
72 ♂♀ (18-61y)	Total body Arm Trunk Thigh Lower leg	20.2 (6.5) kg 2.0 (0.7) kg 8.3 (2.9) kg 7.7 (2.6) kg 2.2 (0.6) kg	19.6 (6.5) kg 1.9 (0.7) kg 8.2 (2.8) kg 7.5 (2.2) kg 2.2 (0.6) kg	Supine	Standing	8/11	(36)
10 Prepubertal children ($\bar{X}$ 9.2y) ♂ ($\bar{X}$ 10.3y) ♀	Total body Arm Trunk Thigh Lower leg	9.9 (1.4) kg 0.8 (0.1) kg 4.3 (0.7) kg 3.6 (0.6) kg 1.1 (0.2) kg	9.5 (1.1) kg 0.7 (0.1) kg 4.7 (0.6) kg 3.3 (0.6) kg 0.7 (0.1) kg	Supine	Standing	8/11	(29)
21 Adolescents ($\bar{X}$ 14.1y) ♂ ($\bar{X}$ 13.8y) ♀	Total body Arm Trunk Thigh Lower leg	17.4 (3.8) kg 1.5 (0.4) kg 7.3 (1.6) kg 6.6 (1.5) kg 2.0 (0.4) kg	17.5 (4.3) kg 1.5 (0.4) kg 7.5 (1.7) kg 6.5 (1.8) kg 1.9 (0.5) kg				
147 ♂♀ (19-77y)	Anterior upper arm	182.2 (65.4) cm^3	179.4 (62) cm^3	Supine	Standing	8/11	(25)
20 ♂♀ (20-41y)	Inner upper thigh			Supine	Standing	9/11	(37)
9 children* (2-6y)	Posterior lower leg medial & lateral	19.70 (9.29) cm^3 11.92 (9.12) cm^3		Did not report	Prone	8/11	(31)
60 ♂ (6-12y)	Total body Arm Trunk Thigh Lower leg	9113 (2241) cm^3 851 (198) cm^3 3495 (795) cm^3 3579 (1026) cm^3 1164 (295) cm^3	8942 (2841) cm^3 825 (194) cm^3 3453 (780) cm^3 3484 (986) cm^3 1180 (323) cm^3	Supine	Standing	8/11	(30)
37 ♀ (6-12y)	Total body Arm Trunk Thigh Lower leg	7688 (2339) cm^3 743 (208) cm^3 2798 (519) cm^3 2905 (905) cm^3 1084 (344) cm^3	7804 (2461) cm^3 719 (232) cm^3 2982 (929) cm^3 3030 (1015) cm^3 1074 (346) cm^3				
60 ♂♀ (51-77y)	Anterior upper thigh	1000 (373.3) cm^3	1019.5 (370.9) cm^3	Supine	Standing	8/11	(38)
61 ♂ ($\bar{X}$ 20.4y)	Total body	38.5 (5.8) kg	38.5 (5.7) kg	Prone	Standing	9/11	(35)

Data in Table 1 is presented in order, from the earliest publication to the latest. *SEM* standard error of mean. *SD* standard deviation. $\bar{X}$ mean in sample. *y* years. ♂ boys and males. ♀ girls and females. *Nine children with bilateral involvement spastic CP (6 ♂ and 3 ♀) in total, 18 lower limbs were evaluated.

Table 2: Muscle thickness sites, equations, and correlations

Segments	Equations	r	SEE cm^3	Reference
EF and EE	$MV_{US} = L \times (\pi \times MT/2)^2$	0.962	50.7 (7.2%)	(34)
EF	$MV_{US} = 2.586\ BH - 1.259\ BW + 7.057\ CIR + 0.524\ (L \times (MT)^2) - 447.46$	0.943	6-8 %	(9)
EE	$MV_{US} = 3.478\ BH - 0.180\ BW + 6.674\ CIR + 0.382\ (L \times (MT)^2) - 559.36$	0.932		
KE	$MV_{US} = (MT \times 311.732) + (L \times 53.346) - 2058.529$	0.824*	175.6 (10.6%)	(32)
EF	$MV_{US} = (MT \times 117.9) + (L \times 12.6) - 494$	0.901*	21.6 (6.9%)	(33)
EE	$MV_{US} = (MT \times 98.1) + (L \times 31.9) - 984.4$	0.866*	41.1 (9.6%)	
KE	$MV_{US} = (MT \times 320.6) + (L \times 110.9) - 4437.9$	0.826*	178.5 (9.8%)	
APF30	$MV_{US} = (MT \times 219.9) + (L \times 31.3) - 1758$	0.833*	92.5 (8.6%)	
♂ Total body (sum of 9 MT)	$MV_{TOT} = 0.641 \times MT_9 \times BH - 12.087$	0.96	2.24 kg	(36)
♂ Total body (sum of 6 MT)	$MV_{TOT} = 0.809 \times MT_6 \times BH - 4.834$	0.96	1.8 kg	
♂ Arm (EF+EE+LAF)	$MV_{TOT} = 0.204 \times MT_{arm} \times BH - 0.517$	0.95	0.22 kg	
♂ Trunk (AB+SUS)	$MV_{TOT} = 1.303 \times MT_{trunk} \times BH + 1.766$	0.88	1.11 kg	
♂ Thigh (KE+KF)	$MV_{TOT} = 0.639 \times MT_{thigh} \times BH - 2.972$	0.83	1.76 kg	
♂ Lower leg (APF30+ADF)	$MV_{TOT} = 0.233 \times MT_{lower\ leg} \times BH - 1.347$	0.83	0.55 kg	
♀ Total body (sum of 9 MT)	$MV_{TOT} = 0.594 \times MT_9 \times BH - 11.32$	0.91	2.75 kg	
♀ Total body (sum of 6 MT)	$MV_{TOT} = 0.831 \times MT_6 \times BH - 7.992$	0.88	2.88 kg	
♀ Arm (EF+EE+LAF)	$MV_{TOT} = 0.132 \times MT_{arm} \times BH + 0.093$	0.53	0.47 kg	
♀ Trunk (AB+SUS)	$MV_{TOT} = 0.937 \times MT_{trunk} \times BH + 1.794$	0.61	1.27 kg	
♀ Thigh (KE+KF)	$MV_{TOT} = 0.532 \times MT_{thigh} \times BH - 2.638$	0.81	1.39 kg	
♀ Lower leg (APF30+ADF)	$MV_{TOT} = 0.237 \times MT_{lower\ leg} \times BH - 1.534$	0.77	0.61 kg	
EF	$MV_{US} = 60.8 \times MT + 6.48 \times L - 0.709 \times AGE + 51.4 \times SEX - 187.4$	0.909*	19.9 (10.9%)	(25)
ADD	$MV_{US} = 5.51 \times MT \times L - 434.9$	0.922		(37)
APF25 medial	$MV_{US} = 2.271 \times L + 15.982 \times MT - 41.493$	0.831*	4.1 (20.6%)	(31)
APF25 lateral	$MV_{US} = 1.479 \times L + 13.347 \times MT - 28.676$	0.779*	3.1 (25.6%)	
♂ Total body (sum of 9 MT)	$MV_{US} = 384.96 \times (MT_9 \times BH) - 3662.1$	0.93*	659	(30)
♂ Arm (EF+EE+LAF)	$MV_{US} = 127.09 \times (MT_{arm} \times BH) - 76.44$	0.71*	124	
♂ Trunk (AB+SUS)	$MV_{US} = 992.53 \times (MT_{trunk} \times BH) + 363.69$	0.65*	565	
♂ Thigh (KE+KF)	$MV_{US} = 463.47 \times (MT_{thigh} \times BH) - 1624.3$	0.84*	419	
♂ Lower leg (APF30+ADF)	$MV_{US} = 176.1 \times (MT_{lower\ leg} \times BH) - 539.29$	0.92*	91	
♀ Total body (sum of 9 MT)	$MV_{US} = 364.87 \times (MT_9 \times BH) - 3523$	0.89*	731	
♀ Arm (EF+EE+LAF)	$MV_{US} = 132.68 \times (MT_{arm} \times BH) - 139.4$	0.8*	89	
♀ Trunk (AB+SUS)	$MV_{US} = 658.79 \times (MT_{trunk} \times BH) + 953.72$	0.57*	561	
♀ Thigh (KE+KF)	$MV_{US} = 425.40 \times (MT_{thigh} \times BH) - 1506.7$	0.9*	286	
♀ Lower leg (APF30+ADF)	$MV_{US} = 166.19 \times (MT_{lower\ leg} \times BH) - 439.17$	0.88*	103	
KE	$MV_{US} = (SEX \times 267.7) + (MT \times 249.3) + (L \times 41.1) - 1663.7$	0.91*	124.4 (12%)	(38)
Total body (sum of 9 MT)	$MV_{TOT} = 0.645 \times (MT_9 \times BH) - 7.821$	0.96		(35)

Data in Table 2 is presented in order, from the earliest publication to the latest. **Equations:** MV_{US} (cm^3) muscle volume US. MV_{TOT} (kg) estimated muscle volume via US in kg. *MT* (cm) muscle thickness obtained via US. *L* (cm) length of the muscle. *CIR* (cm) circumference of upper arm at the same site for MT measuring. *BH* (m) body height. *BW* (kg) body weight. *SEX* refers to biological differences between males 1 and females 0. *AGE* values in years. ♂ boys and males. ♀ girls and females. **Segments:** *EF* elbow flexors MT obtained at 60 % of anterior arm. *EE* elbow extensors, 60 % posterior arm. *KE* knee extensors, midpoint of anterior thigh. *KF* knee flexors, midpoint of posterior thigh. *APF30* ankle plantar flexors, 30 % of the posterior lower leg. *APF25* ankle plantar flexors, 25 % of the posterior lower leg. *ADF* ankle dorsal flexors, 30 % of the anterior lower leg. *LAF* obtained at 30 % of the lateral anterior forearm. *AB* abdominal obtained at a distance 2–3 cm to the right of the umbilicus. *SUS* subscapular, 5 cm directly below the inferior angle of the scapula. *ADD* adductor, 30 % of the medial anterior aspect of the thigh. *r in square instead of r.

FIGURE LEGENDS

Figure 1: Flow chart of the literature search, based on work from The PRISMA Group (47).

Figure 2: Bland-Altman agreement, difference plotted in percentage for the studies that reported MV_{US}. Total body estimates excluded in the plots. Values in plot **A** for the studies that reported data in cm^3: Bias -1 %, SD 4 %, limits of agreement upper 7 % and lower -8 %. Values in plot **B** for the studies that reported data in kg: Bias -6 %, SD 13 %, limits of agreement upper 20 % and lower -32 %.

Figures

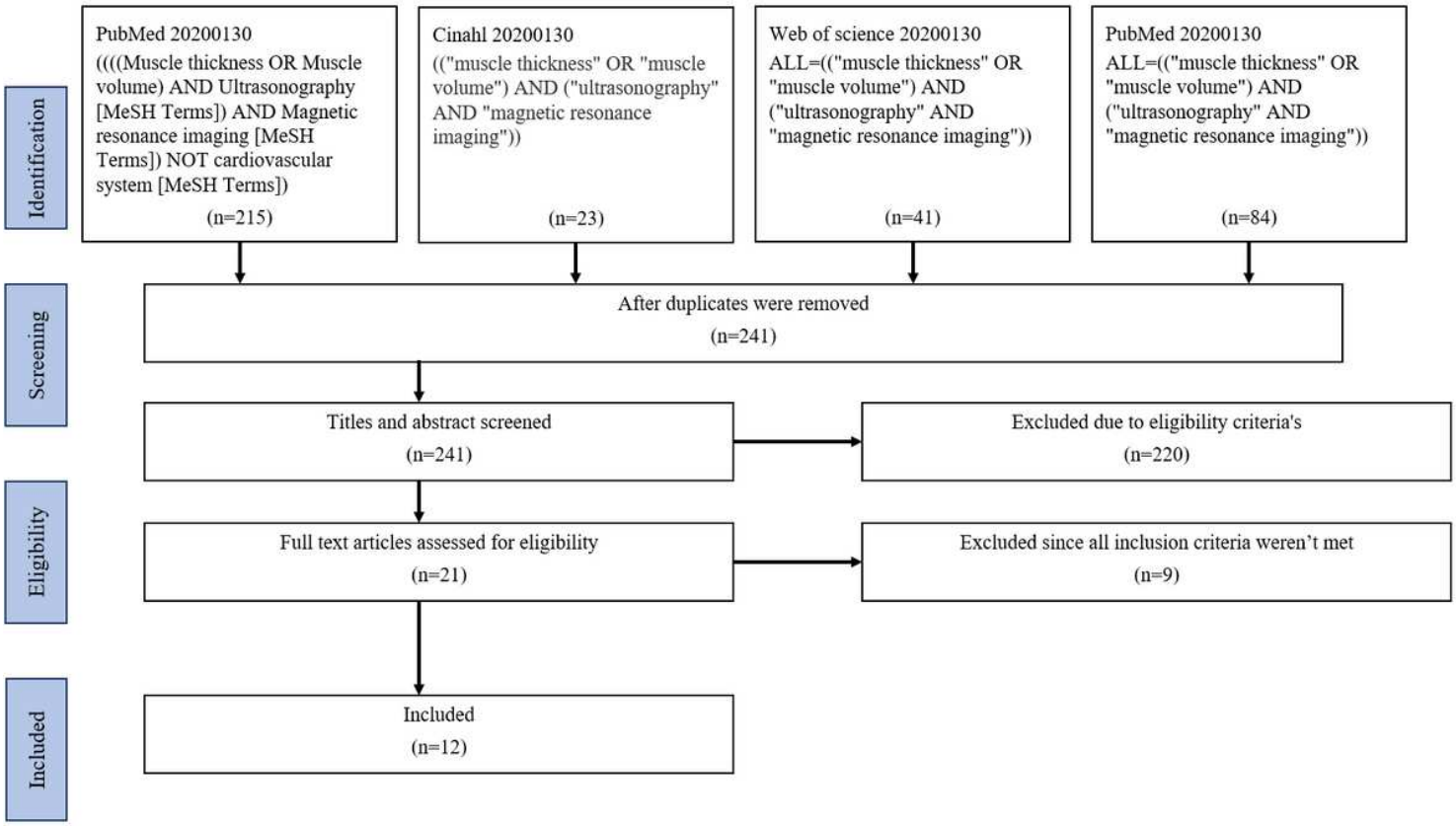

Figure 1

Flow chart of the literature search, based on work from The PRISMA Group (47).

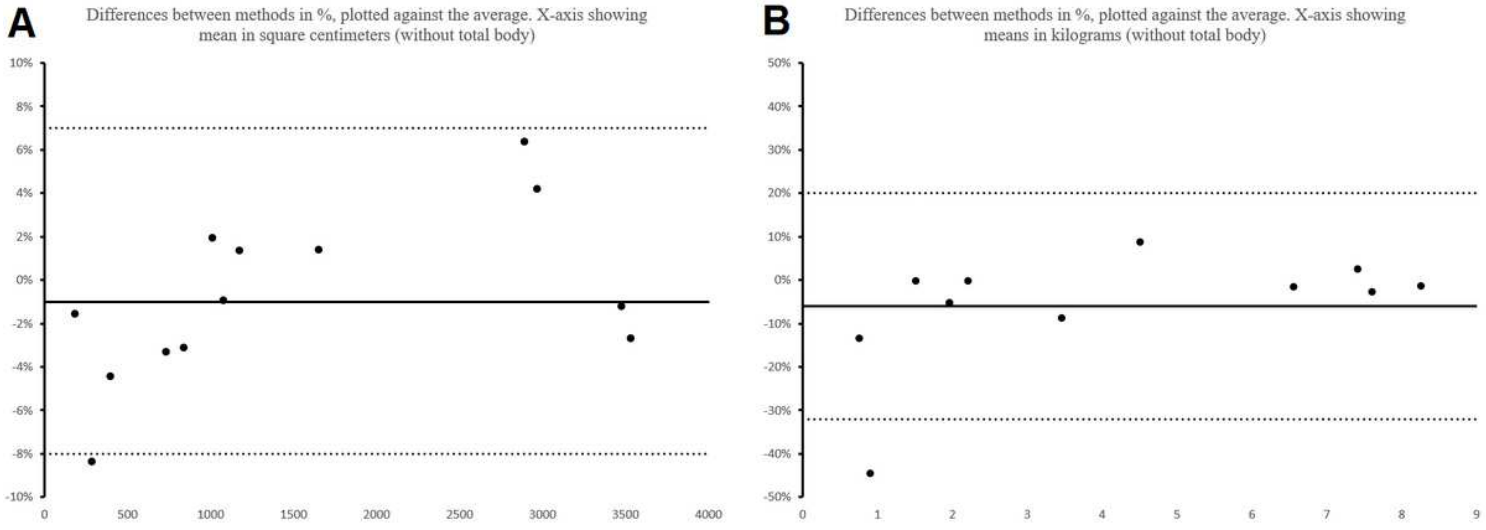

Figure 2

Bland-Altman agreement, difference plotted in percentage for the studies that reported MVUS. Total body estimates excluded in the plots. Values in plot A for the studies that reported data in cm3: Bias -1 %, SD 4

%, limits of agreement upper 7 % and lower -8 %. Values in plot B for the studies that reported data in kg: Bias -6 %, SD 13 %, limits of agreement upper 20 % and lower -32 %.