

Dual energy X-ray absorptiometry positioning protocols in assessing body composition: A systematic review of the literature

A systematic review of the literature

Shiel, Flinn; Persson, Carl; Furness, James; Perez Simas, Vinicius; Pope, Rodney R; Climstein, Michael; Hing, Wayne A; Schram, Ben

Published in:
Journal of Science and Medicine in Sport

DOI:
[10.1016/j.jsams.2018.03.005](https://doi.org/10.1016/j.jsams.2018.03.005)

Licence:
CC BY-NC-ND

[Link to output in Bond University research repository.](#)

Recommended citation(APA):
Shiel, F., Persson, C., Furness, J., Perez Simas, V., Pope, R. R., Climstein, M., Hing, W. A., & Schram, B. (2018). Dual energy X-ray absorptiometry positioning protocols in assessing body composition: A systematic review of the literature: A systematic review of the literature. *Journal of Science and Medicine in Sport*, 21(10), 1038-1044. <https://doi.org/10.1016/j.jsams.2018.03.005>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

Accepted Manuscript

Title: Dual energy X-ray absorptiometry positioning protocols in assessing body composition: A systematic review of the literature

Authors: Flinn Shiel, Carl Persson, James Furness, Vini Simas, Rodney Pope, Mike Climstein, Wayne Hing, Ben Schram



PII: S1440-2440(18)30072-0
DOI: <https://doi.org/10.1016/j.jsams.2018.03.005>
Reference: JSAMS 1826

To appear in: *Journal of Science and Medicine in Sport*

Received date: 2-6-2017
Revised date: 8-3-2018
Accepted date: 11-3-2018

Please cite this article as: Shiel Flinn, Persson Carl, Furness James, Simas Vini, Pope Rodney, Climstein Mike, Hing Wayne, Schram Ben. Dual energy X-ray absorptiometry positioning protocols in assessing body composition: A systematic review of the literature. *Journal of Science and Medicine in Sport* <https://doi.org/10.1016/j.jsams.2018.03.005>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title

Dual energy X-ray absorptiometry positioning protocols in assessing body composition: A systematic review of the literature

Authors

Flinn Shiel¹, Carl Persson¹, James Furness^{1,2}, Vini Simas², Rodney Pope⁴, Mike Climstein^{2,3}, Wayne Hing^{1,2}, Ben Schram^{1,2}.

¹ Physiotherapy Program, Faculty of Health Sciences and Medicine, Bond University, Robina, QLD, Australia

² Water Based Research Unit, Faculty of Health Science and Medicine, Bond University, Robina, QLD, Australia

³ Exercise Health & Performance Faculty Research Group, Faculty of Health Sciences, The University of Sydney, Lidcombe, NSW, Australia

⁴ Physiotherapy Program, School of Community Health, Charles Sturt University, Albury, New South Wales, Australia.

Contact

Dr James Furness, jfurness.bond.edu.au

Word Count – 3584

Abstract word count – 250

Number of Tables – 2

Number of Figures - 1

Abstract

Objectives: To systematically identify and assess methods and protocols used to reduce technical and biological errors in published studies that have investigated reliability of dual energy X-ray absorptiometry (DXA) for assessing body composition.

Design: Systematic Review

Methods: Systematic searches of five databases were used to identify studies of DXA reliability. Two independent reviewers used a modified critical appraisal tool to assess their methodological quality. Data was extracted and synthesised using a level of evidence approach. Further analysis was then undertaken of methods used to decrease DXA errors (technical and biological) and so enhance DXA reliability.

Results: Twelve studies met eligibility criteria. Four of the articles were deemed high quality. Quality articles considered biological and technical errors when preparing participants for DXA scanning. The Nana positioning protocol was assessed to have a strong level of evidence. The studies providing this evidence indicated very high test-retest reliability (ICC 0.90-1.00 or less than 1% change in mean) of the Nana positioning protocol. The National Health and Nutrition Examination Survey (NHANES) positioning protocol was deemed to have a moderate level of evidence due to lack of high quality studies. However, the available studies found the NHANES positioning protocol had very high test-retest reliability. Evidence is limited and reported reliability has varied in papers where no specific positioning protocol was used or reported.

Conclusion: Due to the strong level of evidence of excellent test-retest reliability that supports use of the Nana positioning protocol, it is recommended as the first choice for clinicians when using DXA to assess body composition.

Key words

Test-Retest Reliability, body fat; DXA; lean mass, positioning

Introduction

Dual-energy X-ray absorptiometry (DXA) is a widely accepted method for the assessment of

tissue composition¹. Low bone mineral density (BMD) and associated conditions such as osteoporosis and osteopenia constitute a significant health problem that costs over eight hundred and thirty million dollars annually and osteoporosis is a significant cause of morbidity and mortality^{2,3}. The need to accurately and effectively measure BMD in conditions such as osteoporosis led to the development of the DXA scanner⁴. Now, DXA is considered the gold standard for the assessment of BMD and associated fracture risk⁵. However, DXA is also a valuable clinical tool in the assessment of body composition (BC), due particularly to its ability to assess body segments for lean mass (LM) and fat mass (FM) distributions⁶. The absorption rates of the two different energy levels (40 and 70 KeV) within DXA coupled with the distinctive elements of bone, fat, and lean tissue enable clear imaging of each tissue type and subsequent analysis⁶. Therefore, DXA can be used for assessing segmental body composition (SBC) and is currently used in clinical, sporting and research settings. The data gathered from SBC scans have improved knowledge of malnutrition, growth, aging, obesity and the efficacy of medical treatment interventions (surgical, pharmacological, dietary and exercise)⁷. When used in the sport setting, DXA has enabled the tracking of players overall tissue composition as it has been found that individuals with the lowest start of season BMD and LM values have a greater occurrence of bone related injuries⁸. Nevertheless, the reliability of the DXA scanner is fundamental to the validity of all clinical investigations and research studies that use it to assess BC.

In order to draw valid and reliable conclusions from DXA scan results, the concept of error must be considered. The literature describes biological and technical error as sources for reduced test-retest reliability of the DXA scanner⁹. The International Society for Clinical Densitometry recommends precise measures during preparation of the participant (fasting state, clothing, time of day, physical activity and empty bladder) and consistent positioning⁹. It has been shown that sources of biological error in DXA results include hydration^{1, 9, 10}, stomach content and food consumption^{1, 9, 10}, time of day of scanning⁹ and physical activity^{9, 10}; furthermore sources of technical error include artefacts such as clothing⁹, number of operators used to complete scans¹¹ and position of participant^{1, 9, 12, 13}.

The influence of positioning of the participant on the DXA scanner can be analysed further by considering three identifiable positioning protocols. The first of these is the National Centre for Health Statistics, National Health and Nutrition Examination Survey (NHANES) Body Composition¹² positioning protocol, which the International Society for Clinical Densitometry recommends⁹. The NHANES protocol requires individuals to assume a supine position with feet secured together with a strap, and the palms of the hands flat on the scanning table and not touching the lateral aspect of the body. It should be noted that the Australian and New Zealand Bone Mineral Society (ANZBMS)¹⁴ employs the same body position. The second key protocol, the Nana positioning protocol¹, requires individuals to be in a supine position while placing hands in a neutral position alongside the body and feet in radio-opaque positioning pads. The third approach evident in the literature involves no specific positioning protocol being reported at all.

The study of Kerr et al¹³ is to date the only study that has attempted to compare the reliability of different DXA positioning protocols for assessing BC, to identify which protocol was the most valid and reliable to use in clinical practice. They reported the Nana positioning protocol was the preferred positioning protocol based upon participant comfort when assessing BC with DXA. In their study, the positioning protocols were modified versions of the standard Nana and NHANES protocols. In contrast, most other studies that have assessed the test-retest reliability of their DXA scanner have not compared the reliability of different positioning protocols.

Therefore, the aim of this literature review was to systematically identify and assess methods and protocols used in previously published research that has investigated reliability of DXA, when it is employed to assess BC, to reduce technical and biological errors.

ii. Methods

A search of academic databases was undertaken on 26.09.2016 with the intention of finding studies that have assessed the test-retest reliability of positioning protocols used when assessing

BC by DXA. The search was limited to studies conducted over the recent 10-year period (01.09.2006 to 26.09.2016) to maintain currency. The search was limited to only articles that included the term 'DXA' or a synonym for DXA in the title, as searches not limited in this way provided an excessive number of irrelevant articles. Details of the search strategy and key terms can be found in Figure 1.

Two reviewers (F.S and C.P) assessed the identified literature and removed duplicates. Titles and abstracts were initially screened and articles removed if eligibility criteria were not met. Inclusion criteria included: (1) studies conducted on living human participants, (2) studies of an adult population, and (3) studies primarily investigating reliability of DXA scanning protocols. Exclusion criteria were: (1) non-healthy subjects (eg subjects with: osteoporosis, current fractures, hemiarthoplasty and total joint replacements, rheumatoid or osteoarthritis, current cardiac or pulmonary conditions, or diabetes) (2) studies published prior to September 2006, (3) studies comparing MRI or CT to DXA, and (4) studies not available in English. In the event that insufficient details were provided in the titles and abstracts of articles to allow determination of eligibility, review of full texts was completed, with reference to eligibility criteria and ineligible articles were removed. The remaining articles were included in this literature review. A PRISMA flow diagram (Figure 1) was used to document the study screening and article selection processes¹⁵.

In order to critically appraise the included DXA reliability full text articles, a modified version of the reliability and validity critical appraisal tool (CAT) described by Brink and Louw¹⁶ was utilised, with items designed to appraise studies of validity removed, since the focus of this review was studies of reliability. The thirteen-item CAT was reduced to ten items by removing all items that did not relate to reliability, and was applied by two independent reviewers (F.S and C.P) in order to assess the methodological quality of each study. When both assessors were not in agreement, a consensus was reached by discussion to determine the item's final CAT results. The CAT did not originally include a scoring system; therefore for the purpose of this literature

review, a scoring system was implemented to aid in a quality and reliable analysis, similar to previously published reviews¹⁷⁻²⁰. Studies of higher quality scored $\geq 60\%$ in the modified CAT, and were rated higher due to their superior methodology²¹.

To receive a positive appraisal regarding the appropriateness of statistics in the CAT, each study reporting reliability must have reported an intraclass correlation coefficient (ICC) accompanied with confidence intervals (CI) or a percentage change in mean accompanied with typical error of measurement²². If the only basis for inclusion of a study was that it reported a percentage change in mean, then the calculation of the percentage change in mean must have complied with the guidance of previous work and have included a typical error of measurement in calculations^{23, 24}. Pearson correlation coefficients were not deemed suitable as measures of reliability; as they did not take into account the consistency of measurements from test to retest and the change in average measurements of participants²⁵. The ICC results of the studies that included ICC values were interpreted as indicators of reliability as follows: ICC of 0.00–0.29, very low reliability; 0.30–0.49, low reliability; 0.50–0.69, moderate reliability; 0.70–0.89, high reliability; and 0.90–1.00, very high reliability²⁶. An assessment of high or very high reliability depended primarily upon a reported high or very high ICC (above 0.70) or a low reported percentage change in the mean. When used the reported change in mean needed to be lower than the minimum clinically significant difference ascertained through consultation with practitioners. This ensured that any systematic error in repeated measurements observed during reliability testing was not sufficiently large to obscure clinically important changes or differences in the respective outcome measure – another indication of reliability. Unfortunately, only three studies in this review reported minimum clinically significant differences and therefore this statistic could not be used to compare studies.

Following critical appraisal, data were extracted from the included full text articles and tabulated to identify participant characteristics, the extent of standardisation employed to minimise technical and biological errors, the types of statistical analyses undertaken, and reported results of each study.

A meta-analysis was not undertaken due to the diversity of the methods examined and the statistical analyses employed. Rather, a critical narrative approach was applied to synthesise and analyse the data obtained from the included studies, using a level of evidence approach²⁷. Each positioning protocol identified from included studies was assigned a 'strong', 'moderate', or 'limited' level of evidence, based upon the number of studies that had examined its reliability and the quality of these studies. In order to be rated as having a strong level of evidence, a protocol required consistent findings from ≥ 3 high quality studies; to be assessed as having a moderate level of evidence, a protocol required consistent findings from at least 1 high quality study and 1 or more low quality studies; and to be assessed as having a limited level evidence, a protocol required consistent findings from ≥ 1 low quality study or only having 1 study available.

The use of standardization of methods of measurement to control sources of biological and technical error was assessed based upon the recommendations from the International Society of Clinical Densitometry⁹. Studies that reported having used the appropriate controls were considered more robust. As such the study must have included descriptions indicating how technical (clothing, positioning protocol) and biological (hydration, fasting state, time of day of scanning and physical activity) sources of error were controlled.

iii. Results

The results of the electronic database literature search and subsequent screening and selection process are depicted in Figure 1. The search yielded 128 results. After titles and abstracts were screened and clearly-ineligible studies and duplicates removed¹⁵, the full texts of 33 articles were obtained and further assessed based upon the inclusion/exclusion criteria. Twelve articles were subsequently included in this review.

A total of 724 participants were involved in the twelve (12) studies included in this review. Three hundred and eight (308) were males; two hundred and sixty seven (267) were females; and two

studies^{28, 29} involved one hundred and forty nine (149) participants but did not categorise participants' based on gender. The reported range of participant mean ages was 22.7 - 71.5 years, with the mean being 39.4 years. Nine studies reported mean mass, with the range of mean masses being 68.0 - 98.1 kgs and the mean being 77.1 kgs; similarly, the range of reported heights was 168.0 – 186.0 cm, with the mean being 174.8 cm. Three ³⁰⁻³² studies reported BMI instead of reporting mass and height and in those studies BMI ranged from 26.5 to 27.1 kg/m², with the mean being 26.8 kg/m².

Of the twelve included studies, only four of the studies were assessed as high quality using the CAT. Nine^{1, 10, 13, 28, 30, 31, 33-35} of the twelve studies reported statistics that were deemed appropriate (percentage change in mean accompanied by typical error or ICC). The three studies^{29, 32, 36} that failed to report appropriate statistics were deemed to be of insufficient quality to warrant a high rating from the CAT²¹. Seven^{1, 10, 13, 28, 32-34} of the twelve included studies used methods that were assessed as being reproducible, but only four^{1, 10, 13, 33} of these were rated as high quality studies. A detailed description of all articles assessed using the CAT can be found in Table 1.

The extent of standardisation of procedures to limit biological and technical errors varied significantly between the studies. Only three ^{1, 10, 13} of the twelve included studies reported all of the desirable information on the following standardised procedures: positioning protocol, clothing worn, physical activity completed by participant on day of scan, participant food intake on day of scan, participant hydration status and the time of day that the scanning took place. A further eight studies ^{1, 10, 13, 28, 29, 33-35} reported if clothing was worn, whilst seven studies^{1, 10, 13, 33-36} checked hydration status and six studies^{1, 10, 13, 28, 35, 36} reported scanning participants in a fasted state. Less than half of the studies^{10, 13, 28, 35} reported scanning participants in a rested state. The time of the scan was only reported in four^{1, 10, 13, 28} of the twelve studies.

The 12 included studies reported a variety of statistical representations of reliability, including percentage change in mean with the typical error of measurement, or ICC with CV. Of the studies

that reported ICC, all found the DXA results to have very high test-retest reliability²⁶. All studies that used a percentage change in mean as the test-retest reliability measure reported a change of less than one percent, and all percentage changes in mean were less than the minimum clinically significant difference. A summary of the reliability results from the included studies can be found in Table 2.

When applying a level of evidence approach, it was found that the Nana protocol had a strong level of evidence regarding DXA test-retest reliability, based on high quality articles as assessed by the CAT (Table 1), whilst the NHANES positioning protocol was deemed to have only a moderate level of evidence regarding reliability. This was due to only two high quality studies being reported in the literature for the NHANES positioning protocol, when available studies were assessed using the CAT (Table 1). Where no positioning protocol was reported in a study or a positioning protocol was not detailed, the level of evidence was deemed to be limited.

iv. Discussion

This literature review included twelve studies of test-retest reliability of DXA measurements when used to assess BC in healthy cohorts. The findings of these studies can assist in determining what factors need to be accounted for when using DXA scans to assess individuals for BC, to achieve high test-retest reliability in DXA results. Studies that accounted for both sources of technical error (scanner qualifications, reduction of chance of artefacts effecting results, the positioning protocol followed) and sources of biological error (hydration, stomach content and food consumption, time of day of scanning and effects of physical activity) were found to have superior methodologies and reported greater DXA test-retest reliability.

Additionally, this review examined which DXA positioning protocol for assessment of BC (Nana, NHANES, no specified protocol or no protocol) had the highest level of evidence regarding test-retest reliability. It was evident that the Nana positioning protocol had the highest level of

evidence regarding test-retest reliability of associated DXA results and this protocol was also deemed the most reliable protocol when conducting DXA scans for this purpose.

The Nana positioning protocol requires the use of pads, which are transparent under DXA, to minimise movement as well as increase reproducibility. Assessment of the studies of Nana et al¹, Nana et al¹⁰ and Kerr et al¹³ indicated the Nana positioning protocol was the most reliable based upon three considerations. Firstly, the critical appraisal of the methodological quality of these studies indicated they were high quality studies; secondly, the reliability results reported in these studies indicated high test-retest reliability of the DXA results: and lastly, the methodological provisions employed in these studies to minimise biological and technical errors were robust. The results of this review therefore support the findings of Kerr and colleagues¹³ that the Nana positioning protocol produces quality and reliable results; and also reinforces the original work of Nana¹ in the development of a superior positioning protocol.

The reliability of the NHANES positioning protocol in assessing BC has only been assessed in two studies^{13,33} and therefore can only be judged from a moderate level of evidence. The NHANES positioning protocol requires the participant to assume a supine position, with palms flat on the table and a strap securing the lower limbs to minimise movement¹². According to our CAT assessment, the overall methodological quality of these articles was high. The statistical results and methodological provisions to minimise technical and biological errors also appear to be sound. However, it is important to note that one of the included studies³³ lacks provision for the participant to be rested and standardisation of time of scanning. Ultimately, more high quality research is required for the NHANES positioning protocol before it could be recommended, based upon the criteria used in this review.

The level of evidence is limited from studies^{28, 30-32, 34} which have not followed a specific positioning protocol such as the Nana or NHANES protocol. This is a result of low methodological quality of these studies. The results not surprisingly indicate lower reliability of

DXA results when using such poorly-defined protocols. Additionally, all of the studies of this type did not include methodological provisions to standardise the participants to limit biological and technical errors.

A limited level of evidence was also yielded by studies^{29, 35, 36} that did not include a description of the positioning of the participants in the methods. This omission resulted in poor CAT scores and was associated with fluctuations in reported DXA results and the omission of methodological provisions to overcome sources of biological and technical errors.

Therefore, when scanning individuals using DXA to assess BC it is advised that clinicians use a positioning protocol such as the Nana¹ or NHANES¹² protocols to minimise technical errors and that they ensure the technician performing the scans is qualified. Additionally, accounting for biological sources of error (hydration, stomach content and food consumption, time of day of scanning and effects of physical activity) is vitally important when using the afore mentioned positioning protocols. Of these two protocols, the Nana protocol currently has the highest level of evidence indicating that it should be the preference for clinicians.

Interestingly, Kerr et al¹³ also included a measure of comfort of participants. In this study, they used a modified version of the Nana positioning protocol (adding straps around the waist to secure the arms and the distal lower limb to “mimimise any subject movement”) and a modified version of the NHANES positioning protocol (in which the participants hands were placed against the body but not secured). It could be postulated that these changes to the original positioning protocols may have favoured the Nana positioning protocol, as subjects in the NHANES protocol had to actively hold their arms in a static position during the DXA scan. Perhaps not surprisingly, then, the Nana protocol was favoured by participants based on comfort.

Limitations of this literature review include the non-inclusion of grey literature, and the focus of the literature review being only test-retest reliability. This latter focus may have excluded some

studies which did not report test-retest reliability in their abstracts. The removal of non-English studies and the exclusion criterion of non-healthy subjects may have also reduced the number of included studies in this review. Additionally, this review only focused on whole body BC scans and did not consider hemiscans or compilations of partial scans, as there is a scarcity of studies that have investigated this technique.

Strengths of this literature review include the systematic approach employed and the rigorous methodology followed, using the PRISMA statement¹⁵ as a guide. Additionally, the utilisation of the modified CAT tool and independent reviewers aided and upheld high quality assessments of methodological quality. Furthermore, this is the only literature review to assess multiple variables in the methodology that affect the reliability of DXA measurements of SBC.

This literature review has affirmed the need for more high quality research to assess the test-retest reliability of DXA measurements of BC using the NHANES positioning protocol. Clinicians would benefit from research that more robustly compares the Nana and the NHANES positioning protocols. Robust further research would serve to elevate the NHANES positioning protocol to a similar level of evidence as the Nana positioning protocol.

v. Conclusion

This review aimed to assess the different protocols and methodological approaches used to reduce technical and biological errors in previously published studies that have investigated test-retest reliability of DXA when used to assess BC. The results of this literature review can usefully guide for future clinicians using DXA to assess BC in a variety of settings including elite sport, community health and research. As such, this review indicates that the Nana positioning protocol, when coupled with methodological provisions to minimise biological and technical sources of error, is the positioning protocol with the strongest level of evidence and high levels of test-retest reliability, and thus should be the choice of clinicians when using DXA to assess BC. Currently, moderate level evidence of high test-retest reliability exists for the NHANES positioning protocol

and more high quality research using this protocol is required to enhance the level of available evidence. Not using a positioning protocol or not reporting the protocol employed means studies of DXA reliability are then of low methodological quality; too low to enable recommendations to be made based on their findings.

vi. Practical Implications

Methodological provisions to reduce technical errors and biological errors is of paramount importance to produce reliable DXA measurements of BC.

The use of positioning protocols in such DXA scanning increases the reliability of results.

To minimise technical error, the Nana positioning protocol should be the first choice for clinicians when assessing BC.

vii. Acknowledgements

The authors alone are responsible for the content and writing of this article and did not receive financial assistance to perform the current review.

References

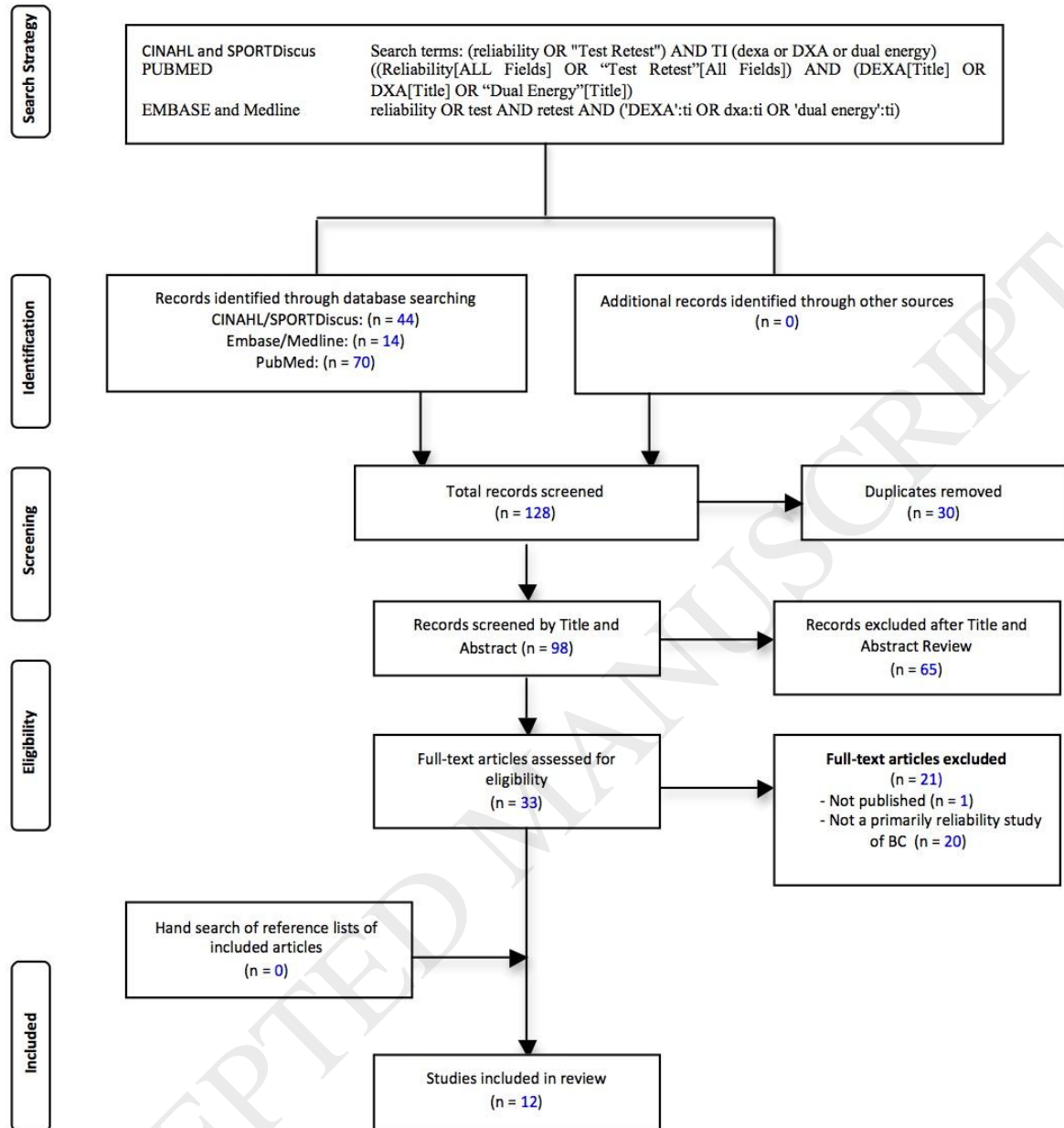
1. Nana A, Slater GJ, Hopkins WG, Burke LM. Effects of daily activities on dual-energy X-ray absorptiometry measurements of body composition in active people. *Med Sci Sports Exerc.* 2012; 44(1):180-189.
2. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporosis International.* 2006; 17(12):1726-1733.
3. Watts JJ, Abimanyi-Ochom J, Sander KM. Osteoporosis costing all Australians. A new burden of disease analysis – 2012 to 2022, *Osteoporosis Australia;* 2013.
4. Lewiecki EM. Clinical applications of bone density testing for osteoporosis. *Minerva Med.* 2005; 96(5):317-330.
5. Blake GM, Fogelman I. The role of DXA bone density scans in the diagnosis and treatment of osteoporosis. *Postgraduate Medical Journal.* 2007; 83(982):509-517.
6. Rothney MP, Brychta RJ, Schaefer EV, Chen KY, Skarulis MC. Body composition measured by dual-energy X-ray absorptiometry half-body scans in obese adults. *Obesity.* 2009; 17(6):1281-1286.

7. Lee SY, Gallagher D. Assessment methods in human body composition. *Curr Opin Clin Nutr Metab Care*. 2008; 11(5):566-572.
8. Georgeson EC, Weeks BK, McLellan CP, Beck BR. Body Composition Change Over a Professional Rugby League Season and Relationship to Rates and Types of Injury. *Med Sci Sports Exerc*. 2011; 43(5 (Suppl 1)):108.
9. Hangartner TN, Warner S, Brailion P, Jankowski L, Shepherd J. The Official Positions of the International Society for Clinical Densitometry: acquisition of dual-energy X-ray absorptiometry body composition and considerations regarding analysis and repeatability of measures. *J Clin Densitom*. 2013; 16(4):520-536.
10. Nana A, Slater GJ, Hopkins WG, Burke LM. Effects of exercise sessions on DXA measurements of body composition in active people. *Med Sci Sports Exerc*. 2013; 45(1):178-185.
11. Kiebzak GM, Leamy LJ, Pierson LM, Nord RH, Zhang ZY. Measurement precision of body composition variables using the lunar DPX-L densitometer. *J Clin Densitom*. 2000; 3(1):35-41.
12. National Health and Nutrition Examination Survey (NHANES). 2011. Body Composition Procedures Manual. https://www.cdc.gov/nchs/data/nhanes/nhanes_13_14/2013_Body_Composition_DXA.pdf
13. Kerr A, Slater GJ, Byrne N, Nana A. Reliability of 2 Different Positioning Protocols for Dual-Energy X-ray Absorptiometry Measurement of Body Composition in Healthy Adults. *J Clin Densitom*. 2016; 19(3):282-289.
14. Australian and New Zealand Bone & Mineral Society (ANZBMS). 2014. Clinical bone densitometry training course. ANZBMS, Sydney.
15. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009; 62(10):e1-e34.
16. Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract*. 2012; 18(6):1126-1132.
17. van der Wurff P, Hagmeijer RH, Meyne W. Clinical tests of the sacroiliac joint. A systematic methodological review. Part 1: Reliability. *Man Ther*. 2000; 5(1):30-36.
18. May S, Chance-Larsen K, Littlewood C, Lomas D, Saad M. Reliability of physical examination tests used in the assessment of patients with shoulder problems: a systematic review. *Physiotherapy*. 2010; 96(3):179-190.
19. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review. *Aust J Physiother*. 2006; 52(2):91-102.
20. Adhia DB, Bussey MD, Ribeiro DC, Tumilty S, Milosavljevic S. Validity and reliability of palpation-digitization for non-invasive kinematic measurement - a systematic review. *Man Ther*. 2013; 18(1):26-34.
21. Barrett E, McCreesh K, Lewis J. Reliability and validity of non-radiographic methods of thoracic kyphosis measurement: A systematic review. *Man Ther*. 2014; 19(1):10-17.
22. Ionan AC, Polley M-YC, McShane LM, Dobbin KK. Comparison of confidence interval methods for an intra-class correlation coefficient (ICC). *BMC Med Res Methodol*. 2014; 14:121.
23. Hopkins WG. Measures of Reliability in Sports Medicine and Science. *Sports Med*. 2000; 30(1):1-15.
24. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. *Med Sci Sports Exerc*. 2009; 41(1):3-13.
25. Rankin G, Stokes M. Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses. *Clin Rehabil*. 1998(12):187-199.
26. Munro B, Visintainer M. *Statistical methods for health care research*, 5th ed, Lippincott, Philadelphia; 2005.
27. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine*. 2003; 28(12):1290.

28. Smith-Ryan AE, Mock MG, Ryan ED, Gerstner GR, Trexler ET, Hirsch KR. Validity and reliability of a 4-compartment body composition model using dual energy x-ray absorptiometry-derived body volume. *Clin Nutr.* 2016.
29. Wilson JP, Fan B, Shepherd JA. Total and regional body volumes derived from dual-energy x-ray absorptiometry output. *J Clin Densitom.* 2013; 16(3):368-373.
30. Covey MK, Smith DL, Berry JK, Hacker ED. Importance of Cross-Calibration When Replacing DXA Scanners: QDR4500W and Discovery Wi. *J Nurs Meas.* 2008; 16(3):155-170.
31. Covey MK, Berry JK, Hacker ED. Regional body composition: Cross-calibration of DXA scanners QDR4500W and Discovery Wi. *Obesity.* 2010; 18(3):632-637.
32. Lohman M, Tallroth K, Kettunen JA, Marttinen MT. Reproducibility of dual-energy x-ray absorptiometry total and regional body composition measurements using different scanning positions and definitions of regions. *Metabolism.* 2009; 58(11):1663-1668.
33. Bilsborough JC, Greenway K, Opar D, Livingstone S, Cordy J, Coutts AJ. The accuracy and precision of DXA for assessing body composition in team sport athletes. *J Sports Sci.* 2014; 32(19):1821-1828.
34. Colyer SL, Roberts SP, Robinson JB, et al. Detecting meaningful body composition changes in athletes using dual-energy x-ray absorptiometry. *Physiol Meas.* 2016; 37(4):596-609.
35. Hurst PR, Walsh DCI, Conlon CA, Ingram M, Kruger R, Stonehouse W. Validity and reliability of bioelectrical impedance analysis to estimate body fat percentage against air displacement plethysmography and dual-energy X-ray absorptiometry. *Nutr Diet.* 2016; 73(2):197-204.
36. Moon JR, Stout JR, Smith-Ryan AE, et al. Tracking fat-free mass changes in elderly men and women using single-frequency bioimpedance and dual-energy X-ray absorptiometry: a four-compartment model comparison. *Eur J Clin Nutr.* 2013; 67 Suppl 1:S40-46.

x. Figure Legends

Figure 1. PRISMA flow diagram of literature search strategy



x. Tables

Table 1: Critical Appraisal Tool

Article	Subjects	Characteristics	Scanner qualified	Rater Intra	Rater Inter	Random order	Test retest Time period	Test retest Period valid	Can reproduce protocol	Withdrawals explained	Statistics reported appropriate	High Quality?
Bilsborough et al. ³³	47	YES	YES	NA	NO	YES	YES	YES	YES	NO	YES	YES
	47M	22.7 ± 3.0 years 84.4 ± 5.62 kg, 186 ± 5 cm					Immediate				ICC %CV	
Colyer et al. ³⁴	53	YES	NO	NO	NO	NO	YES	YES	YES	NO	YES	NO
	34M	23.0 ± 4.0 years					2 days				ICC	
	19F	79.9 ± 18.9 kg, 178 ± 10 cm									%CL	
Covey et al. ³⁰	42	YES	YES	NA	NO	NO	YES	YES	NO	NO	YES	NO
	1M	50.4 ± 9.9 years					1 day				% Difference Limit of Agreement	
	41F	27.1 ± 6.1 kg/m ²										
Covey et al. ³¹	39	YES	YES	NA	NO	NO	YES	YES	NO	NO	YES	NO
	39F	56.6 ± 9.6 years					7-14 days				% Difference Limit of Agreement	
		26.8 ± 5.5 kg/m ²										
Hurst et al. ³⁵	166	YES	NO	NA	NO	NO	YES	YES	NO	NO	YES	NO
	81M	38.9 (36.9-40.9) years					Max 5 days				% Change	
	85F	75.5 kg, 171 cm										
Kerr et al. ¹³	30	YES	YES	NA	NO	YES	YES	YES	YES	NO	YES	YES
	14M	36 ± 11.5 years					Immediate				% Change	
	16F	71.0 ± 7.1 kg, 173.5 ± 0.5 cm									%CV	
Lohman et al. ³²	30	YES	YES	NA	NO	NO	YES	YES	YES	NO	NO	NO
	30M	45.2 (22.0-61.0) years					Immediate				Pearson r value	
		26.5 (17.8-33.9) kg/m ²										
Moon et al. ³⁶	82	YES	NO	NA	NO	YES	YES	NO	NO	NO	NO	NO
	44M	71.5 ± 5.3 years					12 Weeks				Pearson r value	
	38F	71.4 ± 8.1 kg, 168.3 ± 5.3 cm										
Nana et al. ¹	31	YES	YES	NA	NO	NO	YES	YES	YES	NO	YES	YES
	16M	27.0 ± 5.0 years					Immediate				% Change	
	15F	68 ± 7.5 kg, 172.5 ± 6.0 cm									%CV	
Nana et al. ¹⁰	55	YES	YES	NA	NO	NO	YES	YES	YES	NO	YES	YES
	41M	27.7 ± 6.3 years					Immediate				% Change	
	14F	75.5 ± 7.9 kg, 176.4 ± 5.7 cm									%CV	
Smith-Ryan et al. ²⁸	127	YES	NO	NA	NO	NO	YES	YES	YES	NO	YES	NO
		35.8 ± 9.4 years					7-10 days				ICC	
		98.1 ± 20.9 kg, 176.3 ± 9.2 cm									%CV	

Wilson et al. ²⁹	22	YES	YES	NO	NO	NO	NO	YES	NO	NO	NO	NO
		37.8 ± 15.5 years									NO	NO
		70.1 ± 14.8 kg,									Root	
		172.0 ± 11.4 cm									Square	
											Mean	
											%CV	

M Male, *F* Female, *cm* centimetres, *kg* kilograms, *kg/m²* kilograms per square metre, *% Change* Percentage change in mean, *ICC* Intraclass Correlation Coefficient, *%CV* Percentage of Coefficient of Variation *CL%* Confidence Limit Percentage, *SEE* Standard Error Estimation

ACCEPTED MANUSCRIPT

Table 2: Overview of results of studies of test-retest reliability of DXA measurements of BC

Authors	Variable/Condition	Intrarater reliability between scans						High quality
		ICC or % Change in mean* or % Difference** or Pearson Correlation #			CV% (TEM or SEM) or CL%* or Limit of Agreement #			
		BMC	LM	FM	BMC	LM	FM	
Bilsborough et al. ³¹	Lunar - Fan Beam	1.00	1.00	0.99	0.60	0.30	2.50	Yes
Bilsborough et al. ³¹	Lunar - Pencil Beam	1.00	1.00	0.98	1.50	0.50	5.90	Yes
Colyer et al. ³²	-	-	1.00	0.99	-	1.2*	1.2*	Yes
Covey et al. ²⁸	Discovery Wi Machine ##	0.34**	-0.07**	0.29**	-0.04-0.06#	-1.10-1.10#	-0.70-0.77#	No
Covey et al. ²⁸	QDR Machine ##	-0.40**	0.14**	-0.01**	-0.09-0.07#	-1.39-1.51#	-0.87-0.95#	No
Covey et al. ²⁹	QDR Machine ##	-0.50**	0.30**	0.00**	-0.05-0.05#	-1.16-1.27#	-0.67-0.65#	No
Covey et al. ²⁹	Discovery Wi Machine ##	0.20**	-0.10**	0.60**	-0.03-0.03#	-0.81-0.83#	-0.81-0.83#	No
Hurst et al. ³³	-	-	-	0.01*	-	-	-	No
Kerr et al. ¹¹	NHANES Protocol	0.10*	-0.10*	0.20*	0.90	0.80	2.60	Yes
Kerr et al. ¹¹	NANA Protocol	-0.40*	0.20*	-0.20*	1.00	0.60	2.20	Yes
Lohman et al. ³⁰	-	0.99#	0.99#	1.00#	-	-	-	No#
Moon et al. ³⁴	-	-	0.87F 0.95M #	-	-	-	-	No#
Nana et al. ¹	Immediate retest	-0.10F 0.30M*	0.20F 0.00M *	0.00F - 0.40M *	1.00F 0.70 M	0.50F 0.40M	1.30F 1.90M	Yes
Nana et al. ¹	Retest 24 hours later	-0.30F - 0.20M*	0.00F - 0.20M *	-0.40F - - 0.60M *	1.10F 0.70 M	1.00F 0.50M	1.30F 2.10M	Yes
Nana et al. ⁸	Strength Group	0.00*	0.00*	0.10*	1.00	0.60	2.50	Yes
Nana et al. ⁸	Cycling Group	-0.10F 1.90M*	0.30F 0.00M *	-0.10F - - 0.10M *	0.80F 5.20 M	0.80F 1.50M	1.90F 1.40M	Yes
Smith-Ryan et al. ²⁶	-	-	0.996	0.995	-	0.83	0.99	Yes
Wilson et al. ²⁷	-	0.85^ Total mass			1.08 Total mass			No

BMC Bone Mineral Content, LM Lean Mass, FM Fat Mass, ICC Intraclass Correlation Coefficient, CV% Percentage of Coefficient of Variation, CL% Confidence Limit Percentage, ^Root Square Mean Error, ## regional assessment of trunk only, CV% TEM (typical error of measurement) or SEM (standard error of measurement) percentage, F Female, M Male