

Comparative Analysis of Quantitative Whole Body Orthostatic Postural Measures: A Systematic Review

A thesis presented in candidature for the degree of Master of Research - Chiropractic

Curtis Thor Rigney

Doctor of Chiropractic (USA)

Department of Chiropractic, Faculty of Science and Engineering

Macquarie University, Sydney, New South Wales, Australia

6 January 2018

Table of Contents

Abstract.....	iv
Candidate Statement	v
Acknowledgements	vi
List of Figures	vii
List of Tables	viii
List of Abbreviations	ix
Chapter 1: Introduction.....	1
1.1 Background.....	1
1.1.1 Conventional postural measurement.....	2
1.1.2 Relevance of the measurement of posture.....	3
1.2 Evidence based practice	4
1.2.1 Definitions	4
1.2.2 Matter of variables.....	6
1.3 Current status in the literature	7
1.4 Aims of the study	9
Chapter 2: Methods	10
2.1 Protocol and registration	10
2.2 Systematic literature review	10
2.2.1 Data sources and search terms	10
2.2.2 Eligibility criteria.....	11
2.2.3 Study selection	11
2.2.4 Data extraction	11
2.2.5 Quality assessment	12
2.3 Grey literature search	12
2.3.1 Grey literature screening and data extraction	13
Chapter 3: Results	13
3.1 Systematic search results	13
3.1.1 Included studies	15
3.1.2 Photogrammetry protocols	15
3.1.3 Ordinal scale protocol	16
3.1.4 Posture capture protocol	19

3.1.5 Population characteristics and enrollment method	20
3.2 Qualitative analysis	22
3.3 Grey literature	24
Chapter 4: Discussion	25
4.1 Results summary	25
4.2 Discussion on validity and reliability	26
4.3 Discussion of CAT results	27
4.3.1 Quality assessment challenges	28
4.3.2 Alternative outcomes	30
4.4 Statistical analyses	30
4.5 Variables	31
4.5.1 Image capture	31
4.5.2 Ordinal score	33
4.5.3 Photogrammetry	33
4.5.4 Precision	35
4.6 Alignment with convention: generalisability	36
4.6.1 Population characteristics	37
4.6.2 Views	37
4.7 Grey literature contribution	38
4.8 Summary	39
Chapter 5: Strengths and limitations of the review	40
5.1 Strengths	40
5.2 Limitations	41
Chapter 6: Future research	42
Chapter 7: Conclusions	42
References	44
Appendix	49

Abstract

Background: Measuring whole body posture has played a part in the diagnosis of musculoskeletal conditions for decades. However, there is little evidence supporting the supposed impact of faulty posture on the musculoskeletal system. A valid and reliable method of whole body postural measurement is required to build upon the evidence. The aim of this study was to identify and qualitatively assess quantitative methods of whole body orthostatic postural measurement that are valid, reliable, and that reflect conventional practices. Also investigated was the management of variables that affect posture.

Methods: A systematic search through eight electronic databases and a search through grey literature were conducted. Two independent reviewers critically analysed methodologies with a critical analysis tool.

Results: Ten articles retrieved from the literature search and three systems from the grey literature found that photogrammetry and ordinal scaling were two methods of measurement. Four studies were of high methodological quality but did not reflect conventional practice. A narrow aspect of reliability was assessed while validity was undetermined. The protocols poorly managed postural variables.

Conclusions: Photogrammetry has good potential for precise measurements of whole body orthostatic posture. Further research in establishing protocols to improve the reliability and validity of postural measurement is necessary.

Candidate Statement

I certify that the work incorporated in this thesis has not been submitted for a higher degree to any other university or institution.

I certify that the work presented in this investigation is original work except where otherwise acknowledged and referenced in the text of this manuscript.

A handwritten signature in black ink, reading "Curtis Thor Rigney". The signature is written in a cursive style with a large, stylized 'C' and 'R'.

Curtis Thor Rigney

6 January 2018

Acknowledgements

The Master of Research – Chiropractic program was a challenging but rewarding experience. I am glad to have undertaken the degree. Although it is my name on this thesis, I could not have achieved its completion without the support from the Department of Chiropractic, the Faculty of Science and Engineering, and the Macquarie University. There are individuals I wish to acknowledge and express my gratitude for their contributions.

Firstly to my supervisor, Dr. Stephney Whillier, for her patience and attention to detail. To Dr. Roger Engel for his help at the start of my degree process. To Dr. Benjamin Brown for his enormous help with navigating the degree's structure and numerous deadlines.

I owe a debt of gratitude to Hazel Jenkins for her willingness to be the third reviewer. To Aron Downie for helping me through a panic episode related to 'prevalence'. To Dr. Petra Graham for helping me to better understand statistics. And to Christopher Burrell for shouldering much of my teaching responsibilities while I focused on my research.

I will always remember the generous support I received from Rosemary Giuriato, as the Head of the Department of Chiropractic and Dr. Barbara Messerle, the Executive Dean of the Faculty of Science and Engineering. I am grateful to my superiors for their support through the resource of time.

The Macquarie University offered many workshops, online tutorials, and guest lecturers that were helpful in guiding my journey through the program. I am thankful for the endeavours taken by the University to help students succeed. The Macquarie University library is an extraordinary resource with incredible access to information and research librarians. Fiona Jones and Jo Hardy have forever improved my quality of life by helping me crack the 'EndNote' code and revolutionising my search skills. I am so grateful for these librarians.

Finally, I wish to express my gratitude to Joseph Curry. His consistent presence, love, and support has helped ease the pain of the struggle and maximise the joy of the success.

List of Figures

Figure 1	Example plumb line placement.....	2
Figure 2	PRISMA results flow diagram.....	14
Figure 3	Search engine search results flow diagram.....	24

List of Tables

Table 1	Search syntax example.....	10
Table 2	Included articles in alphabetical order, aims, statistics, conclusions...	17
Table 3	Photogrammetry data.....	18
Table 4	Posture capture protocol data.....	20
Table 5	Population characteristics data.....	21
Table 6	Methodological quality evaluation using Brink & Louw tool.....	23
Appendix	Brink & Louw critical analysis tool.....	49

List of Abbreviations

A –	Anterior
ASIS –	Anterior superior iliac spine
BMI –	Body mass index
CAT –	Critical analysis tool
CI –	Confidence interval
COSMIN –	Consensus-based Standards for the selection of health Measurement Instruments
CR –	Curtis Rigney
EBM –	Evidence based medicine
ICC –	Intraclass correlation coefficient
L –	Lateral
MeSH –	Medical search heading
MSK –	Musculoskeletal
O –	Oblique
P –	Posterior
PAS –	Postural Assessment Software
PRISMA –	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QAREL –	Quality Appraisal of Diagnostic Reliability Studies
QUADAS –	Quality Assessment of Diagnostic Accuracy Studies
S –	Superior
SEm –	Standard error of measurement
STARD –	Standard for the Reporting of Diagnostic Accuracy
STROBE –	Strengthening the Reporting of Observational Studies in Epidemiology
SW –	Stephney Whillier

Chapter 1: Introduction

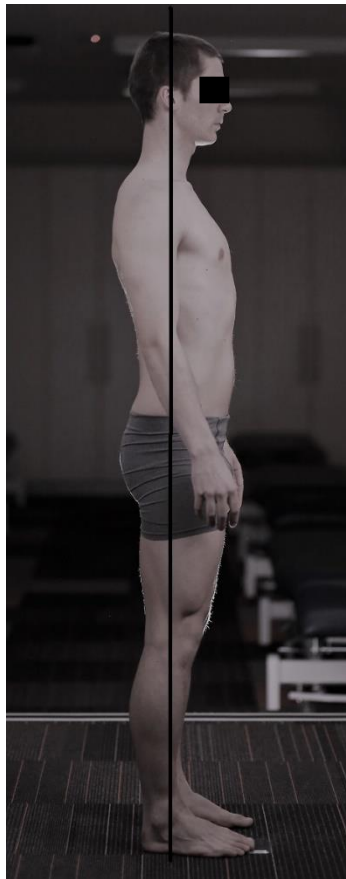
1.1 Background

The concept of posture often conjures an image of someone standing or sitting. Posture is defined as the position of body parts or regions in relation to each other¹. Numerous types of posture are adopted including standing, sitting, squatting, and lying. Although posture is an arrangement of parts, all body parts are interconnected such that the displacement of one part can affect displacement in other parts¹. For this reason, it has been advised that clinicians consider the entire body when measuring posture².

Postural measures have been a component of the physical examination for decades. In the early 20th century, postural measurement focused on observing and classifying a person as having a particular postural type³⁻⁵. Correct or good posture during this era was modelled after ancient Greek statues of Olympians; in particular, “the Spear Thrower” by Polycleitos⁶. Posture’s effect on general health was the concern. For example, the “slender” body type was thought to be more prone to disease compared to those deemed “Spartan-like” or “ontomorphic”⁵. The culture around posture started to shift in the middle of the 20th century.

The modern approach to the measurement of posture shifted the concern from health and fitness to the carriage of the centre of mass. Considerations of biomechanics, the use of plumb lines and the association of musculoskeletal pain with faulty posture became the standard⁷. A plumb line is a weighted string that represents the vertical pull of gravity. Placed next to a person, the plumb line illustrates the pull of gravity on the body’s weighted average of its different regions; its centre of mass⁸ (Fig 1). Because of the vertical downward pull from gravity on the centre of mass, it is commonly referred to as the centre of gravity⁹. Its placement is believed to influence the entire body⁸. The plumb line, representing the centre of gravity, is fundamental to the understanding of the modern “ideal” posture. The body’s observed placement along this line demonstrates its balance. It illustrates how well the body’s bones, joints, muscles, and nervous system are maintaining an upright posture against the pull of gravity.

Figure 1. Example of plumb line placement



Ideal alignment of the body is how Kendall *et al*¹⁰ defines the ‘ideal’ or standard posture. This hypothetical position of alignment and equilibrium is achieved when the body’s distribution of weight is balanced around a centrally placed gravitational line, described as the intersection of the mid-coronal and mid-sagittal planes. In conditions of ideal alignment, less force and energy are required to maintain the upright posture. Fewer muscles are activated and body weight is carried through bones rather than joint or soft tissues. Which results in less energy consumption, less adverse force placed on muscles, joints, and soft tissues, and less pain.

1.1.1 Conventional postural measurement

The Kendalls, Florence and Henry, have had significant influence on the development of modern postural measurement. Their first publication on the topic was in 1952¹¹ with a fifth edition published in 2005¹⁰. Due to their long-standing history, influence, and continued relevancy, this

paper will refer to the Kendall method of postural measurement as being representative of a conventional measurement of whole body orthostatic posture. This method of postural measurement attempts to capture an individual's usual or habitual posture and requires the individual to be minimally dressed allowing visualisation of the body in four standard views; posterior, anterior, left lateral, and right lateral. The body contours must be seen as they relate to underlying musculoskeletal structure. A plumb line is used to illustrate the gravitational line from a fixed point, the individual's feet when in the standing position. The head, shoulders, trunk, pelvis, and lower extremities are considered to be in ideal alignment when they are symmetrically aligned with the plumb line. Deviations from the hypothetical ideal or standard position are considered postural faults. According to the Kendalls, postural deviations should be graded as slight, moderate, or marked. Observation of body contours can reveal additional postural faults through visual assessment of joint angles and muscle characteristics.

1.1.2 Relevance of the measurement of posture

Numerous health care professions measure posture as a component of the physical examination of patients. Chiropractic, physical therapy, osteopathy, and orthopaedics all use textbooks that describe an observational method for the measure of whole body orthostatic posture^{10,12-15}. The reason for this prevalence is the widely held belief that poor posture contributes to the development or maintenance of musculoskeletal (MSK) pain^{1,2,7,10,12-16}. A profession concerned about the diagnosis, management and prevention of MSK pain is essentially interested in posture because of the prevalence of this pain.

Musculoskeletal conditions are the most common pain complaints seen by physicians^{17,18}. Low back and neck pain, together, are the fourth leading cause of disability globally¹⁹. Back pain in itself is the fifth most common reason for all physician visits^{20,21}. Given this high prevalence of MSK conditions, and the burden it places on the health system, it is prudent to strive continually to strengthen the methods of diagnosis, management, and prevention of MSK pain conditions. Furthermore, given that the measurement of posture is in common use amongst MSK clinicians, and that MSK conditions are commonly seen and are a public health concern, it behooves clinicians to look to the development of evidence-based methods of measurement that are precise, accurate and reliable.

1.2 Evidence based practice

The ‘evidence based medicine’ (EBM) paradigm, as described by Sackett and others²²⁻²⁴, provides a guide for health care professionals in this new era of well-supported diagnosis and management. The EBM practitioner utilises the best available research and clinical experience to aid in clinical decisions related to diagnosis and treatment. These decisions are tempered and balanced by taking into account the needs of individual patients and the clinical environment, in a holistic approach to the patient’s wellbeing and care. The measurement of whole body orthostatic posture has a role in physical examination and diagnosis. Therefore, it falls into the diagnostic accuracy research domain²⁵. Valued attributes of diagnostic tests are validity and reliability. A common method for validating the results provided by a diagnostic test is by comparing it under blinded conditions with a gold standard²⁶. A diagnostic test that is able to repeatedly distinguish between individuals with or without a ‘target condition’ is considered to be reliable²². Demonstrating these two attributes is a challenge for whole body orthostatic measurement because a gold standard does not yet exist and the target condition is unclear. The terminology regarding validity and reliability can be confusing due to a lack of consistency. Attempts to standardise the related taxonomy have been made²⁷. However, it remains necessary to provide the operational definitions relevant to this paper.

1.2.1 Definitions

Validity is a reflection of the degree that a measurement method or tool truly measures what it purports to measure²⁷. It asks the question; does the diagnostic test actually measure what it says it measures? Sensitivity and specificity are measures of validity. They relate to a test’s ability to correctly identify individuals with and without a target condition²⁴. The calculation of these properties require a comparison with a gold or reference standard. Accuracy is a term that refers to the level of agreement between a measurement tool and a gold standard²⁸. Accuracy, then, is a statement of validity.

Reliability represents the degree to which a measurement is free from measurement error²⁷. It reflects the ability to which a measurement tool can differentiate between people, those with and without the target condition^{27,29}. Reproducibility is a property of reliability and represents the extent to which scores from repeated measures are the same for a stable person³⁰. These repeated

measures can occur over time (test-retest), by the same rater on different occasions (intra-rater), or between two raters on the same occasion (inter-rater)²⁷. Precision is a measure of reliability as it reports the closeness repeated measures are between raters²⁸. Measurement error is a significant determination of reliability. The result from a diagnostic tool is comprised of two measurements; the true measure and the measure that comes from error. When a second measure is different from the first, it is important to understand the source of that difference. Was this measurement difference the result of a true change or a change that resulted from an error? Some sources of measurement error can come from the instrument, from the operators of the tool or both³⁰. Additional sources can come from natural variations of the individual being measured. For example, a person's body sways even when standing still⁸. Measuring posture at different points within the sway may result in inconsequential differences.

The process of validation for a diagnostic test involves a comparison with a 'gold standard'. As defined by a cohort of researchers behind the Standard for the Reporting of Diagnostic Accuracy (STARD) initiative, a gold standard is measurement that is free from error^{25,31}. The group admits that few diagnostic tests are error-free and suggest an alternative term, 'clinical reference standard'. A reference standard represents the 'best available' method in determining the presence of a target condition. For postural measures, radiographic images have been suggested to be the gold standard³². This may be true for the measurement of spinal curves, but it is not the case for whole body posture. In the case of measuring whole body orthostatic posture, the best available reference standard is the subjective visual analysis described by Kendall¹⁰ and others^{14,33}.

A target condition is the clinical entity, disease, condition, or response from therapy, that the diagnostic test is designed to detect³¹. The target condition that whole body postural measurement seeks to identify is poor posture, the less than ideal alignment of the body's parts. This target condition presents some unique challenges. Firstly, poor posture in itself is not a definitive diagnosis. Its measurement may lead to an understanding of cause or prognosis of pain but the degree to which poor posture contributes to this pain is uncertain. It is not known how significant a postural fault must be before it becomes a threat to MSK health. Another challenge is the ubiquitous nature of poor posture. Poor posture is common¹⁰. Brown⁴ used

schematography to measure the standing posture of 746 young adult men. This method involved tracing the outline of a person in a reduced form. In Brown's examination, the vertical line (similar to the gravitational line), head position, abdominal shape, and lateral spinal curves were measured and then compared to a standard. This standard posture was deemed normal, which was similar to today's ideal posture. Of the 746 men, as many as 90% were determined to have poor posture⁴. It is worth noting that a high prevalence of a target condition affects the psychometric assessment of the condition's measurement tool and is a potential threat to the validity of its reported accuracy³⁴. The absence of an appropriate clinical reference standard and a clear target condition creates challenges for validating quantitative methods of measuring whole body orthostatic postural. However, these challenges are not insurmountable. Developing a reliable method of whole body orthostatic postural measurement would be a step closer toward validity.

1.2.2 Matter of variables

The measurement of whole body orthostatic posture is associated with numerous non-pathological variables that affect posture. These variables challenge the development of a reliable postural measurement tool as they are potential sources of measurement error. A reliable method of postural measurement must address these variables and demonstrate repeatable results. The variables are both intrinsic and extrinsic in nature. Examples of intrinsic variables are the presence of symptoms, fatigue, and mood. Symptomatic individuals have demonstrated greater postural faults than those that are asymptomatic³⁵. Mood and emotion can affect posture^{36,37}. Cureton commented on how one's mental attitude can affect posture saying that, "physical poise is related to mental poise"³⁸. Fatigue also affects posture. A person's posture can differ in the afternoon from how it started in the morning. The same is true for before and after physical activity³⁹. Because these variables can transiently affect posture, they need to be taken into account and controlled. Otherwise, measured changes in posture may not be true changes. The presence of symptoms (pain) is a challenge for the establishment of a reliable method of postural measurement. Most individuals undergo postural measurement because of pain and therefore, will have symptoms.

Extrinsic variables relate to the protocols employed for postural measurements. Examples include the dress code imposed on the subject, the instructions given to position the subject, and the length of time required for the measurement. The placement of the plumb line bob is another example. This placement has not been standardised for lateral views. Some suggest placing it ‘slightly anterior’ to the lateral malleolus^{10,14}. Others suggest placing the plumb bob at the lateral malleolus³³. This small variation in bob placement will shift the perceived alignment of the centre of gravity, affecting the results of postural measurement. The central body segments that would be considered in an appropriate alignment when the bob is placed anterior to the lateral malleolus would be considered anterior weight bearing if the bob was placed at the lateral malleolus. Related to subject postural instruction is the placement of feet. Feet placement has an affect on posture^{40,41}. For example, an individual might normally stand with the right foot more anterior and rotated laterally than to the left foot. This individual’s posture may shift from his or her normal posture if the two feet were forced into a predetermined position. To capture the habitual orthostatic posture, a common instruction is to stand in a comfortable, relaxed, or normal stance⁴²⁻⁴⁴. Another approach is more involved and has the individual engage in a number of actions in an attempt to obtain their habitual posture. An example of this is to have the individual march in place, move the head in a circle, roll their shoulders, and bend sideways before adopting their “resting” position³⁹. The establishment and strict adherence to a postural measurement protocol is necessary to control these extrinsic variables. No standard instructions for the positioning of a subject have been established.

1.3 Current status in the literature

Several literature reviews on posture measurements have recently been published^{32,45-50}. Two reviews sought to find and report on the validity and reliability of published quantitative methods that measure posture in general terms^{45,46}. The remaining reviews limited their aims, focusing their attention on either particular methods of measurement^{47,48,50} or particulars of sagittal spinal curvature^{32,49}. Macedo Ribeiro *et al*⁵⁰ reviewed a method for measuring whole body orthostatic posture. However, they looked at one particular method on one particular population; Postural Assessment Software (PAS) on healthy adult females. The remaining reviews included postures other than whole body static standing with some also including the measurement of movement.

Qualitative analyses were conducted in four of the reviews. One used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool⁴⁹ while another used the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) tool⁵⁰. The remaining two reviews qualitatively assessed their included studies with a new critical analysis tool created by Brink *et al*^{32,47}. The number of databases searched in these reviews ranged from as few as 2-3^{45,46,48} to as many as 4-6^{32,47,49,50}. None of the search methods included a search through a chiropractic database such as MANTIS.

The unanimous conclusion of these literature reviews was that more investigation into the validity and reliability of quantitative methods for the measurement of posture is necessary. A number suggested that future methodologies should better limit or remove biases. The quantitative methods of postural measurement include the use of ordinal scales, goniometry (goniometer, inclinometer), measurement of distance (ruler, caliper, measuring tape, flexi-rule), and photography. Photogrammetry showed prominently within these publications with three specifically looking into some version of photogrammetry^{47,48,50}. Two others concluded that photogrammetry showed good potential^{46,49}.

The current systematic review also looked into the validity and reliability of quantitative methods for the measurement of posture. However, its focus was on a conventional method of measurement, that of whole body orthostatic posture. The selection of this particular posture was because its measurement is common practice and because of the interconnected nature of the different body regions. Recall that postural faults are believed to cause, contribute to, or promote MSK pain. Therefore, it is reasonable to measure the whole body posture in an effort to consider potential contributions of postural faults that are distal to the site of pain. In addition to the focus on whole body measurement, this review will include a search through the chiropractic database, MANTIS. The previous literature reviews did not include a chiropractic database. Some relevant information may have been missed as a result of this omission. Another benefit of the MANTIS database is that it includes grey literature. Benzies *et al*, defines grey literature as publically available, open source information that is not available through usual channels such as publications or book sellers⁵¹. Additional grey literature will also be searched in an attempt to capture potentially relevant methods of whole body orthostatic postural measure. The decision to

search in the grey literature was based on the expectation of finding a low volume of evidence and low quality of evidence in the peer-reviewed literature. The complexities associated with whole body postural measurement also contributed to the decision to search the grey literature⁵¹.

The traditional standard set by the visual inspection of the whole body in the orthostatic posture may be suitable for individuals in private practice but it is not appropriate in today's era of evidence-based medicine because it is regarded as a subjective measure, prone to error. Such a measurement tool would not be appropriate for investigating the many suppositions that have been made about posture and its relationship with health and MSK. It is necessary to investigate these hypotheses and their potential to improve the care of individuals with MSK pain conditions. The first step for the rigorous investigation into the measurement of whole body orthostatic posture is to establish a reliable method of measurement. A measurement that can repeatedly measure posture with minimal error. Sources of error need to be identified and managed. Not only should it withstand the rigours of evidence-based practice, it should be affordable, easy to use, and amenable for private practice. Once a reliable tool for postural measurement is established, investigations into matters of validity as well as correlations with certain conditions can continue with greater confidence.

1.4 Aims of the study

The aims of this study was to identify and qualitatively assess quantitative methods of whole body orthostatic postural measurement that; 1) presented with evidence of validity and/or reliability, 2) addressed the management of variables that affect postural measurement, and 3) aligned with conventional postural measurement. Two searches were undertaken. One was a systematic search through published peer reviewed literature for studies that assessed the validity and/or reliability of quantitative whole body orthostatic postural measurements. The second was a search through grey literature for whole body postural instruments used in the private or commercial sectors. The characteristics of the methods found in this second search will be compared with those found in the literature review.

Chapter 2: Methods

2.1 Protocol and registration

Details of the protocol for this systematic review were registered on PROSPERO and can be accessed at: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017055558.

Two searches were undertaken to identify a quantitative method for the measurement of whole body orthostatic posture with the greatest reliability: 1) a systematic literature review of studies that assessed the validity and/or reliability of quantitative methods for measuring whole body orthostatic posture, and 2) a search for relevant postural measures in grey literature.

2.2 Systematic literature review

2.2.1 Data sources and search terms

A systematic search of the scientific literature was conducted across multiple databases to identify research into the reliability and validity of quantitative methods for whole body orthostatic postural measurements. The searched databases included Medline (EBSCO), CINAHL (EBSCO), AMED, Embase (OVID), MANTIS (OVID), SCOPUS, PEDro, and Cochrane Library. Subject headings and keywords included “posture”, “body”, “erect”, “standing”, “measurement”, “analysis”, “assessment”, “clinical assessment tool”, “tool”, “instrument”, “procedure”, “test”, “quantitative”, “reliability”, “accuracy”, and “validity”. An example of the search syntax can be found in Table 1. Hand searches through relevant bibliographies were also conducted.

Table 1. Search syntax example

1	“posture”, OR “body posture”, OR “erect posture”, OR “standing posture”, OR “standing body posture”, OR “whole body posture”, OR “posture alignment”
2	“analysis”, OR “measurement”, OR “assessment”, OR “clinical assessment tool”, OR “determination”, OR “tool”, OR “instrument”, OR “procedure”, OR “test”
3	“reliability”, OR “reliab*”, OR “precision”, OR “quantitative”
4	“validity”, OR “valid*”, OR “accuracy”
5	1 AND 2 AND 3 AND 4
6	5 NOT: “movement”, OR “motion”, OR “range of motion”, OR “animal”

2.2.2 Eligibility criteria

Articles that were eligible for inclusion had to use a quantitative measurement of the whole body orthostatic (still standing) posture and were published in English after 1960. The selection of this period was to capture possible worthy methods of measurements that came after the establishment of the Kendall approach and that may have been overlooked with the passage of time. Studies measuring motion were excluded. Whole body was specifically defined as, at minimum, the assessment of the head, torso and lower extremities. Studies requiring the use of X-ray or animals were also excluded.

2.2.3 Study selection

Two reviewers (CR and SW) searched through the titles and abstracts of the initially identified papers and excluded those that did not meet the selection criteria. Full texts of the remaining studies were collected and screened for selection criteria by the two reviewers. In the instance of disagreement, a third reviewer mediated the decision. Only those articles meeting the selection criteria were included in the study.

2.2.4 Data extraction

The nature of measurement as well as the type of validity or reliability investigated by the included studies was extracted. Population data were extracted and included sample size, how the size was determined, the nature of participant enrollment, age, gender, symptomatic status, and body mass index (BMI). Data from the posture capture protocol were extracted and included: the degree of instruction, placement of arms and feet, the number of views captured, the body regions assessed, whether a qualitative outcome was made and the setting in which the measurement took place. Data in relation to photogrammetry were also extracted and included the number of cameras used, landmarks marked, and measurements made. The nature of the marking of landmarks was extracted as well. Data required for the quality analysis were extracted. This included the reporting of population characteristics, qualification of raters, details about the reference standard, blinding of raters, examination order, timing between use of different measurement instruments, stability of target condition, independence of testing, level of description of tests, explanation of withdrawals, and quality of statistical analysis.

2.2.5 Quality assessment

The Brink and Louw critical analysis tool (CAT) was chosen to evaluate the methodological quality of the included studies⁵² (see Appendix). The authors of this tool combined items from the QUADAS and Quality Appraisal of Diagnostic Reliability Studies (QAREL) tools to create a novel 13 item tool. Their intent was to appraise diagnostic accuracy studies of index tests whose measurements do not result in a particular target condition or diagnosis. The results of the critical analysis were expressed as a percentage score. Studies that aimed to evaluate validity and reliability were calculated with all 13 items. Where the number of items receiving a ‘yes’ score were divided by the total number of related items; 13. Reliability studies were calculated from the nine items related to reliability. The number of reliability related times that received a ‘yes’ score would be divided by the nine; the total number of items related to reliability. Studies with a score greater than 60% will be considered ‘high quality’³². The CAT authors did not establish this criterion nor has it been validated for this particular tool. The decision to use of this criterion was based on precedent⁵³⁻⁵⁶. Two raters independently evaluated the quality of the included studies using the Brink and Louw tool. When disagreement occurred, discussion took place until agreement was reached.

2.3 Grey literature search

A search of the non-research grey literature was performed to identify methods of whole body orthostatic posture available in the private and commercial sectors. The intention was to increase the chance of finding additional and promising whole body orthostatic postural measurement. The methods identified in the private sector would be compared for similarities with those available in the academic literature. The grey literature was limited to web based search for postural measurement websites. Two popular search engines were searched; Google and Bing. These web-based search engines use a complex strategy to improve their search results. This strategy takes into account the frequency and number of links associated with web pages (PageRank), the anchor texts, and the proximity to related work, in order to present the most ‘relevant’ information in the first 10s of results⁵⁷. Because of the comprehensive nature of the search engines, it was not necessary to use numerous synonymous search terms. Moreover, these search engines do not use the Boolean operators. After a trial of different search terms, “body

postural assessment” was the term used for this search. The first 100 ‘hits’ from each search engine were screened for relevance and possible inclusion.

2.3.1 Grey literature screening and data extraction

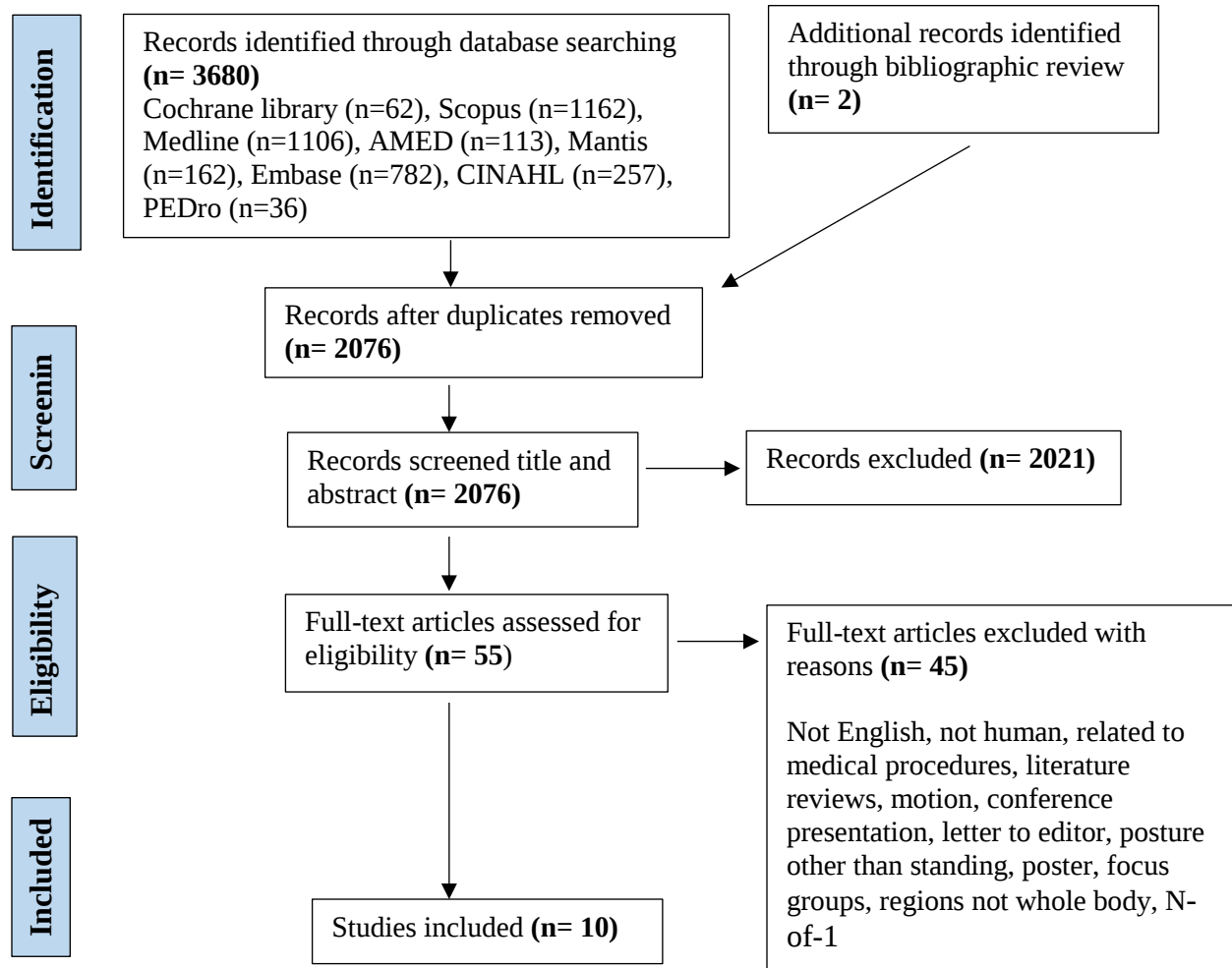
Websites with quantitative methods of whole body orthostatic postural measures were included. Excluded were methods that measured movement, instructional visual assessment, measurement for marketing purposes, postures other than standing, or were peer reviewed research papers. A single reviewer (CR) conducted the search and determined the selection due to time constraints. Data extraction focused on characteristics of external validity and included; type of measurement, the technology required, views assessed, space required, and cost. Any claims of validity and/or reliability were also extracted and evaluated with considerations from the Brink & Louw tool.

Chapter 3: Results

3.1 Systematic search results

From the eight database searched, 3,680 records were retrieved. Two records were identified through backward citation searching. After the removal of duplicates, 2,076 remained. Studies that used animals, x-ray, or assessed motion were excluded as well as those that were not in the English language (2,021 were excluded). Fifty-five full-text articles were then assessed for eligibility. Thirty-eight papers were rejected because they were literature reviews, poster presentations, conference proceedings, pilot studies, or letters to the editor. Others did not assess orthostatic posture, rather they involved medical interventions, motion, postures other than standing, or focus groups. The two reviewers agreed on eight studies but differed on two. A third party evaluated their eligibility. The total number of included studies was ten. Figure 2 presents the PRISMA search results flow diagram.

Figure 2: PRISMA flow diagram describing selection process for included studies



3.1.1 Included studies

All ten studies included in this systematic review used a form of photographic capture to quantify postural measurements. Each study assessed at least one aspect of reliability (intra-rater, inter-rater, test-retest, or measurement error). One of the included studies stated an assessment of accuracy was one of their aims. Table 2, details the studies aims, statistical analyses used, and their conclusions. Each study concluded that the photographic method under study achieved a result of satisfactory reliability. All but one study involved software supported photogrammetry of images captured photographically^{43,44,58-62}, digital video cameras⁶³, or a scanner made of multiple cameras⁶⁴. The non-photogrammetric study used printed photographs to measure posture through a novel method that scored qualitative postural deviations⁶⁵. This method scored ten postural variables using an ordinal scale representing three postural categories; no deviation, moderate deviation, or marked deviation. Images depicting each category were used to guide the raters' scoring.

3.1.2 Photogrammetry protocols

Photogrammetry is the process of making goniometric measurements such as angles and distances taken from a photographic image. An object of known parameters is occasionally included in the image, which serves to calibrate the measurements. Photogrammetry for posture uses anatomical landmarks as points of reference. Distance can be determined by measuring the space between two landmarks or between a representation of the gravitational line and deepest aspect of the lumbar lordosis for example. Angles can be determined by measuring the degrees produced at the intersection of two lines. For example, a line from a representation of the true horizontal intersecting a line formed by connecting two points on unleveled shoulders.

Anatomical landmarks are usually boney structures such as the lateral malleolus or the anterior superior iliac spine (ASIS). On occasion, soft structures like the joint space of the knee or the navel are used. All studies that assessed a form of photogrammetry involved the marking of landmarks with stickers, pegs, or Styrofoam balls. The number of landmarks marked as well as the number and type of measurements taken varied between studies. The number of landmark ranged from 5 to 32. The range of measurements taken from the photographs was from 5 to 29. The types of measurements include distances and angles from a horizontal and vertical reference line. These all involved different regions of the body depending on the study parameters. The

Newton and Neal study⁶³ took measurements of the upper extremities. These were in addition to measurements that were similar to the other studies; those of the head, neck, torso, and lower extremity. No studies measured muscles nor did they comment on them. The photogrammetric studies used a single individual, such as a research assistant, who was not a rater, for the placement of markers on the selected landmarks. The exception was Normand *et al*⁶¹ who employed a methodology where each rater placed the required markers on each subject and on both occasions. The results of the photogrammetry parameters and the software systems used to process the measurements are found in Table 3.

3.1.3 Ordinal scale protocol

The authors of the ordinal scale study⁶⁵ reviewed the photographic images taken of 114 boys. Each subject had four postural views captured; anterior, lateral, oblique, and posterior. The authors selected photographs of subjects that represented the level of each possible score; good posture, moderate deviation, and marked deviation. Each of these categories was assigned the number 5, 3, or 1 respectively. These representative images were then used to score the posture of the 114 subjects. No landmarks were marked prior to the image capture in this study.

Table 2. Included articles in alphabetical order, their aims, statistical methods, and conclusions

Author/Year	Aims	Statistics	Conclusions
Ferreira <i>et al</i> 2010	Assess the accuracy of software inter- & intra-rater reliability	ICC _{2,1} : 86% of were acceptable or higher ICC _{3,1} : Averaged acceptable or higher 81%	PAS is accurate tool. Good inter- & intra reliabilities
Hazar <i>et al</i> 2015	Investigate the inter- & intra-rater reliability of photographs & MB-Ruler	ICC _{2,1} : All acceptable or higher ICC _{3,1} : All Excellent	Photographic method with software reliable and repeatable
Hough <i>et al</i> 2013	To establish the intra-rater reliability of the Postural Analysis Toolkit (PAT)	Spearman's correlation: 81% had good correlation	PAT proved to be reliable for standing postural alignment
Iunes <i>et al</i> 2009	Verify agreement between visual raters, photo raters, and visual & photo raters	Cramer's V coefficient: 76% with high level of agreement	Better agreement in photo, poor in others comparisons
McEvoy <i>et al</i> 2005	To assess the variability of posture using repeated measures	ICC _{1,3} (CI): 100% excellent	Posture didn't change significantly on repeated measure
Newton <i>et al</i> 1994	Assess a number of posture parameters for suitability for quantification	Pearson correlation coefficient Intra-day Mean $r=0.77$. Inter-day $r=0.56$	Was accurate & reliable in detecting small changes
Normand <i>et al</i> 2007	Examine the intra- and inter-examiner reliability of software process	ICC _{2,1} : 45% were acceptable or higher ICC _{3,1} : All good or excellent	Highly reliable, SEM small, observer error small
Pausic <i>et al</i> 2010	To examine inter-item reliability of two measurement software programs	ICC (A,3)(CI) All good or excellent ICC (A,1)(CI) All good or excellent SEm	Satisfactory interitem reliability of photographic assessment. Slight fluctuations affecting SEM
Tomkinson <i>et al</i> 2013	1) Quantify the repeatability of direct measurements of standing posture with 3D scanning 2) quantify the magnitude of postural and technical errors	Intra-rater test mean, retest mean, absolute change, effect size, chance of change in mean, typical error, typical error/SD _{pooled} , limits of agreement 73% with good statistical results	Most postural measurement had good repeatability except head/neck, poor Low technical error, good precision.
Watson <i>et al</i> 2000	1) Describe assessment criteria for 10 aspects of posture (categories) 2) develop quantitative posture rating scale 3) establish reliability of assessment	Descriptive % with CI	Reliable and suitable for clinical use

Table 3. Photogrammetry data

Author/ Year	Camera # / type	Landmark #/ marker	# measures/ other regions	Marked by rater	Software
Ferreira 2010	2 digital cameras	32/ 1 marker	29, none	No	Postural assessment software
Hazar 2015	1 digital camera	7/ 1 marker	5, none	No	MB-Ruler
Hough 2013	1 digital camera	11/ 1 marker/ SAME photos	11, none	No	Postural analysis Toolkit
Iunes 2009	1 digital camera	21/ 1 marker	14, none	No	ALCimagen-2000 Manipulating Images 1.5
McEvoy 2005	1 digital camera	5/ not clear / marks left on	5, none	No	ImageTool UTHSCA version 2
Newton 1994	3 video cameras	24/ not clear	17, upper extremities	No	Matrox PIP1024B image processing board fitted in IBM computer
Normand 2007	1 digital camera	29/ 3 raters	15, none	Yes	PosturePrint®
Pausic 2010	1 digital camera	10/ 1 marker/ marks left on	9, none	No	Posture Image Analyzer & ImageTool UTHSCA
Tomkinson 2013	Body scanner	18/ 1 marker	11, none	No	DigiSize v2.3 ²

3.1.4 Posture capture protocol

The person's 'typical' posture is the interest of many health care professionals. Once the landmarks have been marked, the subject is positioned for image capture. The degree or amount of instruction given prior to image capture varied from a low to a high level of instruction. The majority provided a low degree of instruction. An example of a low degree of instruction is to ask participants to adopt a standing posture that was "comfortable" or "usual" for them^{43,44,60,62,63}. This comfortable standing position was standardised by aligning the subjects' feet to a line taped on the floor. Three studies provided a high degree of instruction or action. Tomkinson *et al*⁶⁴ and Normand *et al*⁶¹ asked their subjects to perform activities which are believed to position a person in a more habitual or typical posture. Activities included marching in place as a means to identify normal foot position and determining normal head position by bending the head and neck forward and backward before returning to a comfortable position. Watson *et al*⁶⁵ showed their participants an image of good posture and asked that they stand in a posture as similar as possible to that in the image. These participants were also asked to place their feet and knees such that they were touching and facing forward. All studies that reported posture instructions asked for a particular placement of the arms. These positions were either the arms in a neutral position, across the chest, or with the elbows bent to 90 degrees of flexion. Two studies did not clearly report on postural instructions^{58,59}.

Depending on the number of cameras involved in the measurement protocol, the postural instructions would be repeated for the capture of the additional views. As stated previously, convention calls for a minimum of four postural views. Two of the reported protocols assessed the conventional number of four postural views^{58,65}. However, of the two, Watson and Mac Donncha did not follow the convention of capturing each of the lateral views, but instead captured one lateral and one oblique view⁶⁵. This view was unique when compared with all other included studies. Three studies assessed three views^{44,61,63}. Newton and Neal⁶³, took a superior view rather than the more typical anterior view. Normand *et al*⁶¹, captured two lateral views foregoing the posterior view. The remaining protocols assessed one or two views. When two views were captured, they were of the anterior and lateral views. The lateral view was the priority when only one view was captured.

The space used for the posture capture was categorised as ‘lab’ if it was a large space or ‘clinic’ if the space used reflected a treatment room in a private practice. One study captured the images of their participants in a small space that was clinic-like⁴³. One study’s report on the space used to capture their participants’ images was unclear⁶⁴. The remaining studies all used a large space. The postural protocol results can be found in Table 4. This table also displays those studies that provided qualitative outcomes resulting from the postural measurements^{44,65}.

Table 4. Posture capture protocol data

Author/ Year	Degree of instruction	Feet/arm position	Views	Qualitative outcome	Setting type
Ferreira 2010	unclear	no/no	A, P, 2L	no	lab
Hazar 2015	low	no/yes	L	no	clinic
Hough 2013	unclear	no/no	A, L	no	lab
Iunes 2009	low	no/yes	A, P, L	yes	lab
McEvoy 2005	low	no/yes	L	no	lab
Newton 1994	low	no/yes	P, L, S	no	lab
Normand 2007	high	no/yes	A, 2L	no	lab
Pausic 2010	low	no/yes	A, L	no	lab
Tomkinson 2013	high	yes/yes	A, L	no	unclear
Watson 2000	high	yes/yes	A, P, L, O	yes	lab

Legend: A, anterior; P, posterior; L, lateral; S, superior; O, oblique

3.1.5 Population characteristics and enrolment method

The population characteristics can be seen in Table 5. Hough and Nel enrolled a single subject to examine the intra-rater reliability between 15 raters⁵⁹. The gender and age of this subject was unclear. Due to the methodological nature of this study, its single subject was excluded from the subsequent discussion on population. This study, however, was one of three that determined the number of subjects to enrol either through a power analysis or through citing other references. Pausic *et al*⁶² performed a power calculation and through a method of convenience enrolled 273 healthy boys between the ages of 10 and 13 years. Ferreira *et al*⁵⁹ cited previous literature to determine their enrolment number of 22. The characteristics of their 22 participants were not reported or were unclear. One study reported to enrol their subjects through a process of randomisation. Watson and Mac Donncha⁶⁵ reported a random enrollment of 114 healthy males

between the ages of 15-17 years but did not explain how the randomisation occurred. Two studies used a convenience method to enrol participants of both genders with a mixed status of either healthy or symptomatic. McEvoy and Grimmer examined 38 children between the ages of 5 to 12 years⁶⁰ while Normand *et al*⁶¹ enrolled 40 individuals with a mean age of 24.4 years. Tomkinson *et al*⁶⁴ used a sampling of convenience but was unclear with reporting gender. They enrolled 52 participants with a mixed status (healthy or symptomatic) and a mean age of 35 years. The remaining three studies did not report their method of enrolment but all reportedly enrolled healthy subjects. Hazar *et al*⁴³ enrolled 30 teenagers of mixed gender that were 16 and 17 years old. This study also reported on body mass index (BMI). Iunes *et al* evaluated 21 young healthy male and female adults who were between the ages of 22 and 26 years⁴⁴. Newton and Neal selected 20 healthy males with a mean age of 20.2 years⁶³. In summary, the sample sizes ranged from 20 to 273. The ages of the studied population ranged from early childhood to middle age. A healthy population was reportedly used in 60% of the studies. A population of mixed gender was reportedly used in 40% of the studies.

Table 5. Population characteristics data

Author/ Year	Sample size/ Power	Enrolment	Age/gender	Status	BMI
Ferreira 2010	22 / lit based	unclear	unclear / unclear	unclear	-
Hazar 2015	30 / no	unclear	both/16-17	healthy	y
Hough 2013	1 / for raters	unclear	unclear /18-30	healthy	-
Iunes 2009	21 / no	unclear	both/22-26	healthy	-
McEvoy 2005	38 / no	convenient	both/5-12	mix	-
Newton 1994	20 / no	unclear	male/ 20.2 mean	healthy	-
Normand 2007	40 / no	convenient	both/24.4 (1.9 SD)	mix	-
Pausic 2010	273 / yes	convenient	male/10-13	healthy	-
Tomkinson 2013	52 / no	convenient	unclear /35 (12 SD)	mix	-
Watson 2000	114 / no	random	male/15-17	healthy	-

3.2 Qualitative analysis

Prior to assessing the ten studies with the CAT, the raters carefully discussed each of the thirteen items within the tool. Explanations of what constituted a 'yes' score were agreed upon. The raters then independently scored each study. In spite of the raters' best efforts to interpret the scoring protocol, 21 out of the 130 possible scores were not similarly scored. The proportion of disagreement was 16%. Each of the 21 criteria were mutually evaluated and discussed until an agreed score could be determined.

Results from the critical analysis tool used to qualitatively evaluate the ten included studies can be found in Table 6. Studies with scores above 60% were considered to be of high quality. Studies assessing reliability would require six 'yes' responses out of the nine reliability items to be considered high quality. Validity and reliability studies would require eight of the possible thirteen items. The average of the quality score was 56% (range 38-67%). The only study that aimed to measure validity and reliability scored 38% and therefore, was considered to be of poor quality⁵⁸. Four studies of reliability scored at 67% (6 'yes' responses of 9 reliability items) and were therefore considered to be of high quality^{43,61,62,64}. The remaining five reliability studies received scores of either 44% or 56% and were considered to be of poor quality. The majority (8 of the 10) papers were deemed to have used satisfactory methods of statistical analyses (item 13 in the CAT). Intraclass correlation coefficient (ICC) calculations were used to evaluate the results from the repeated measurements. It is preferred to present the ICC results with a 95% confidence interval (CI). However, ICC results without CI were judged as reasonable. Two of the three studies that evaluated measurement error appropriately calculated the standard error of measurement (SEm). The reason two of the studies did not obtain this approval was the reliance of descriptive analyses to arrive at their conclusion of reliability.

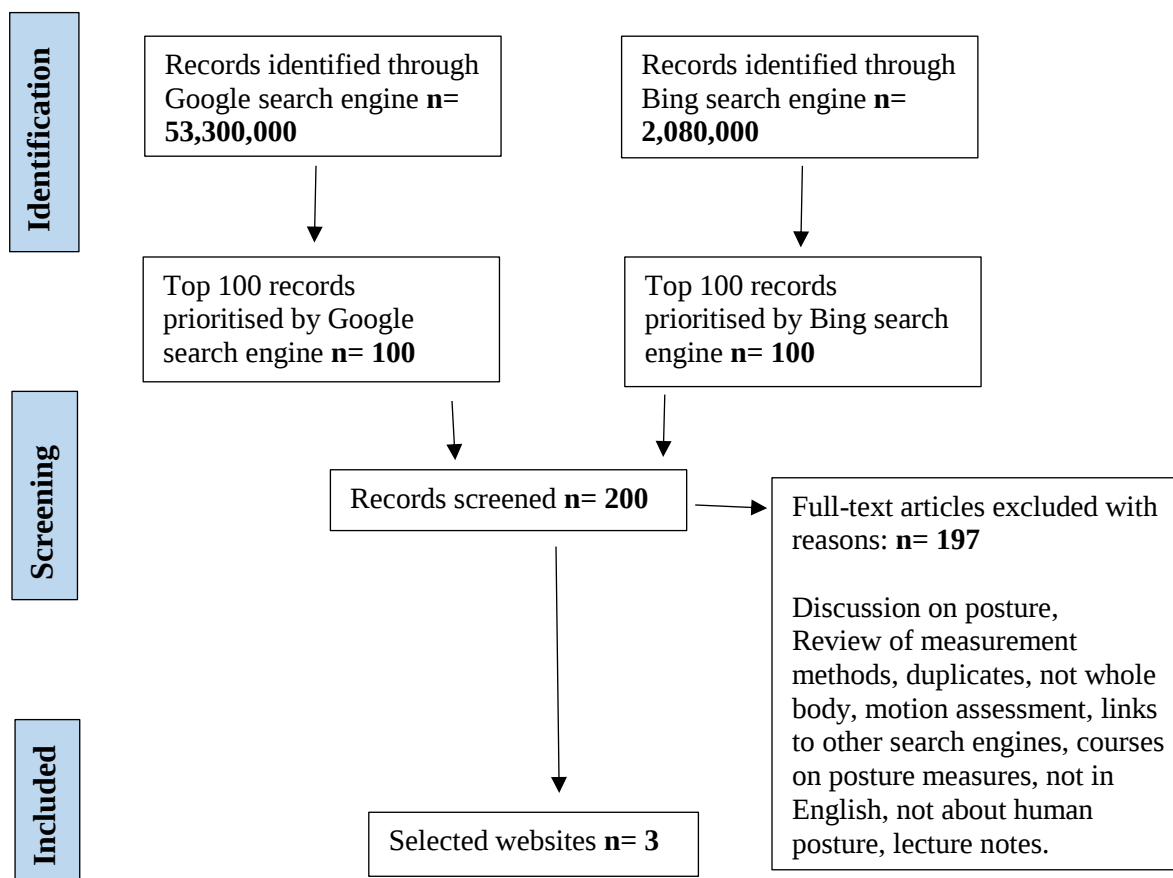
Table 6. Methodological quality evaluation using the Brink & Louw critical analysis tool

Author / Year	1	2	3	4	5	6	7	8	9	10	11	12	13	High quality >60%
Ferreira 2010	x	✓	n/a	x	x	✓	n/a	✓	n/a	✓	n/a	x	✓	5/13 = 38%
Hazar 2015	✓	x	n/a	✓	x	x	n/a	✓	n/a	✓	n/a	✓	✓	6/9 = 67%
Hough 2013	x	✓	n/a	✓	x	x	n/a	✓	n/a	✓	n/a	✓	x	5/9 = 56%
Lunes 2009	✓	✓	✓	x	n/a	x	x	n/a	✓	✓	✓	✓	✓	5/9 = 56%
McEvoy 2005	✓	x	n/a	n/a	x	x	n/a	✓	n/a	✓	n/a	✓	✓	5/9 = 56%
Newton 1994	x	x	n/a	n/a	x	x	n/a	✓	n/a	✓	n/a	✓	x	4/9 = 44%
Normand 2007	✓	✓	n/a	✓	✓	✓	n/a	✓	n/a	✓	n/a	✓	✓	9/9 = 100%
Pausic 2010	✓	✓	n/a	n/a	x	x	n/a	✓	n/a	✓	n/a	✓	✓	6/9 = 67%
Tomkinson 2013	✓	✓	n/a	n/a	x	x	n/a	✓	n/a	✓	n/a	✓	✓	6/9 = 67%
Watson 2000	x	✓	n/a	x	x	x	n/a	✓	n/a	✓	n/a	✓	x	4/9 = 44%

3.3 Grey literature

The Google search engine was utilised on the 25th of November 2017. The total number of retrieved responses was 53.3 million. As indicated earlier, the first 100 hits were investigated. Forty-six links out of the first 100 were opened and investigated for inclusion criteria and this determination was based on the link title and the description. Forty-four of the 46 were excluded. Two websites were included in this study: PostureScreen®⁶⁶ and PostureZone©⁶⁷. The search engine Bing was also searched on the same date. The number of retrieved responses was 2.08 million. Nineteen links were opened for evaluation of inclusion criteria. Chinesport⁶⁸ was selected (see Figure 3).

Figure 3. Search engine search results



All relevant products from each included site measured whole body orthostatic posture using a photogrammetric software program. PostureScreen® and PostureZone© offer photogrammetric software programs that take measurements from digital images. The cost of PostureScreen® starts at US\$19.99 and PostureZone© costs US\$399.00. Both can measure up to four views (anterior, posterior, left and right lateral). Chinesport offered a photogrammetric software along with all additional component such as cameras, force plates, viewing mirrors, plumbelines, and computers. The Chinesport system offered a fifth view, which was from the top looking down. The cost was not provided and although the site claimed reliability, there was no evidence for its repeatability claim. PostureScreen® provided a link to a published article about the application's intra- and inter-rater agreement⁶⁹. This article scored 56% on the critical analysis tool. A number of reporting criteria were lacking for its reported methodology: intra-rater blinding, order of subject examination, and a clear description of the measurement protocol.

Chapter 4: Discussion

4.1 Results summary

The aim of this paper was to identify a method of whole body postural measurement that was valid and reliable. The method should manage variables that affect the measurement of posture as well as be aligned with conventional practices. Such a method of postural measurement would be beneficial for the future investigations involving posture. The systematic literature review found ten studies of reliability for whole body postural measures, one of which also aimed to assess validity. All methods involved the use of photography. A search of grey literature found three methods of orthostatic whole body measurement, which also utilised photography. Twelve of the thirteen methods involved photogrammetry. The one non-photogrammetric study measured posture with a novel ordinal scale that quantified ten postural deviations.

To help determine which method of postural measurement presented with the most potential for future investigative use, a critical analysis tool was used to assess the methodological and reporting quality of the studies included from the peer-reviewed literature search. Additionally, an evaluation of the management of numerous variables affecting posture was also considered. These postural variables have the potential to negatively affect measurement reliability if not

controlled. A quantitative method of postural measurement that reflects conventional practice may improve the generalizability and translation of the method into clinical practice. Therefore, the studies' measurement protocols were compared with those of conventional practice.

4.2 Discussion on validity and reliability

Diagnostic tests should accurately identify individuals with a condition of interest. These tests should repeatedly differentiate between individuals with or without the condition. Studies of diagnostic accuracy should determine the evidence of these qualities and appropriately report the findings. Inappropriate reporting may lead to over interpretation or misreporting. In an attempt to assist researchers to improve their reporting, the Standards for Reporting of Diagnostic Accuracy (STARD) initiative was published in 2003²⁵ and updated in 2015³¹. The members of the STARD committee developed a thirty item checklist that covered matters related to methodological and reporting quality. Examples of such matters include study design, study participants, and test methods. The reporting of quality helps to prevent 'spin' or the misreporting of results. Unfortunately, 'spin' continues to be frequently published in high impact journals⁷⁰. The assessment of study quality has become an important component of systematic reviews.

A number of quality checklists, in addition to that of STARD, have been developed to aid the analysis. This literature review did not use the STARD checklist because of its strong demands for a target condition and reference standard. Instead, a critical analyse tool (CAT) created by Brink and Louw was used because of its softer approach toward the target condition⁵². The role of the reference standard remains an important component of the CAT but mostly for studies investigating validity. The Brink and Louw tool has a 13 item checklist with three possible responses; yes, no, or not applicable (n/a). Each of these items relate to concerns of either validity, reliability, or both. An advantage of the CAT was the expectation that a reliability study would have an 'n/a' in each of the four items related to validity and that this did not affect the measure of the study's quality.

4.3 Discussion of CAT results

Four of the ten included studies received a CAT score of 67% and are considered to be high quality according to the Cohen *et al*³² approach to interpreting the CAT results. The remaining six studies did not meet the greater than 60% threshold and were considered poorer in quality. Regardless of the overall score, there were common strengths and weaknesses regarding the methodologies and reporting within all of these studies. There were three criteria incorporated in the tool where the studies performed well according to the raters' agreed criteria. The first was the degree to which the authors described the execution of the index test. In this case, the description of how the measurement of posture was performed. All studies described how the postural measurement was performed in enough detail as to allow its replication. Secondly, all but one study accounted for or explained subject withdrawals from their studies. More accurately, it was agreed by the raters using the CAT that if no withdrawals occurred, a 'yes' would be indicated for this criteria. In actuality, fewer studies would have received a 'yes' rating if it were based on the reporting of withdrawals. This outcome may be related to the generally low subject numbers, which may have made it easier to keep track of and manage subjects as compared to studies with larger subject numbers. The third criterion relates to the stability or theoretical stability of the measurement of the posture. Although spinal posture has been determined to be stable⁷¹, it is not well established for the measurement of whole body orthostatic posture. The raters agreed that if a repeat measurement occurred within one week or less it would be considered a 'yes' for this item in the tool. This was regardless if the authors discussed their rationale or not.

The areas of weakness within the studies, as illustrated by the results of the CAT, were intra-rater blinding and whether the study varied the assessment order of the subjects. Intra-rater blinding is generally a challenge when assessing the reliability of postural measures. Some methods include measurements of the face, which eliminates the option of blocking out recognisable features. Moreover, the posture of individuals are recognisable. Most studies included in this review did not report on methods for intra-rater blinding. Varying the order in which the subjects are reassessed is a viable method of reducing the potential bias that results from remembering prior outcomes when repeating a measurement. Only two studies reported varying the order of the subjects' repeated measurement.

An interesting observation in relation to the assessment of quality is that four of the studies that were deemed to be high quality using the CAT were conducted after the first STARD publication. These studies may have benefited from this publication, although, none referred to the STARD initiative. The determination of the STARD initiative's influence on reporting diagnostic accuracy for methods of whole body orthostatic postural measurement is difficult. There are few study numbers and of those studies conducted after its publication, there are equal numbers with high quality as there are with poor quality.

4.3.1 Quality assessment challenges

Evaluating the quality of studies investigating the validity and/or reliability of measuring whole body orthostatic posture is challenging. This is not only in relation to the absence of an appropriate clinical reference standard or a clear target condition. There is also a lack of standardised and appropriate reporting in spite of the efforts of the STARD initiative. These challenges generated difficulties in using the CAT. The difficulties came in two forms. One was related to the lack of consistency in reporting on the part of the study authors and the second came from the structure of the CAT itself.

An aspect of the CAT structure that was a source of confusion and required clarification from the raters was that, although its purpose is to evaluate quality, some of the thirteen items in its checklist relate to the quality of reporting and not necessarily about the study's methodological quality. Raters had an initial tendency to rate based on methodological quality, and after further discussion, had to reassess the way they were evaluating the studies. Another challenge that resulted from the structure of the CAT was the multiple components in a single criterion or item. An example of this is in the CAT's item 1 (see Table 2). This item relates to the study's population. Three components are described in this item; the nature of recruitment, characteristics of the subjects, and selection criteria. The raters had to negotiate the rating terms of this item because not all studies reported on all three components. Most studies reported on one or two.

The lack of consistent use of terms and reporting methods in the included studies presented an additional challenge when using the CAT. In relation to reporting, one would expect to find the

majority, if not all, information about a study's methodology in their 'methods' section. This was not the case in many of the studies evaluated within this literature review. The information required to respond to a number of item criteria pertaining to methods in the CAT was often found scattered throughout the articles. Relevant information could be found in the 'results' or 'discussion' sections of the paper. The different uses of terms in the studies presented yet another challenge for the raters. The term 'accuracy' is an example that demonstrates this challenge. As previously defined, the determination of accuracy relies on a comparison of results with a reference standard. It is a term used to describe validity in this regard. The study that claimed to assess the validity and reliability of a whole body orthostatic postural measure did not use any comparison of results, let alone results from a reference standard. As it turned out, this study did not evaluate validity, rather, it only assessed aspects of reliability.

The term 'reliability' presented its own challenges. Recognising the confusion that a lack of consistent use of terminology causes when reporting research results and the potential impact that this confusion has on the translation of research into clinical practice, the COSMIN group aimed to align terminology²⁷. In their attempt to define reliability, however, they used the term 'reliability' as both a domain name as well as a term to describe a property of measurement. The definition of reliability as a domain name is "the degree to which the measurement is free from error". As a property of measurement, reliability is defined as "the proportion of the total variance in the measurement which is due to the true difference between patients²⁷". Both definitions relate to measurement error. People are not necessarily required for the determination of measurement error as defined by the domain name. However, people are necessary for the determination of reliability as defined by a property of measurement.

The COSMIN group defined the essence of reliability in terms of measurement error. However, reliability is often determined by a test's ability to produce similar results on a stable individual in repeated measures. It should also be able to differentiate between individuals with or without a target condition. Reliability can be evaluated through a number of repeated measurement methods; intra-rater, inter-rater, or test-retest. It has been previously stated that all studies included in this literature review evaluated the reliability of a particular method of postural measurement. They did not evaluate the same aspects of reliability though. Some evaluated

intra-rater reliability. Others evaluated inter-rater reliability. Regardless of which form of reliability was evaluated, the CAT was applied uniformly to all studies. The thorough CAT item discussions and debate resulted in a proportion of disagreement between raters of 16 percent. One hundred percent agreement was reached through further discussion of each item where disagreement occurred.

4.3.2 Alternative outcomes

To illustrate the challenge and affect terminology has on the assessment of quality using the CAT score, outcomes of two studies will be discussed. One study overtly stated validation as one of their aims, while the other study could have evaluated validity but did not state this as an aim. The title of the study by Ferreira *et al*⁵⁸ stated an aim for validation and reliability. The stated objective within their paper was the determination of accuracy and reliability. The CAT was applied to this paper in accordance with their stated aims, which resulted in a score of 38% (5 ‘yes’ responses out of 13 items). When scored as a reliability study, their results would be 56%. The main issue with this paper was a lack of reporting relevant information. Regardless, this is a demonstration of how an appropriate use of terminology can complicate the evaluation of quality. The paper by Iunes *et al*⁴⁴ had a contrary situation. They stated a desire to assess levels of agreement between different methods of postural measurement, one of which was visual measurement. However, the stated statistical analysis used in this study was reliability (inter-rater). As a reliability study, the CAT result was poor with 56% (5 ‘yes’ responses out of 9 items). If, on the other hand, they had statistically compared the two methods as a means to evaluate validity, their score would be high quality with 62% (8 of 13 times with a ‘yes’ response). Validity was not the stated aim of this paper and therefore the quality of the paper’s reliability was reported.

4.4 Statistical analyses

Each study in this literature review presented statistical results that supported their claims of good reliability. The studies that calculated ICC statistics, presented results that were considered acceptable to excellent an average of 92% of the time (ranging from 76% to 100%). These include results from the evaluation of intra-rater, inter-rater, and test-retest reliability. Three of the four studies that were deemed by the CAT results to be of high quality used ICC analyses to

evaluate their measurements. One hundred percent of these analyses resulted in acceptable or higher levels of agreement. However, when Normand *et al*⁶¹ analysed reliability with a conservative ICC method, as opposed to a liberal method, the results were not as favorable. The conservative method resulted in an acceptable or better level of agreement in 45% of the measurements that were calculated. Inter-rater reliability was poor with only two of eleven measurements reaching the acceptable threshold. Tomkinson *et al*⁶⁴ used results from systematic and random error calculations to determine their reliability and showed that 8 out of 11 (73%) of their measurements had good repeatability. These outcomes indicate that the measurement of whole body orthostatic posture through photogrammetric means has promise. However, the ensuing discussion will show cause for caution regarding the authors' claims.

4.5 Variables

Managing the variables that can affect posture is an important endeavour in the establishment of a reliable quantitative method for the measurement of whole body orthostatic posture. If left unmanaged, perceived differences in postural measurements may not reflect true change but rather an error in measurement. The variables affecting posture can be from intrinsic or extrinsic factors. The discussion of variable management is presented in relation to the protocols used in acquiring the image to be measured (image capture) and the protocols used for the measurements themselves (ordinal scale or photogrammetry).

4.5.1 Image capture

All of the studies included in this literature review use photography as a central component to their methods of postural measurement. This then requires a person to be placed in a quiet standing position in preparation for the photograph to be taken. This positioning process is true for most methods of whole body orthostatic postural measurement and is a potential source of variance and error. The instructions provided to the subjects influence the stance they adopt. Convention calls for a neutral or habitual stance. Required of this is to stand on both feet with the body weight equally distributed. Janda suggested that individual's actual habitual stance is not with the weight on both feet but on one foot. He proposes that the posture of standing on one foot is more representative of the individual's function⁷². Janda's suggestion is an illustration that standing on two feet with the body weight equally distributed between them may not actually be

a habitual or even a comfortable posture and that a person needs to be instructed into the position in which he or she will be measured.

Of the authors that clearly reported on a degree of instruction provided to their participants, the majority gave a minimal degree of instruction. An example of this type of instruction was to ask a person to stand in a comfortable position while standing on both feet. A concern with such a low degree of instruction is the transient nature of the perception of 'comfortable'. As one becomes more comfortable with the measurement process, one can become more relaxed. Their 'comfortable' position could change as a result. Three studies provided their participants with a high degree of instruction. Two studies had their subjects run through active manoeuvres to help them identify their most comfortable posture. These movements would help a person find their natural and hypothetically truer posture. Contrary to conventional practice, the third study that provided a high degree of instruction did not attempt to find a habitual posture. Instead, a picture of good posture was shown to the participants and then were asked to stand in as similar a position as that shown in the photograph. These researchers also asked that their subjects stand with the feet parallel to and touching each other. Their knees were also to be facing forward and touching. The purpose of this particular position was to aid in the reporting of knee posture. Feet unable to touch could indicate a certain degree of genu valgus.

In keeping with the convention of measuring a natural posture, two studies asked subjects to stand with their arms resting comfortably at their sides. The remaining studies that reported on the details of postural positioning asked their subjects to place their arms in a particular position; one other than natural. In these cases, the subjects were asked to place their arms either across their chests or at their sides but bent to 90° at the elbows. The purpose of such arm positions was to avoid obscuring the pelvis. However, undesirable alterations to the natural posture may result from this modification. A person may lean backward to compensate for the centre of gravity's forward shift. The weight bearing alignment could be affected by this shift. Additionally, muscles of the shoulder girdle would contract to stabilise the arms in this raised position. This muscle activation could effect a change in the position of the head, neck, and shoulders. Moreover, the scapulae may shift into a 'winging' position in response to the weight of the forearms or the crossing of the arms. In regards to the feet, only one study standardised a natural

foot position by tracing the outline of their subjects' feet after having had them march in place for a few steps. The subjects used these foot tracings when being repositioned for subsequent measurements. This marching in place is believed to place a person's feet in a natural position. However, this process may position the feet closer than a person would naturally stand on two feet. A wider stance is a more stable and comfortable stance. The remaining studies merely asked subjects to place their feet in line with a tape on the floor. There is no standard for the degree or type of instruction to provide a person when positioning them for the measurement of their orthostatic posture.

4.5.2 Ordinal score

The study that applied an ordinal scale to a photographic method of postural measurement quantified ten commonly assessed postural faults. Examples of the faults relate to head position, spinal curves, carriage of the centre of gravity, and lower extremity alignment. The newly developed scale provided a systematic means for ranking postural findings, which the authors conclude produced reliable methods for postural measurement. This method may have an ability to better measure change as compared to the conventional method. However, the statistical analyses used to determine this method's reliability were descriptive and did not provide confidence in their conclusions. In addition to the statistical methodology of this study, the conclusions drawn by Tyson's review of the literature casts further doubt on ordinal postural measures. She stated that they were unreliable⁴⁵. Although there are some concerns about the level of reliability when using an ordinal scale, it appears to be a possible improvement over the conventional visual method of measurement on face value.

4.5.3 Photogrammetry

The remaining studies used the subjects' photographed image to take goniometric measurements. This process is referred to as photogrammetry. The essence of the photogrammetric method of analysis is the capture of an individual's digital image while in a still standing posture and then the uploading of the image into a software program. The photogrammetric program takes goniometric measurements from the image either with assistance from a rater or automatically as part of its program. The essential equipment for photogrammetry is a tripod, camera, and a

photogrammetric software program. Additional equipment could include a plumbline, an object of known parameters, additional cameras, or some type of grid behind the subject.

Seven studies used a single camera. Of these, two photographed a single view of posture. The repositioning of subjects was required when measurements from different views were needed. One study used two cameras in their protocol. The second camera was set at 90° from the first. This additional camera increased the space requirement for the image capture protocol. However, the use of two cameras halved the time spent when capturing four postural views, making the posture capturing process more convenient for the patient and practitioner. Two studies presented a method of image capture that was termed ‘three dimensional’. A 3D scanner comprised of multiple cameras was used for one study. Interestingly, this 3D scanning method only presented measurements that resulted from two postural views; anterior and lateral. A system with three video cameras was used for the other method of 3D analysis. The measurements taken from the images captured by the three video cameras came from three views; posterior, lateral, and superior. The superior view is not conventional but reveals information about the body’s placement around the vertical axis. Other than the vertical view, the advantage of these two 3D methods was not clear. Their equipment and space requirements may render them less conducive to private practitioners.

Common among the photogrammetric studies was the placement of markers on anatomical landmarks. Reflective tape, coloured stickers, wooden pegs, and Styrofoam balls are examples of the different types of marker used within these studies. Markers were used as reference points that facilitate the taking of measurements from the photographic images. The marking of landmarks is another potential source of variance and error. As the number of landmarks increase, so too does the potential of error. The consistent placement of these markers can be challenging as explained by do Rosario⁴⁸. He found, for example, that a marker placed on smaller landmarks such as the anterior superior iliac spine was more consistent than when placing markers on larger landmarks such as a patella. Additionally, more consistency occurred when marking lean individuals. His statement about lean individuals raises a question regarding the ability to measure the posture of an obese patient through photogrammetric means. None of the studies within this review referred to obesity nor was it discussed in any inclusion or

exclusion criteria. The marking of anatomical landmarks is relevant to the number and type of postural measurements taken from the images. The results in this review indicate a correlation between the number of landmarks and the number of measurements. As the number of landmarks increased, so too did the number of measurements. Two studies took five measurements of whole body orthostatic posture from as few as five landmarks. On the other end of the spectrum, one study took 29 measurements from 32 landmarks. There is no standard for the number of landmarks and the correct number of measurements has not been determined.

Marker training and experience are important for the minimisation of the potential error that results from the marking of landmarks. The use of a well trained and experienced marker can improve measurement standardisation and consistency. Eight studies used a signal marker, one who was not a rater, to mark the landmarks on all participants. This person also positioned the subject for image capture and took their photograph. The experience of the marker was not always clear. The authors focused their reporting on experience on the raters, the individuals processing the images for measurement, rather than on the marker. The use of a single marker for all subjects reduced or eliminated the effect of the landmark confounder on these studies' results. Therefore, the interpretability of their results is limited. There was an exception where one study had each rater mark the necessary landmarks. The three raters in this study independently marked each of the subjects for each session of image capture.

4.5.4 Precision

To meaningfully assess the reliability of a postural measure, one must simulate clinical practice and assess the entire procedure. The studies that used a single individual to mark and position the subjects reduced the interpretation of reliability. The placement of landmarks and the setting up of the participant for image capture were significant sources of variation that were eliminated. It can therefore be said that most studies assessed only one part of the entire procedure, and their conclusions are confined to that section of the protocol alone.

The section of the measurement method that was essentially assessed was related to the downloading of the posture images into the photogrammetric software and the subsequent identification of landmarks or their corresponding markers. An example method of landmark

identification is to mouse click on the relevant landmarks. At this point, the software program would calculate the postural measurements. Some software programs self-identify the landmarks and perform the calculations without the assistance of the rater. As it turns out, the majority of these studies were determining the degree to which the raters (intra- or inter-) using the photogrammetric software programs could repeatedly achieve similar measurements. In other words, these studies were measuring the protocols' precision. The Pausic *et al* paper⁶² had the largest sample size of the photogrammetric studies. Their subject number of 273 was determined by a power calculation. The results of their study provides us with a stronger estimation of precision.

4.6 Alignment with convention: generalisability

Whole body orthostatic postural measurement includes the analysis of the head, spine, torso, pelvis, and the extremities. Spinal curves both in the sagittal and coronal planes are a component of the analysis as is the alignment of the body to the centre of gravity. Muscle evaluation is another component of postural measurement. Kendall¹⁰ further explains that to fully measure posture, the subject should be minimally dressed, stand in a habitual posture and be viewed from four directions; posterior, anterior, and from the right and left lateral. The definition of 'whole body' was relaxed for the inclusion criteria set in this literature review to the minimum analysis of the head, torso and lower extremity. The relaxation of this definition was due to the anticipation of a lack of studies that evaluated the conventional definition of whole body.

The measurement of whole body orthostatic posture is performed for a number of reasons. The most common reason is for determining if postural faults are a contributing factor to the experience of MSK conditions. Postural measurements could also be performed for prognostic or preventive purposes. People of all ages and of both genders undergo whole body orthostatic postural measurement. Some may be symptomatic while others are healthy and pain-free. The representative characteristics of the population that have their posture measured is heterogeneous and quite varied. None of the studies included in this literature review reflected the conventional practice in its entirety. No study actually measured the whole body as defined by convention nor did any study discussed or evaluated the measurement of muscles. However, there were some similarities to convention with regard to population characteristics and measurement methods.

4.6.1 Population characteristics

Three studies in this literature review measured the posture of a heterogeneous population. Their subjects were of both genders and were a mix of healthy and symptomatic individuals. However, they each limited their studies to different age groups. The McEvoy *et al*⁶⁰ study measured the posture of children. The Normand *et al*⁶¹ study measured young adults while the Tomkinson *et al*⁶⁴ study measured slightly older adults with a mean age of 35 years. In addition to measuring a heterogeneous population, the two studies that measured adults were considered to be high quality according to the CAT results.

4.6.2 Views

The advantage of measuring the four conventional views is self-evident. Each view provides additional sources of information. Only the Ferreira *et al*⁵⁸ study captured postural images in all four conventional views. This was also the study that evaluated the greatest number of postural measurements, that of 29. Watson *et al*⁶⁵ assessed four views of posture but used an oblique view in place of one of the lateral views. The intention of this view was to enable better interpretation of the thoracic kyphosis as the scapula can obscure it. The oblique view is a realistic reflection of the visual method of analysis. A health care practitioner does not typically stand in one place when measuring an individual's posture. He or she would move around the patient to capture different views, which would include the oblique views. It may be beneficial to add an oblique view to the photogrammetric process.

Three postural views are sometimes assessed in clinical practice. This may be to save time or for convenience. When three views are measured, they usually consist of an anterior, posterior and one lateral. The assumption is made that one lateral view is sufficient. The lateral view provides information about the sagittal curvature of the spine and carriage of weight as it relates to the gravitational line. One could argue that these measurements do not require the assessment of the opposite lateral view. However, one would not see the positioning of the upper or lower extremities of the opposite side. Iunes *et al*⁴⁴ took 14 measurements from these three views. Newton *et al*⁶³ also took three postural views but captured a superior view and omitted the anterior view. The superior view allows visualisation of the body regions' rotation around the

central axis and can give insight into torsional effects on the spine. This was the only study that included measurements of the upper extremities.

The six remaining studies did not take measurements from the posterior view. An argument could be made that the assessment of symmetry between bilateral structures such as the shoulders and iliac crests can be performed on either an anterior view or a posterior view; making only one view necessary. Indeed, the anterior view can demonstrate knee height, pelvic level, shoulder level, and eye level. However, the posterior view provides information about the tone of the gluteus maximus, coronal spinal curves, and scapular placement. Two studies captured only one lateral view and took five postural measurements from the image. Five was the fewest number of measurements taken among the studies.

The Ferreira *et al*⁵⁴ study illustrates that it is possible to capture and take goniometrical measurements of posture in the four conventional views. The authors that limited the number of views presented no rationale for their decision. Perhaps it is a reflection of measurement priority. There is no standard number of measurements. Moreover, it is not known which measurements provide the most valuable clinical information.

4.7 Grey literature contribution

All three methods for the measurement of whole body orthostatic postural found in the grey literature involved a version of photogrammetry. These commercially available and potentially affordable methods share a number of similarities to those presented in the academic literature. Prior to the photographic capture of a person's orthostatic posture, markers are placed on specified anatomical landmarks. However, the details of how the landmarks are identified and marked were not presented in the websites. Nor were the details of how the person was positioned for the image capture. None-the-less, these findings demonstrate the potential acceptance of quantitative postural measurement methods in general practice. PostureScreen® did provide a link to an evidentiary study on its method's reliability; however, more rigorous investigations into this and the other commercially available products should be performed to further evaluate the reliability of their methods.

4.8 Summary

The search for a valid and reliable quantitative method for measuring whole body orthostatic posture reflective of conventional practice has led to photogrammetry; the process of taking goniometric measurements from photographic images. The findings in the grey literature illustrates the commercial appeal and viability of this method of postural measurement. Establishing the validity and reliability of a whole body postural measure is a challenging process. This is evidenced by the low volume and quality of available research on the topic. However, the research quality has improved in recent years. This improvement may reflect an advancement in methodological design, reporting standards, as well as in technology.

Although each study included in this review claimed its method of postural measurement was valid and/or reliable, misuse of terminology and a methodological flaw necessitated a cautious interpretation of their conclusions. The study that stated an aim of determining validity or accuracy failed to compare the results of their method of postural measurement with a clinical reference standard. As a result, it evaluated a type of measurement error, an aspect of reliability. The methodological flaw that resulted in the need to limit the interpretation of reliability was the use of a single marker and photographer. The use of such an individual removed the confounders that reflect clinical reality, that of marking and positioning the subject before capturing the postural images. Unless the entire postural measurement protocol is assessed, the determination of the method's reliability is limited. The interpretation of reliability in the methods presented from these studies is limited to precision and measurement error. The results from these studies lead us to understand that the use of photogrammetric analysis of whole body orthostatic posture may provide measurements that have good precision and low levels of measurement error.

There were many variations among the different photogrammetric methods. There were variations in anatomical landmarks used, subject positioning, views assessed, and number of measurements. Some of these variations have the potential of contributing to measurement error and require further investigation as to their effect on reliability. Others may simply reflect the authors' opinions about postural measurement priorities and have no bearing on reliability. There were also differences in technology use. Most methods used a single digital camera. Others used multiple digital video cameras or even a digital 'scanner'. Based on the results of this study, a

single camera would be sufficient. A second camera may reduce the time spent with the subject but space could be a limiting factor for its use.

Common to all photogrammetric methods presented in this review was the use of photogrammetric software programs. Each method used a different software program. It appears that the differentiating factor determining the method with the most precision is the software program. The most precise software program has yet to be determined. In addition to precision, a valuable characteristic of a clinically useful photogrammetric software program would be flexibility. The variables stated earlier indicate that different health care professionals prioritise different postural measures. An appropriate software program should allow the expression of these different priorities. For instance, the measurement of full body could include muscles and the upper extremities in multiple views. In terms of research, it would be important to explore the possible value of different postural measurements in a reliable manner.

The desirable characteristics of a future photogrammetric method to measure whole body orthostatic posture include reliable measurements of the entire body (head, neck, trunk, spine, pelvis, upper and lower extremities, and muscles) in multiple views (anterior, posterior, right and left lateral). A standardised dress code and positioning instructions are necessary. The method should be able to be performed within a private practice setting, easy to use, and require a short period of time.

Chapter 5: Strengths and limitations of the review

5.1 Strengths

The strengths of this study lie in the scope and rigour of its search. Eight databases were searched and numerous keywords were employed in an attempt to capture as many relevant methods of quantitative whole body orthostatic postural measurement as possible. The chiropractic database, MANTIS was searched with the intention of including this profession's body of work on the topic. Not all of the literature associated with chiropractic can be found in databases such as Medline or CINAHL. The MANTIS search retrieved 162 responses, seven of

which were unique. However, none of the responses contributed to the final number of included studies.

The measurement of posture is associated with many terms. Previous systematic reviews do not take into account the magnitude of these variations. For example, the term measurement has the following synonyms; analysis, assessment, tool, test, and clinical assessment tool. The last search term was a medical search heading or MeSH. Grey literature was searched to identify postural measurement methods used in the private and commercial sectors. The combined search strategies allowed for a thorough search for methods of whole body orthostatic postural measures.

The selection process for the research studies found through the literature databases was rigorous. Two individuals evaluated the retrieved responses for inclusion and exclusion criteria. The study was included when both reviews agreed. In instances where disagreement occurred, a third reviewer determined the outcome of the final decision. An additional strength of this paper is that we assessed the quality of the included studies with the use of a critical analysis tool. This tool was created specifically for index tests that do not have a clear target condition and has been previously used for a systematic literature review involving postural measurement. Lastly, the detailed analysis of the identified methods of quantitative postural measurements combined with their comparison to conventional methods of measurement provided a unique overview. It called into question the practice of all methods and highlighted the extent to which future investigations are needed in order for whole body orthostatic postural measurements to meet current expectations of evidence-based practice.

5.2 Limitations

Some limitations of this paper stem from the search of the grey literature. The search was limited to two search engines. A more extensive search through additional sources such as poster presentations, conference proceedings, and unpublished papers would potentially uncover additional methods of whole body postural measurements. Moreover, the grey literature search was conducted by and eligibility was determined by a single individual. The participation of a minimum of two individuals would strengthen the results of this activity. The critical analysis

tool may have been over simplified with only thirteen criteria assessed. The use of a tool that has greater focus on methodological quality rather than quality of reporting would be beneficial.

Chapter 6: Future research

The results from this study demonstrate that photogrammetric software programs can provide postural measurements with good precision. Further research is required to identify the program that offers the best potential for research and clinical practice. Reliability, flexibility, ease of use, time demands, and cost are considerations for a diagnostic test when considering translation into clinical practice. The program should allow the measurement of the entire body from multiple views

Once a precise and flexible photogrammetric software program is identified or created, investigation into the multiple variables that affect posture and other potential sources of error could begin. The variables of particular interest relate to photogrammetric protocols and patient positioning. After a protocol that manages the postural variables is established, reliability studies should be performed. These studies should begin with the assessment of inter- and intra-rater reliability of the entire protocol as well as determining the degree of measurement error. Following on from this would be to trial the whole body orthostatic postural measurement tool on a heterogeneous population comprised of healthy and symptomatic individuals of both genders.

Once a reliable method of postural analysis is established, research into the value of each postural view and postural measurement could progress. The ultimate aim is to explore the validity of the many hypotheses and suppositions associated with posture with confidence.

Chapter 7: Conclusions

The authors of the studies included in this review have contributed well to the scientific advancement of quantitative measurement for whole body orthostatic posture. Based on their work, photogrammetric presents as a method with potential for being a precise postural measurement. The photogrammetric software program that provides the greatest precision with the most flexibility remains unclear. The degree of variability among the protocols outlined in

these studies did not provide insight into how best to manage the variables that can affect posture. In addition, their methods of postural measurements did not conform well to the conventional methods of practices. Many methods took measurements of too few aspects of the body from too few views. Photogrammetric appears to be a promising quantitative method to measure whole body orthostatic posture. Additional research is required to develop a method with greater validity, reliability, and flexibility. Such a method would have the potential to become the reference standard for orthostatic postural measurement and could be used to put the numerous theories about posture to the test. The association of faulty posture with MSK symptoms, the prognostic capabilities of postural measurement, and the possible prevention of MSK conditions are examples of such theories. The true clinical utility of measuring posture would then be better understood.

References

1. Kuhns JG, Klein A, Regen E, Williams PC, Crowe HE. Posture and its relationship to orthopaedic disabilities. Report. Chicago, Illinois: American Academy of Orthopaedic Surgeons; 1947.
2. Page P, Frank C, Lardner R. Assessment and treatment of muscle imbalance: the Janda Approach. Champaign, IL: Human Kinetics; 2010.
3. Goldthwait JE. Anatomic form and posture, important factors in the treatment of pulmonary tuberculosis. *Boston Med Surg J.* 1916;175(3):88-93.
4. Brown LT. A combined medical and postural examination of 746 young adults *JBJS.* 1917;s2-15(11):774-87.
5. Forrester-Brown M. Poture as a factor in health and disease. *Br Med J.* 926;1(3407):690-3.
6. Schwartz L. A resume, with comments, of the available literature relating to posture. *Public Health Rep.* 1927;42(18):1219-48.
7. Burt HA, editor Effects of faulty posture. *Proceedings of the Royal Society of Medicine;* 1949 November 9, 1949; Royal Society of Medicine. 1949;42:187-194.
8. Benda BJ, Riley PO, Krebs DE. Biomechanical relationship between center of gravity and center of pressure during standing. *IEEE Trans Rehabil Eng.* 1994;2(1):3-10.
9. Winter DA. Human balance and posture control during standing and walking. *Gait Posture.* 1995;3(4):193-214.
10. Kendall F, McCreary E, Provance P, Rodgers M, Romani W. *Muscles: testing and function with posture and pain.* 5th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2005. Chapter 2, Posture:p. 49-117.
11. Kendall Historical Collection Baltimore MD: University of Maryland; 2017 [cited 2017 December 1]. Available from: <https://www.hshsl.umaryland.edu/resources/historical/kendall/>.
12. Bickley LS. *The Musculoskeletal System.* Bate's guide to physical examination and history taking12 ed. Philadelphia, PA: Wolters Kluwer, Lippincott, Williams & Wilkins; 2016. p. 458-527.
13. Greenman P. *Principles of manual medicine.* 2nd ed. Baltimore MD: Williams & Wilkins; 1996.
14. Magee DJ. *Assessment of posture. Orthopedic physical assessment.* 6 ed. St. Louis, MO: Elsevier Saunders; 2014. p. 1017-53.
15. Morris C, Chaitow L, Janda V. Functional examination for low back syndromes. In: Morris C, editor. *Low back syndromes: integrated clinical management.* New York, NY: McGraw Hill; 2006. p. 333-416.
16. Marcus M, Gerr F, Monteilh C, Ortiz DJ, Gentry E, Cohen S, et al. A prospective study of computer users: II. Postural risk factors for musculoskeletal symptoms and disorders. *Am J Ind Med.* 2002;41(4):236-49.
17. Gureje O, Von Korff M, Simon GE, Gater R. Persistent pain and well-being; a World Health Organization study in primary care. *JAMA.* 1989;280(2):147-51.

18. Jordan K, Joud A, Bergknut C, Croft P, Edwards J, Peat G, et al. International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. *Ann Rheum Dis*. 2014;73:212-8.
19. Kasselbaum K, Collaborators GDaH. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and health life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016;388:1602-58.
20. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. National Surveys. *Spine*. 2006;31(23):2724-7.
21. Hart LG, Deyo RA, Cherkin DC. Physician office visits for low back pain: frequency, clinical evaluation, and treatment patterns from a U.S. National Survey. *Spine*. 1995;20(1):11-9.
22. Sackett D, Haynes R, Guyatt G, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. 2nd ed. Boston MA: Little, Brown and Co; 1991.
23. Streiner D, Norman G. *Health measurement scales: a practical guide to their development and use*. 2nd ed. Oxford: Oxford University Press; 1995.
24. Straus S, Richardson W, Glasziou P, Haynes R. *Evidence-based medicine: how to practice and teach EBM*. 3rd ed. Edinburgh: Elsevier, Churchill, Livingstone; 2005.
25. Bossuyt PM, Reitsma JB, Bruns DE, et al. The stard statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med*. 2003;138(1):W1-12.
26. Herbert RD, Sherrington C, Maher C, Moseley AM. Evidence-based practice -- imperfect but necessary. *Physiother Theory Pract*. 2001;17(3):201-11.
27. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-45.
28. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 2005;37(5):360-3.
29. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
30. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? *Injury*. 2011;42(3):236-40.
31. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: An updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. 2015;277(3):826-32.
32. Cohen L, Kobayashi S, Simic M, Dennis S, Refshauge K, Pappas E. Non-radiographic methods of measuring global sagittal balance: a systematic review. *Scoliosis Spinal Disord*. 2017;12(1):30.
33. Oatis CA. Characteristics of normal posture and common postural abnormalities. *Kinesiology: the mechanics and pathomechanics of human movement*. 2nd ed. Baltimore MD: Lippincott Williams & Wilkins; 2009. p. 875-91.
34. Leeftang MMG, Bossuyt PMM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J Clin Epidemiol*. 2009;62(1):5-12.

35. Szeto GPY, Straker L, Raine S. A field comparison of neck and shoulder postures in symptomatic and asymptomatic office workers. *Appl Ergon.* 2002;33(1):75-84.
36. Dael N, Mortillaro M, Scherer KR. Emotion expression in body action and posture. *Emotion.* 2012;12(5):1085-101.
37. Rosario JLP, Diogenes MSB, Mattie R, Leite JR. Can sadness alter posture? *J Bodywork Mov Ther* [Internet]. 20012. Available from: <http://dx.doi.org/10.1016/j.jbmt.2012.12.001>.
38. Cureton TK. Bodily posture as an indicator of fitness. *Res Q Exerc Sport.* 1941;12:sup2:348-67.
39. Tomkinson GR, Shaw LG. Quantification of the postural and technical errors in asymptomatic adults using direct 3D whole body scan measurements of standing posture. *Gait Posture.* 2013;37(2):172-7.
40. Kirby RL, Price NA, MacLeod DA. The influence of foot position on standing balance. *J Biomech.* 1987;20(4):423-7.
41. Mouzat A, Dabonneville M, Bertrand P. The effect of feet position on orthostatic posture in a female sample group. *Neurosci Lett.* 2004;365(2):79-82.
42. Dunk NM, Chung YY, Compton DS, Callaghan JP. The reliability of quantifying upright standing postures as a baseline diagnostic clinical tool. *J Manipulative Physiol Ther.* 2004;27(2):91-6.
43. Hazar Z, Karabacak GO, Tiftikci U. Reliability of photographic posture analysis of adolescents. *Journal of Physical Therapy Science.* 2015;27(10):3123-6.
44. Iunes DH, Bevilaqua-Grossi D, Oliveira AS, Castro FA, Salgado HS. Comparative analysis between visual and computerized photogrammetry postural assessment. *Rev Bras Fisioterapia.* 2009;13(4):308-15.
45. Tyson S. A systematic review of methods to measure posture. *Phys Ther Rev.* 2003;8(1):45-50.
46. Fortin C, Feldman DE, Cheriet F, Labelle H. Clinical methods for quantifying body segment posture: a literature review. *Disabil Rehabil.* 2011;33(5):367-83.
47. Brink Y, Louw Q, Grimmer-Somers K. The quality of evidence of psychometric properties of three-dimensional spinal posture-measuring instruments. *BMC Musculoskelet Disord.* 2011;12.
48. do Rosário JLP. Photographic analysis of human posture: A literature review. *J Bodyw Mov Ther.* 2014;18(1):56-61.
49. Adami Sedrez J, Tarragô Candotti C, Silveira Furlanetto T, Fagundes Loss J. Non-invasive postural assessment of the spine in the sagittal plane: a systematic review. *Motricidade.* 2016;12(2).
50. Macedo Ribeiro AF, Bergmann A, Lemos T, Pacheco AG, Mello Russo M, Santos de Oliveira LA, et al. Reference Values for Human Posture Measurements Based on Computerized Photogrammetry: A Systematic Review. *J Manipulative Physiol Ther.* 2017;40(3):156-68.
51. Benzies KM, Premji S, Hayden KA, Serrett K. State-of-the-evidence reviews: advantages and challenges of including grey literature. *Worldviews Evid Based Nurs.* 2006;3(2):55-61.
52. Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract.* 2012;18(6):1126-32.

53. van der Wurff P, Hagmeijer R, Meyne W. Clinical tests of the sacroiliac joint: a systematic methodological review. part 1: reliability. *Man Ther.* 2000;5(1):30-36.
54. May S, Littlewood C, Bishop A. Reliability of procedures used in physical examination of non-specific low back pain: a systematic review. *Aust J Physiother.* 2006;52:91-102.
55. 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;96:179-190.
56. Abhia D, Bussey M, Ribeiro D, Tumilty S, Milosavljevic S. Validity and reliability of palpation-digitization for non-invasive kinematic measurement: a systematic review. *Man Ther.* 2013;18:26-34.
57. Brin S, Page L. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer Networks.* 2012;56(18):3825-33.
58. Ferreira EA, Duarte M, Maldonado EP, Burke TN, Marques AP. Postural assessment software (PAS/SAPO): Validation and reliability. *Clinics.* 2010;65(7):675-81.
59. Hough R, Nel R. Intra-Rater Reliability of the Posture Analysis Tool kit. *S Afr J Occup Ther.* 2013;43(1):15-20.
60. McEvoy MP, Grimmer K. Reliability of upright posture measurements in primary school children. *BMC Musculoskelet Disord.* 2005;6:35.
61. Normand MC, Descarreaux M, Harrison DD, Harrison DE, Perron DL, Ferrantelli JR, et al. Three dimensional evaluation of posture in standing with the PosturePrint: an intra- and inter-examiner reliability study. *Chiropr Osteopat.* 2007;15:15.
62. Pusic J, Pedisic Z, Dizdar D. Reliability of a photographic method for assessing standing posture of elementary school students. *J Manipulative Physiol Ther.* 2010;33(6):425-31.
63. Newton RU, Neal RJ. Three-dimensional quantification of human standing posture. *Gait Posture.* 1994;2(4):205-12.
64. Tomkinson GR, Shaw LG. Quantification of the postural and technical errors in asymptomatic adults using direct 3D whole body scan measurements of standing posture. *Gait Posture.* 2013;37(2):172-7.
65. Watson AWS, Mac Donncha C. A reliable technique for the assessment of posture: assessment criteria for aspects of posture. *J Sports Med Phys Fitness.* 2000;40(3):260-70.
66. PostureCo. PostureScreen: accurate postural assessment [Internet]. Trinity (FL)2017 [cited 2017 Nov 25]. Available from: <https://www.postureanalysis.com/posturescreen-posture-movement-body-composition-analysis-assessment/>.
67. PostureZone. PostureZone software [Internt]. Alpharetta (GA)2017 [cited 2017 Nov 25]. Available from: www.posturezone.com.
68. Chinesport. Body posture analysis [Internet]. Udine, Italy: Chinesport; 2017 [cited 2017 Nov 25]. Available from: <http://www.chinesport.com/catalogue/posture-analysis/body-posture-analysis/>.
69. Boland DM, Neufeld EV, Ruddell J, Dolezal BA, Cooper CB. Inter- and intra-rater agreement of static posture analysis using a mobile application. *J Phys Ther Sci.* 2016;28(12):3398-402.
70. Ochodo EA, de Haan MC, Reitsma JB, Hooft L, Bossuyt PM, Leeftang MMG. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of “spin”. *Radiology.* 2013;267(2):581-8.
71. Bullock-Saxton J. Postural alignment in standing: A repeatability study. *Aust J Physiother.* 1993;39(1):25-9.

72. Janda V. On the Concept of Postural Muscles and Posture in Man. Aust J Physiother. 1983;29(3):83-4.

Appendix

Brink & Louw Critical Analysis Tool⁵²

1	If human subjects were used, did the authors give a detailed description of the sample subjects used to perform the (index) test?	Validity & reliability
2	Did the authors clarify the qualification, or competence of the rater(s) who performed the (index) test?	Validity & reliability
3	Was the reference standard explained?	Validity
4	If inter-rater reliability was tested, were the raters blinded to the findings of other raters?	Reliability
5	If intra-rater reliability was tested, were raters blinded to their own prior findings of the test under evaluation?	Reliability
6	Was the order of examination varied?	Reliability
7	If human subjects were used, was the time period between the reference standard and the index test short enough to be reasonably sure that the target condition did not change between the two tests?	Validity
8	Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time interval between repeated measures?	Reliability
9	Was the reference standard independent of the index test?	Validity
10	Was the execution of the (index) test described in sufficient detail to permit replication of the test?	Validity & reliability
11	Was the execution of the reference standard described in sufficient detail to permit replication of the test?	Validity
12	Were withdrawals from the study explained?	Validity & reliability
13	Were the statistical methods appropriate for the purpose of the study?	Validity & reliability