Measuring spinal and trunk shape using an electromagnetic sensor

By Eleanor Shalini Daniel

Bsc (Hons)

A thesis submitted in partial fulfilment of the requirements for a Master of Research degree awarded by Bournemouth University.

May 2021

Bournemouth University

Abstract

A critical component in the clinical assessment of spinal and trunk disorders is the analysis of posture. Currently the gold-standard is restricted by repeated radiation exposure and whilst alternative surface methods are available, these are limited to detection of spinal shape only. To date, no surface method has been extended to also quantify trunk shape. In order to address this, the aims of this research were 1) develop a method for measuring spinal and trunk shape using an electromagnetic system; 2) determine the validity and reliability of this method and 3) explore the optimal data processing for this method.

Using a repeated measures design, data were collected on phantom models of different shapes using an electromagnetic system. This provided the three-dimensional co-ordinates from which spine and trunk angles were derived.

The 6th order polynomial fit was deemed optimal for spinal shape measurements with an electromagnetic system. These measurements were highly reliable (ICC = >0.999), highly repeatable (MDC = <0.018º, SEM = <0.007º) and shown to be valid compared to a flexicurve method. The Lowess function was recommended for trunk shape measurements as it yielded good-to-excellent repeatability (ICC = 0.809-0.999), high absolute reliability (MDC = 0.18-4.0º, SEM = 0.06-0.07º) and angles derived were valid compared to a flexicurve method.

This study addressed a clinical need by developing a novel method for measuring trunk shape in addition to spinal shape using a surface method which was shown to be valid and reliable. Exploration of the method's optimal data processing techniques found the 6th order polynomial fit and Lowess function to be best for spinal shape and trunk shape measurements respectively. Additionally, whilst it is recommended that tangent lengths should not be used interchangeably, the tangent length chosen should not significantly affect measurements if used consistently. Meanwhile, the method's non-invasive, non-ionising and low-cost features make it clinically attractive. Therefore, this research holds future prospects for the examination and monitoring of disease and treatment outcomes as well as, the understanding of many disorders, such as scoliosis. Although further research is warranted, this method has the potential for use in routine clinical practice.

Contents

List of figures

Chapter 1 – Review of related literature

Chapter 2 – Development of methods

Chapter 3 – Main methods

Chapter 4 – Results

List of tables

Chapter 1 – Review of related literature

Chapter 2 – Development of methods

Chapter 3 – Main methods

Chapter 4 – Results

List of abbreviations

º – degrees

3-D – three dimensional/ three dimensions

6DOF – six degrees of freedom

95%CI/ CI95% – ninety five percent confidence interval

AIS – adolescent idiopathic scoliosis

AM – angular motion

BMI – body mass index

cm – centimetre

CMC – coefficient of multiple correlation

CoV – coefficient of variation

DMT – digitised manual technique

EM – electromagnetic

ICC – intraclass correlation coefficient

LBP – Lower back pain

LLB – left lateral bending

LPSIS – left posterior superior iliac spine

m – metres

MDC – minimal detectable change

MDC95 – minimal detectable change at 95 percent confidence level

mm – millimetre

MRI – magnetic resonance imaging

PCC – pearson’s correlation coefficient

QMA – quantitative motion analysis

r^2 – R squared

RLB – right lateral bending

RMS – root mean square

RMSE – root mean square error

ROM – range of motion

RPSIS – right posterior superior iliac spine

sd – standard deviation

SEM – standard error of measurement

US – ultrasound

VF – video fluoroscopy

Acknowledgements

This MRes has been both a challenging and rewarding journey and I would like to acknowledge a number of people who have helped make it possible.

Firstly, I would like to thank my supervisors Dr Jonathan Williams and Prof Raymond Lee. I am particularly grateful to Jon for everything he has done for me throughout this process. Without his continuous support, patience and understanding, this work would not have been possible.

In addition, I extend my thanks to Simon Hatcher for his willingness to lend me his time and architectural skills, for which I am immensely appreciative.

Finally, this thesis would not have been accomplished without my friends and family. For their unwavering support, encouragement and motivation, I would like to thank my Mum, Dad, Grandma, Josh and Nigel. I am also very thankful to Polly, Ruth, Georgia, Andy and Hannah for always being there for me.

I feel very lucky to have received such support over the past two years, without which this MRes would not have been possible.

Declaration

This thesis is the sole work of the author except for the bespoke MATLAB algorithms used to determine spine and trunk shape that were devised by Dr Jonathan Williams.

Organisation of thesis

Chapter one highlights gaps in current knowledge, providing the background to the development of this study's research question. It begins with a systematic review of the current literature related to the assessment of segmental range of motion in the lumbar spine using four imaging modalities. It then continues with a contemporary scoping review of literature exploring the reliability of thoracolumbar surface curvature measurements. This chapter concludes with the statement of the problem and need for the study.

Chapter two presents the development of the study's methods. This includes detailed description of the instrumentation, data capture and processing techniques used in the thesis.

Chapter three describes the study's main methods, outlining the study's aims, procedures, data processing and analysis techniques.

Chapter four presents the results of the study, providing recommendations and conclusions for the method's use in spinal and trunk shape measurements.

Chapter five presents a general discussion of the thesis' findings.

Chapter six details the limitations and conclusions of the study before outlining recommendations for future work.

Chapter 1. Review of related literature

This chapter reviews current literature relating to the assessment of spinal segmental range of motion (ROM) and spinal shape in-vivo. This section highlights gaps in current knowledge and provided the background to development of this study's research question. The definition of segmental ROM used in this thesis refers to the motion of one spinal motion segment (i.e. vertebral body, vertebral endplate, vertebral disc, apophyseal joints, transverse processes, spinous processes and neural arch) on an adjacent spinal motion segment (e.g. L4/5). A systematic review of the literature investigating four non-invasive imaging modalities commonly used for clinical assessment of segmental ROM in the lumbar spine is presented first. The essence of this review aims to provide insight into, and future direction for, the methodology and tools required for exploration of long held segmental spinal ROM notions and also, a step change in the use of imaging for spinal pathologies. The reliability of video fluoroscopy, ultrasound imaging, magnetic resonance imaging and radiography is discussed with their advantages and disadvantages outlined, including their limitations. Following this, a change in the direction of this project's research question is explained. This chapter then continues with a scoping review of the literature exploring reliability of surface curvature measurements of the thoracolumbar spine. This review provides a description and critique of each measurement method before reliability results from previous research is discussed. Lastly, a summary of this section, recommendations for future work and the justification for this projects research question concludes this chapter.

1.1 The reliability of video fluoroscopy, ultrasound imaging, magnetic resonance imaging and radiography for measurements of segmental range of motion in the lumbar spine in-vivo: A systematic review

1.1.1 Introduction

Lower back pain (LBP) is the principal cause of disability worldwide and the sixth leading contributor to overall disease burden (Buchbinder et al. 2013). LBP affects approximately 540 million people globally at any one time (Global Burden of Disease and Injury Incidence and Prevalence Collaborators 2017). International studies have reported LBP point prevalence rates between 12 and 35% and lifetime prevalence rates ranging from 49 to 80% (Maniadakis and Gray 2000). As a result, LBP is one of the most common reasons for an individual to seek medical attention (Ramdas and Jella 2018). In the United Kingdom alone, the estimated direct cost of healthcare for LBP exceeds £1 billion per year (Maetzel and Li 2002).

Despite this substantial economic burden, the pathophysiology of LBP is poorly understood (Murray et al. 2012). However, evidence suggests individuals commonly display differences in movement behaviour (Reeves et al. 2008; Hodges et al. 2009; Karayannis et al. 2013), some of which are believed to reflect changes in segmental spinal motion (Kulig et al. 2007; Golightly et al. 2016). One example would be spondylolisthesis, where structural change is associated with potentially excessive segmental motion. Spondylolisthesis has a prevalence of 11.5% in the general population (Kalichman et al. 2009) but this rises significantly in sporting populations such as gymnastics and cricket fast bowling with reported prevalence's of 15% and 25% respectively (Hellström et al. 1990; Elliot et al. 1992). In addition, alterations to segmental motion are clearly the target for some treatments, with rates of spinal fusion more than doubling between 2004 and 2015 (Martin et al. 2019).

Therapeutic models of LBP assessment and treatment across a range of professions are firmly embedded in this notion of change in segmental ROM. In addition, segmental ROM assessment is also critical for enhancing the understanding of existing spinal diseases, aiding spinal diagnoses and evaluating contemporary treatment or surgical intervention. For these reasons, the measurement of lumbar spinal ROM is clinically important (Trudelle-Jackson et al. 2010) yet the ability to

measure an individual's segmental ROM non-invasively remains a challenge (Fritz et al. 1998).

Kinematics of the lumbar spine have been studied using a range of techniques including implantable bone pins (Gercek et al. 2008; Rozumalski et al. 2008) and implanted ball bearings (Park et al. 2012). However, these methods are usually the basis of research investigation and due to the invasive nature of these methods, they are unlikely to become routine clinical practice. Non-invasive methods including radiography (Dombrowski et al. 2018), video fluoroscopy (Okawa et al. 1998; Takayanagi et al. 2001), magnetic resonance imaging (McGregor et al. 2002; Huang et al. 2009) and ultrasound (Chleboun et al. 2012) are alternate methods reported in the literature which are available in current clinical practice. However, to date, no contemporary synthesis of the literature exploring these non-invasive methods to assess segmental ROM has been completed. Understanding these current methods will provide insight into, and future direction for, the tools required for exploration of long held segmental ROM notions and a step change in the use of imaging for spinal pathologies.

The purpose of this study was to review the reliability of four current non-invasive modalities (Video Fluoroscopy (VF), Ultrasound (US) imaging, Magnetic Resonance Imaging (MRI) and Radiography) used for measuring segmental ROM in the lumbar spine in-vivo, through systematic examination of the literature.

1.1.2 Materials and Methods

1.1.2.1 Search Strategy

In January 2021, a systematic literature search of electronic databases including: CINAHL complete, Academic Search Ultimate, MEDLINE Complete, ScienceDirect, Complementary Index, PsycINFO and Supplemental Index was conducted using key terms and Boolean logic for each modality, as listed in Table 1. Each search was limited to peer-reviewed articles, published in the English language. Table 1 shows the number of articles yielded for each modality after exact duplicates were removed.

Articles were initially screened by title, abstract, and where necessary full text, against inclusion and exclusion criteria (as listed in Table 2) by the author; with any uncertainty resolved by consensus with a supervisor (JW). All studies deemed appropriate for this review were also checked and confirmed by an additional author (JW). A detailed flow chart of the search can be seen in Figure 1.

Table 1. Search strategy

Search Strategy	**Number of results (with duplicates removed)**
Fluoroscopy AND spine AND motion AND reliability OR validity OR consistency OR repeatability NOT cervical OR thoracic	41
Ultrasound OR ultrasonography OR US OR USS OR ultrasound imaging AND lumbar AND reliability OR validity OR consistency OR repeatability NOT scoliosis OR musc* OR cervical OR thoracic	157
Magnetic resonance imaging OR MRI OR MRI scan AND lumbar spine AND motion OR kinematics AND reliability OR validity OR consistency OR repeatability	32
X ray OR radiology OR radiograph* AND reliability OR validity OR consistency OR repeatability AND lumbar AND motion OR kinematics NOT videofluoroscopy OR fluoroscopy OR scoliosis	200
Total	**430**

(* (asterisk), truncation).

1.1.2.2 Inclusion and exclusion criteria

Studies needed to investigate segmental ROM of the lumbar spine in-vivo (human participants) using VF, US, MRI or Radiography. Consideration of the modality's

psychometric properties was also required by the articles. See Table 2 for detailed inclusion and exclusion criteria and Table 3 for reasons for article rejection.

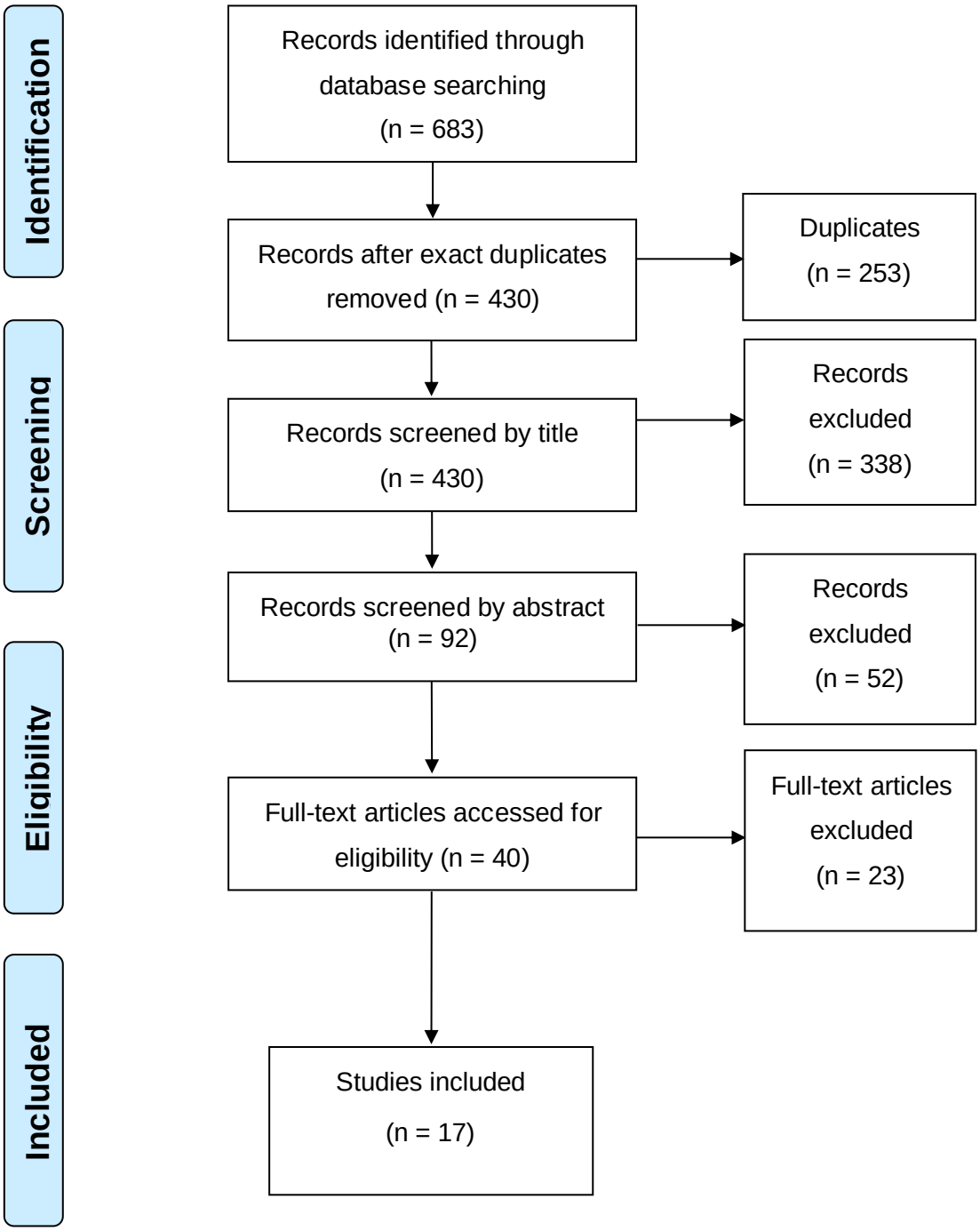


Figure 1. PRISMA flow diagram of search (n, number; =, equals)

Table 2. Inclusion and exclusion criteria

Inclusion Criteria	Exclusion criteria
Human participants	Animal studies/ studying in-vitro
Measuring segmental or intersegmental ROM defined as angular rotation of one vertebral body on another (or a representation of this)	Articles solely investigating linear translation of a vertebrae
Measurements at the lumbar spine	Measuring ROM in only cervical and/ or thoracic spine
Measuring ROM with VF, US, MRI or radiography	Not using modality for imaging
Investigating reliability or validity	Studies published after January 2021
	Non-objective psychometric outcome
	Studies published not in the English language

(ROM, range of motion; VF, video fluoroscopy; US, ultrasound; MRI, magnetic resonance imaging).

Table 3. Reason for article rejection after accessing full-text citation

Reason for full-text citation rejection	No. of citations in this category
Not examining modality's psychometric properties	7
Examination in-vitro	4
Examination of an animal spine	1
Not VF/ US imaging/ MRI/ radiography	2
No reliability statistics included	2
Not assessing kinematics, motion or ROM	1
Measuring only linear translation	4
A report	1
Comparing a method against modality	1
Total no. of citations rejected	23

(No., number of; ROM, range of motion; VF, video fluoroscopy; US, ultrasound; MRI, magnetic resonance imaging).

1.1.2.3 Quality assessment

Critical appraisal of the methodological quality of each article was completed by the author using an assessment tool for observational cohort and cross-sectional studies, taken from the National Heart, Lung and Blood Institute (NHLBI) (2019). The appraisal criteria consisted of 14 questions concerning study population and sample size, exposure measures and assessment as well as, statistical analyses. These could be answered yes, no, cannot determine, not applicable or not reported. Then, an overall quality rating was given based on these answers. The results can be seen in Table 4. This tool was used because its design draws focus to the key concepts of a study facilitating evaluation of its internal validity (NHLBI 2019).

Table 4. Quality assessment

Criterion	Study																
	Haas et al. (1990)	Maigne et al. (2003)	Breen et al. (2006)	Cakir et al. (2006a)	Cakir et al. (2006b)	Landel et al. (2008)	Pearson et al. (2011)	Sui et al. (2011)	Chleboun et al. (2012)	Mellor et al. (2014a)	Yeager et al. (2014)	Cuesta-Vargas (2015)	Tozawa et al. (2015)	Du Rose and Breen (2016)	Tozawa et al. (2016)	Breen et al. (2018)	Mahato et al. (2019)
1	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y
3	CD	Y	CD	CD	CD	CD	N	CD	CD	Y	Y	CD	CD	CD	CD	CD	CD
4	Y	Y	N	Y	Y	Y	Y	N	CD	Y	Y	Y	CD	Y	CD	Y	CD
5	N	N	N	N	N	N	N	N	N	Y	N	Y	N	Y	N	N	N
6	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
10	N	N	N	Y	Y	N	N	N	N	N	N	N	N	Y	N	Y	Y
11	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
12	NA	NA	NA	NA	NA	NA	NA	NA	NA	N	N	NA	NA	NA	NA	NA	NA
13	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NR	NA	NR	Y
14	N	N	N	N	N	N	N	N	N	Y	N	N	N	N	N	Y	N
Overall	Fair	Good	Fair	Good	Good	Good	Good	Fair	Fair	Good	Good	Good	Fair	Good	Fair	Good	Good

(Y, yes; N, no; CD, cannot determine; NA, not applicable; NR, not reported; overall, overall quality rating)

Table 5. Data Extraction

Author	Modality	Participant characteristics	Psychometric outcome(s) assessed	Vertebral level(s)	Measurement Methods
Haas et al. (1990)	Radiography	58 participants Mean age 28 Students from Western States Chiropractic College Asymptomatic	Kappa SEM	L1-L5	Investigated tilt into side bending and rotation in standing and lateral bending positions.
Maigne et al. (2003)	Radiography	74 participants with Chronic LBP 42 had pain immediately on sitting down and relieved on standing up from sitting (mean age 54.9, 6 males, 36 females) 32 age and gender matched participants who did not have symptom described above (mean age 57.5, 4 males, 28 females)	Mean difference LoA	L1-L5	Investigated angle change between adjacent vertebral endplates of lateral flexion-extension radiographs in standing and sitting positions.
Breen et al. (2006)	VF	30 male participants Aged 18-40 Asymptomatic 4 subjects assessed	RMSE	L3-L5	Investigated non weight bearing side flexion, flexion and extension. One movement trial for each individual. Manual identification of first frame then automated analysis using vertebral corners as reference points.
Cakir et al. (2006a)	Radiography	24 participants. 10 males, 14 females All with monosegmental degenerative disc disease Mean age 40.2	PCC	L4-S1	Investigated flexion-extension radiographs in standing. Three examiners, two took measurements between-day.

Cakir et al. (2006b)	Radiography	24 participants. 10 males, 14 females following a monosegmental total disc replacement at L4-5 or L5-S1 Mean age 40.2	PCC	L4-S1	Investigated flexion-extension radiographs in standing. Follow up study. Three examiners, two took measurements between-day.
Landel et al. (2008)	MRI	29 participants 13 Males, 16 Females Aged 18-45 Diagnosis of non-specific LBP Recent onset of centralised LBP	ICC SEM	L1-L5	Investigated P-A force in non-weight bearing. Segmental mobility quantified by measuring the change in the intervertebral angle between the resting position and the end range of the P-A force application.
Pearson et al. (2011)	Radiography	30 participants with spondylolisthesis at L4-5 37% males, 63% females Randomly selected from the spine patient outcomes research trial Mean age 66	ICC SEM	L1-S1	Investigated intervertebral rotation of flexion-extension radiographs. Measurements made with a digitised manual technique by three raters and by a quantitative motion analysis software by three different raters.
Sui et al. (2011)	VF	12 participants 10 healthy, 2 lumbar spondylolisthesis 8 Males, 4 Females Aged 19-38	ICC	L1-S1	Investigated seated flexion and extension. Automated tracking after manual marking of four vertebral corners.
Chleboun et al. (2012)	MRI US	6 participants 2 Males, 4 Females Aged 22-35 Asymptomatic	ICC CoV	L1-L5	Investigated supine flexion and extension postures. MRI – distance between inferior edge of caudal and cranial spinous processes measured.

					US – distance between the peak of the curvature of caudal and cranial spinous processes measured. Manual/visual method used digitally.
Mellor et al. (2014a)	VF	80 participants 44 Males, 36 Females Aged 21-50 40 with Chronic non-specific LBP 40 asymptomatic volunteers Convenience sample	ICC SEM	L2-L5	Investigated lying flexion and extension. Sequences processed using automatic tracking algorithms after manual template application to first image. Landmark used was vertebral body corners.
Yeager et al. (2014)	VF	61 participants 52% Male, 48% Female Aged 31-62 34 Asymptomatic 14 preoperative (with confirmed pathology) 13 post operative (with a previous lumbar procedure)	ICC SEM	L1-S1	Investigated flexion and extension in upright and lying positions. Intervertebral rotation and intervertebral translation measured using automated vertebral motion analysis tracking algorithms as well as a manual technique.
Cuesta-Vargas (2015)	US	15 male participants Convenience sample Asymptomatic	ICC SEM MDC	L4-L5	Investigated flexion in upright position. 10 reps forward bending and neutral. Manual/visual identification of spinous process then semi-automated orientation estimation.
Tozawa et al. (2015)	US	10 participants Healthy Males Aged 20-23	ICC MDC	L2-L3	Investigated prone, prone on elbows and kneeling with flexed spine postures. Measured distance between caudal end of L2 spinous process to cranial end of L3 spinous process and from top of L2 spinous process to top of L3 spinous process.

					Manual/ visual identification of landmarks.
du Rose and Breen (2016)	VF	18 male participants Mean age 27.6 No history of LBP	ICC SEM	L2-S1	Investigated flexion and extension in upright position. Sequences processed using automatic tracking algorithms after manual template application to first image.
Tozawa et al. (2016)	US	10 male participants Aged 20-23 No history of orthopaedic disease or dysfunctions	ICC MDC	L1-L2	Investigated prone, prone on elbows and kneeling with Lx fully flexed postures. Manual/ visual identification of spinous process. Measured distance between caudal end of L1 and cranial end of L2 spinous processes.
Breen et al. (2018)	VF	109 participants 66 Males, 43 Females Aged 21-80 Healthy volunteers Convenience sample	ICC MDC	L2-S1	Investigated flexion, extension, left side flexion and right side flexion in recumbent and standing positions. Single motion sequences. Manual first image registration then frame-to-frame automatic tracking.
Mahato et al. (2019)	MRI	10 participants Aged 18-60 Volunteers Asymptomatic	ICC CoV	L2-L4	Investigated side flexion in weight-bearing upright position. Measured changes in intervertebral axes positions using cranial to caudal vertebrae measurement and displacements in individual vertebrae within a calibrated imaging space.

(SEM, standard error of measurement; L, lumbar; LoA, limits of agreement; VF, video fluoroscopy; RMSE, root mean square error; PCC, pearsons correlation coefficient; MRI, magnetic resonance imaging; LBP, lower back pain; ICC, intra-class correlation co-efficient; P-A, posterior-anterior; %, percent; US, ultrasound; CoV, coefficient of variation; MDC, minimal detectable change; Lx, Lumbar spine).

1.1.3 Results

A total of 17 studies were eligible for this review (Haas et al. 1990; Maigne et al. 2003; Breen et al. 2006; Cakir et al. 2006a; Cakir et al. 2006b; Landel et al. 2008; Pearson et al. 2011; Sui et al. 2011; Chleboun et al. 2012; Mellor et al. 2014a; Yeager et al. 2014; Cuesta-Vargas 2015; Tozawa et al. 2015; du Rose and Breen 2016; Tozawa et al. 2016; Breen et al. 2018; Mahato et al. 2019). Table 5 summarises the data extracted and Table 6 summarises the findings.

Six studies used VF to measure ROM (Breen et al. 2006; Sui et al. 2011; Mellor et al. 2014a; Yeager et al. 2014; du Rose and Breen 2016; Breen et al. 2018), five used radiography (Haas et al. 1990; Maigne et al. 2003; Cakir et al. 2006a; Cakir et al. 2006b; Pearson et al. 2011), four articles used US imaging (Chleboun et al. 2012; Cuesta-Vargas 2015; Tozawa et al. 2015; Tozawa et al. 2016), and two citations investigated MRI (Landel et al. 2008; Mahato et al. 2019). However, Chleboun et al. (2012) also included MRI results as a gold standard comparator for US.

Overall, 600 participants were included in this review; of which at least 289 were male and 243 were female. Two studies (Haas et al. 1990; Mahato et al. 2019) did not report the breakdown of male to female participants. 250 participants were symptomatic whilst 350 were classed as healthy or asymptomatic.

Most studies involved only healthy participants (Haas et al. 1990; Breen et al. 2006; Chleboun et al. 2012; Cuesta-Vargas 2015; Tozawa et al. 2015; du Rose and Breen 2016; Tozawa et al. 2016; Breen et al. 2018), whereas some had a mixture of symptomatic and asymptomatic individuals (Sui et al. 2011; Mellor et al. 2014a; Yeager et al. 2014; Mahato et al. 2019) and others studied specific populations (Maigne et al. 2003; Cakir et al. 2006a; Cakir et al. 2006b; Landel et al. 2008; Pearson et al. 2011). These included participants with LBP (Maigne et al. 2003; Landel et al. 2008), spondylolisthesis (Pearson et al. 2011), monosegmental degenerative disc disease (Cakir et al. 2006a) and monosegmental total disc replacement (Cakir et al. 2006b).

Articles measured segmental ROM during flexion and/ or extension (Maigne et al. 2003; Cakir et al. 2006a; Cakir et al. 2006b; Landel et al. 2008; Pearson et al. 2011;

Sui et al. 2011; Chleboun et al. 2012; Mellor et al. 2014a; Yeager et al. 2014; Cuesta-Vargas 2015; du Rose and Breen 2016). Others investigated flexion, extension and side flexion (Breen et al. 2006; Breen et al. 2018), two studies (Tozawa et al. 2015; Tozawa et al. 2016) quantified motion of the lumbar spine from three static positions; whilst one study looked at neutral positioning and lateral bending motion (Haas et al. 1990).

The psychometric properties of each modality analysed varied between reliability (Haas et al. 1990; Maigne et al. 2003; Cakir et al. 2006a; Cakir et al. 2006b; Breen et al. 2006; Pearson et al. 2011; Sui et al. 2011; Chleboun et al. 2012; Mellor et al. 2014a; Yeager et al. 2014; du Rose and Breen 2016; Tozawa et al. 2016; Breen et al. 2018; Mahato et al. 2019) or a combination of reliability and validity (Landel et al. 2008; Cuesta-Vargas 2015; Tozawa et al. 2015). All but five studies (Haas et al. 1990; Maigne et al. 2003; Breen et al. 2006; Cakir et al. 2006a; Cakir et al. 2006b) used intra-class correlation coefficient (ICC) as a metric of reliability. Additional outcomes studied amongst the articles were standard error of measurement (SEM) (Haas et al. 1990; Landel et al. 2008; Pearson et al. 2011; Mellor et al. 2014a; Yeager et al. 2014; Cuesta-Vargas 2015; du Rose and Breen 2016), coefficient of variation (CoV) (Chleboun et al. 2012; Mahato et al. 2019), pearson correlation coefficient (PCC) (Cakir et al. 2006a; Cakir et al. 2006b), kappa (Haas et al. 1990), root mean square error (RMSE) (Breen et al. 2006) and minimal detectable change (MDC) (Cuesta-Vargas 2015; Tozawa et al. 2015; Tozawa et al. 2016; Breen et al. 2018).

All studies had an overall quality rating of fair or good based on the 14-point appraisal checklist (NHLBI 2019) but demonstrated similar methodological flaws and thus, shared common threats to validity.

1.1.3.1 Methodological Analysis

Only three studies (Mellor et al. 2014a; du Rose and Breen 2016; Cuesta-Vargas 2015) justified their sample size or provided a description of study power. This methodological element is important to ensure an adequate number of participants are studied to yield valid estimates of reliability. As sample size varies considerably across the studies, it is likely that the power also varies significantly and this should be considered when extrapolating the findings.

All studies, excluding five (Cakir et al. 2006a; Cakir et al. 2006b; du Rose and Breen 2016; Breen et al. 2018; Mahato et al. 2019) took ROM measurements from images at only one stage during the study period, thus exploring within-day repeated measures reliability. Whilst this is likely to result in more consistent movement patterns; conclusions regarding reliability of between-day repeated movements are not possible.

Additionally, aside from two studies (Mellor et al. 2014a; Breen et al. 2018), key potential confounding variables were not reported. Confounding factors are characteristics which may influence the dependant variable and thus, alter the findings of a study. For example, US imaging can be more difficult in individuals with a high body mass index (BMI) (Modica et al. 2011) and likewise, this category of participants may require a stronger radiation dose for VF (Cushman et al. 2016) and radiographs (Metaxas et al. 2019). Similarly, the quality of MRI images can be affected by permanent cosmetics, including tattoos (Durbridge 2011). In the absence of the consideration of confounding factors, it is difficult to determine if their presence or absence affected the results.

1.1.3.2 Reliability

1.1.3.2.1 Video Fluoroscopy

1.1.3.2.1.1 Segmental ROM values

Segmental ROM of flexion across the studies were similar, ranging from 4.05 to 7.10 degrees (º) for lying (Mellor et al. 2014a; Yeager et al. 2014; Breen et al. 2018) and 9º to 14º for standing (Breen et al. 2018). Extension has been less frequently studied with segmental ROM values of 4.11º to 5.31º (Mellor et al. 2014a), 2.00º (Yeager et al. 2014), 5.33º for lying and 2.01º in standing (Breen et al. 2018). Landel et al. (2008) and Sui et al. (2011) did not report individual segmental ROM values.

1.1.3.2.1.2 Intra-rater reliability of segmental ROM measurement

In VF, automated tracking algorithms are commonplace; where bony boundaries are automatically tracked by a computer from which ROM calculations are made (Breen et al. 2012). In most cases, an operator is required to manually mark the first image,

or few images, from which the tracking algorithm commences (Davis et al. 2015). This manual identification is known to be an important source of error both between individuals and within individuals (Plocharski et al. 2018). To this end, a body of work has concentrated on quantifying the variability this manual marking of images affords (Breen et al. 2006; Plocharski et al. 2018; Breen et al. 2012). The methodology involves participants completing one movement in the fluoroscope, from which multiple mark ups and analysis are completed. This is either repeated by the same individual or between individuals.

Using a mixture of individuals with pathology and those asymptomatic, Yeager et al. (2014) demonstrated excellent reliability (ICC 0.98, CI95% 0.98-0.99, SEM 0.10º, SEM% = 2% for flexion, 5% for extension) for the same investigator repeatedly marking-up and processing the same VF sequences. These included sagittal plane motions only and were a mixture of upright and recumbent movements. Similar findings were reported by Mellor et al. (2014a) for lying motion, where excellent reliability was established for sagittal plane motions (ICC 0.92-1.0, CI95% 0.72-1.0, SEM 0.10º to 0.35º, SEM% = 3% flexion, 8% extension). In addition, similar findings have been reported for recumbent side bending movements where ICCs ranged from 0.99-1.0, CI95% 0.95-1.0, SEM 0.08º to 0.17º, SEM% = 3% lateral bending (Mellor et al. 2014a).

These results demonstrate that if the same individual marks up and processes the images; VF can be used to reliably measure lumbar sagittal ROM in recumbent and upright as well as, recumbent side-bending with a small SEM.

1.1.3.2.1.3 Inter-rater reliability of segmental ROM measurement

Investigation of the inter-rater reliability of processing the same images show sagittal ICC values remain good-to-excellent but are slightly lower for extension (ICC 0.74-0.99, CI95% 0.23-0.99) (Mellor et al. 2014a), Yeager et al. (2014) (ICC 0.96, CI95% 0.95-0.97). It should be noted that the confidence interval for extension was large; with lower estimates suggesting poor reliability. In addition, the SEM values were also higher at 0.22º (or 5% flexion and 11% for extension) (Yeager et al. 2014), and 0.17º to 0.31º for flexion (or 7%) and 0.27º to 0.77º for extension (or 19%) (Mellor et al. 2014a). It is not clear why Yeager et al. (2014) have much lower SEMs compared with

Mellor et al. (2014a) but it is apparent that Mellor et al. (2014a) contained outliers for extension which could have affected results. Furthermore, the presented percentage SEM of 19%, reported at L2/3 by Mellor et al. (2014a), is the largest expected. It is possible that this was at the upper edge of their fluoroscope and therefore, was affected by image quality. Comparatively, Yeager et al. (2014) investigated L1/2 as their upper segment; suggesting a wider field of view. To this end, the L4/5 segment assessed by Mellor et al. (2014a) affords much better reliability for extension measurement (ICC 0.99, CI95% 0.96-1.0, SEM 0.27º, SEM% = 5%).

Altogether, these results indicate that larger variation is seen when different individuals process the same VF motion sequences, even though automated algorithms are used. Nevertheless, the ICCs remain good-to-excellent. Furthermore, although some larger errors are noted for extension; errors were small, especially for flexion.

1.1.3.2.1.4 Repeated Movements

The measurement of repeated movements is not common in VF research, presumably due to repeated participant exposure to radiation. However, establishing this enables more than just the error in marking up of VF images to be explored. Humans have a natural variance in movement (Stergiou and Decker 2011; Krüger et al. 2017), and this variance needs quantifying prior to any methods being employed for repeated measures in clinical studies. To date, only one study has investigated this. Breen et al. (2006) conducted a baseline measurement and follow-up measurement approximately 30 minutes later. Unfortunately, due to some technical issues, repeated measures reliability was only reported for side bending; with root mean square (RMS) errors of 2.75º to 2.91º (Breen et al. 2006). Raw data ranges were not reported but using those from Mellor et al. (2014a), who had a similar methodology; this represents around 52% error.

As a result, even with the same individual marking up images, this suggests that errors are quite large when exploring repeated measurements with VF.

1.1.3.2.1.5 Between-day measures

To explore between-day reliability some studies have taken a VF sequence, processed it and then reprocessed it sometime later to explore between-day intra-rater reliability (du Rose and Breen 2016). Excellent reliability for all vertebral levels was established with SEM values between 0.23º and 0.54º (du Rose and Breen 2016). However, this just represents errors associated with processing, rather than the additional biological variation of repeated movements.

This variation has been recently studied in 55 participants and over 200 motion segments, both in lying and standing, without pain or known pathology (Breen et al. 2018). ICC values suggest excellent reliability (0.80 for lying flexion and extension, 0.82-0.91 for standing flexion and extension) and small confidence intervals (lowest ICC ninety five percent confidence interval (95%CI) = 0.74) (Breen et al. 2018). Rather than reporting SEM, the authors chose to present the MDC at the 95% confidence level (MDC95). The MDC95 values are high suggesting significant variance in the repeated measurements. For example, the MDC95 value for flexion in lying was 4.66º (Breen et al. 2018). This means that with 95% confidence, a change greater than 4.66º represents true change beyond normal variation expected with repeated testing. The total range measured was 5.14º (Breen et al. 2018) thus, a change of 4.66º on 5.14º indicates 91% change is needed before there is confidence that this represents real change. As percentages, the magnitude of MDC95 was 91% for flexion in lying, 97% for extension in lying, 100% for flexion in standing and 176% for extension in standing (Breen et al. 2018). Therefore, a change from 9.1º average flexion in standing to over 18º would be required to provide 95% confidence that is was true change beyond natural variation. Previous studies would suggest this may not be physiologically possible or at least, puts the segmental ROM in the top 2% of normal (Mellor et al. 2014a). Similar findings were observed for side bending with good ICC values but high MDC95 values (60-69% lying, 97-98% standing) (Breen et al. 2018).

In summary, it is clear that within-day reliability of marking up and processing VF sequences is excellent for both intra- and inter-rater. However, the intra-rater reliability of measuring repeated movements within-day demonstrates larger errors, and these are even greater when investigating between-day reliability. Therefore, if using VF to investigate interventional changes across days, large change in segmental ROM are

needed to be sure these are greater than natural variability. This suggests low sensitivity to change of measuring repeated movements with VF.

1.1.3.2.2 Ultrasound

1.1.3.2.2.1 Segmental ROM values

In order to quantify segmental ROM using US, many studies (Chleboun et al. 2012; Tozawa et al. 2015; Tozawa et al. 2016) opted to visualize and then measure the linear distance between two adjacent spinous processes. Therefore, reporting of segmental ROM was commonly as a linear distance measurement in the units of millimetres (mm). Three studies investigated a 'neutral', flexed and extended position in either prone (Tozawa et al. 2015; Tozawa et al. 2016) or supine (Chleboun et al. 2012); whilst the other study investigated 'neutral' in standing and forward bending motion (Cuesta-Vargas 2015).

Values for spinous process separation in flexion ranged from 25.6mm to 32.3mm (Chleboun et al. 2012) and 29.2mm to 30.1mm (Tozawa et al. 2015). Distance measures for extension ranged from 21.5mm to 26.9mm (Chleboun et al. 2012) and were reported only in this study. Actual flexion ROM, taken from neutral, ranged from 3.0mm to 4.4mm and were only reported in one study (Chleboun et al. 2012). Segmental ROM was reported in degrees for Cuesta-Vargas (2015) using an image rotation method; yielding values of 15.4º to 16.3º for segmental ROM during flexion.

1.1.3.2.2.2 Intra-rater reliability of segmental ROM measurement

Intra-rater reliability estimates were reported as excellent by Chleboun et al. (2012) (ICC 0.94, CI95% 0.85-0.97), Tozawa et al. (2016) (ICC CI95% 0.963-0.999) and good-to-excellent (ICC CI95% 0.79-1.0, or with one examiner removed CI95% = 0.92-1.0) by Tozawa et al. (2015). Small coefficient of variances have been reported (1.8%) (Chleboun et al. 2012), along with moderate MDC95 values of 0.29mm (around 10%) (Tozawa et al. 2015). However, these could be as large as 1.8mm (around 60%) (Tozawa et al. 2016) based on segmental ROM of 3.0mm.

Both Tozawa et al. (2015) and Chleboun et al. (2012) positioned the participant in one position and collected all three measurements in that same position prior to then

moving onto the new position, henceforth eliminating the biological variation of repeated movement measurements. Nevertheless, this method doesn't replicate the type of method required to determine the repeated measures reliability that is more normal in biomechanical studies. This includes the biological variation of the human completing repeated movements.

1.1.3.2.2.3 Inter-rater reliability of segmental ROM measurement

Inter-rater reliability was explored in two studies (Tozawa et al. 2015; Tozawa et al. 2016) with good-to-excellent reliability reported by Tozawa et al. (2015) depending on the measurement method (ICC 0.914, CI95% 0.80-0.97; ICC 0.725, CI95% 0.55-0.87) and excellent reliability seen in their follow up study (ICC 0.969, CI95% 0.90-1.00) (Tozawa et al. 2016).

1.1.3.2.2.4 Repeated movements

Only one study investigated repeated movements (flexion) measured with US (Cuesta-Vargas 2015). Excellent estimates of reliability were reported for both within-day (CI95% = 0.995-0.999) and between-day (CI95% = 0.996-0.999) (Cuesta-Vargas 2015). Moreover, MDC95 estimates were made from the SEM reported in the article (SEM = 0.54º, MDC95 = 1.5º or 10%) (Cuesta-Vargas 2015), indicating change greater than 10% would represent true change.

Overall, these US results show that if the same individual captures repeated images without altering the participant's position; excellent intra-rater reliability should be expected. This expectation is further extended to between individuals. In addition, MDC95 values could be up to 60%, but these have not been established for between individuals. Consequently, this is an important consideration when designing test-retest studies. The values of MDC95 provide estimates as to the sensitivity of change, which is important when designing future experiments. Lastly, repeated movements have been less well investigated but estimates from a single study show promising reliability within- and between-day.

1.1.3.2.3 MRI

1.1.3.2.3.1 Segmental ROM values

The studies included in this review focussing on MRI often had primary aims not aligned to proving the utility of MRI for segmental ROM testing. Some used it as a gold standard comparator (Chleboun et al. 2012), others for validity of manual therapy (Landel et al. 2008). Only Mahato et al. (2019) focused on segmental ROM.

The distance between spinous processes were reported as a surrogate of flexion and extension with values ranging from 24.6mm to 35.6mm for flexion, 19.9mm to 29.4mm for extension and segmental ROM estimates, from neutral of 1.8 to 4.9mm for flexion and 0.9 to 4.3mm for extension (Chleboun et al. 2012). Actual segmental ROM values for right side bending were reported between 8.5º and 17.3º depending on the segment (Mahato et al. 2019).

1.1.3.2.3.2 Reliability

Regarding reliability, a synthesis of the studies is difficult due to a large degree of heterogeneity evident in the methodology.

Chleboun et al. (2012) utilised supine positioning with wedge placement to induce extension and flexion and three measures were taken without moving from each position. This method is unlikely to achieve full ROM and it also removes all biological variation due to repeated movement. As a result, reliability estimates were excellent (ICC = 0.98, CI95% = 0.95-0.99; CoV = 1.6%) (Chleboun et al. 2015). Landel et al. (2008) completed a prone MRI during manual palpation and 'accessory spinal mobility assessment'. They quantified the curvature change during 'posterio-anterior' pressure. However, they did not report any actual values of curvature. Reliability estimates for intervertebral curvature in prone for five participants across two visits were excellent (ICC 0.95-0.99; SEM 0.40º to 0.66º) (Landel et al. 2008). Mahato et al. (2019) completed MRI during right side bending between days. ICC estimates for between-day reliability of segmental ROM (side bending) were excellent 0.93-0.94 and CoV values at 14-15% (Mahato et al. 2019).

In summary, regardless of the methods employed, it appears that MRI for segmental ROM measurements is highly reliable in both the sagittal and frontal plane for end of range static positions. Despite this, the CoV seems to depend on the movement being

measured and the method of analysis. Similarly, the effect of different assessors and of true repeated movements is not clear.

1.1.3.2.4 Radiography

1.1.3.2.4.1 Segmental ROM values

Since the aim of this review was to investigate reliability, the search of radiography papers was limited to those investigating this psychometric outcome. As a result, the citations included in this review are not inclusive of the exhaustive list of radiography studies that report segmental ROM values. Readers interested in this area are directed to papers such as Yukawa et al. (2019) and Galbusera et al. (2021).

Measurements of lumbar segmental ROM from radiographs varied between the included studies. Three studies reported at least one plane of lumbar segmental rotation including side bending and rotation (Hass et al. 1990) and flexion-extension (Maigne et al. 2003; Pearson et al. 2011). Individual segmental ROM values were not reported in three studies (Haas et al. 1990; Cakir et al. 2006a; Cakir et al. 2006b).

Using similar conceptual methods, segmental ROM was quantified from flexion-extension radiographs in two studies by reporting the angle change between adjacent vertebral endplates (Maigne et al. 2003; Pearson et al. 2011). Pearson et al. (2011) found an average change in intervertebral rotation of 5.1º and 5.7º for the digitised manual technique (DMT) and computer-assisted quantitative motion analysis (QMA) method respectively. However, it is not known whether these results are in relation to weight bearing or recumbent postures as they did not detail the positioning of participants during imaging.

Maigne et al. (2003) also reported segmental ROM values but in sitting and standing positions of participants with chronic LBP. Some had pain that occurred immediately on sitting down which was relieved on standing up (patient group) and participants who did not have these symptoms were matched to the patient group based on age and gender (control group) (Maigne et al. 2003). Angular motion (AM) for positional change from extension to flexion was 13.9º ± 4.5º (patient group) and 7.5º ± 4.3º (control group) (Maigne et al. 2003). Similar values were seen for positional change

from extension to sitting (AM = 10.0º ± 4.5º (patient group); 6.2º ± 4.0º (control group)) (Maigne et al. 2003). It is not immediately clear why such large ROM was observed in the patient group. However, the influence of LBP on the motion in this group could offer explanation as well as, the sample being largely female and the method's specific focus on achieving maximal ROM.

1.1.3.2.4.2 Intra-rater reliability of segmental ROM measurement

In radiography research, reliability analysis usually involves one or several raters measuring segmental ROM from the same radiographs on one or multiple occasions. However, due to variability in methodology and presented reliability statistics, synthesis of the studies included is difficult.

Using two raters and two measurement methods, Cakir et al. (2006a) and Cakir et al. (2006b) investigated the intra-rater reliability of measurements from standing flexion-extension radiographs, with measurements taken from the same images on two separate occasions. Intra-rater reliability estimates for segmental ROM were reported as strong for measurements made by the same rater using the same method (PCC = 0.782-0.916) with small mean differences between the two measurements (-0.17º to 0.04º) (Cakir et al. 2006a). However, the 95% confidence intervals for these differences ranged from ±4.0º to ±6.8º) (Cakir et al. 2006a) suggesting that despite a small mean, there was a large range of differences between two measurements.

Similar outcomes were observed for their follow up study where the method was adapted to measure the intervertebral segment which had received a total disc replacement (Cakir et al. 2006b). Strong intra-rater reliability estimates (PCC = 0.903-0.962) with small mean differences (-0.08º to 0.08º) were reported but there were wide confidence intervals between these two measurements (±2.0º to ±3.3º) (Cakir et al. 2006b)

In Pearson et al. (2011) study, 30 flexion-extension radiographs were measured twice by six raters, over a four week period, using either the DMT or QMA method. Intervertebral rotation intra-rater reliability ICCs were higher for the QMA method (ICC = 0.997) with small SEM (0.5º) compared to the DMT (ICC = 0.870, SEM = 2.5º) (Pearson et al. 2011).

For end-plate angle in extension, flexion and sitting, Maigne et al. (2003) analysed the intra-rater reliability of one rater extensively by opting to investigate if there was a difference between repeated measurements. They determined no significant difference between repeated measurement of the same images, reporting that the mean difference between two measurements was ≤0.31º (Maigne et al. 2003). However, the 95% confidence interval for the limits of agreement between the measures was at best -3.0º to 2.4º suggesting significant variability between repeated measures (Maigne et al. 2003).

1.1.3.2.4.3 Inter-rater reliability of segmental ROM measurement

Inter-rater reliability estimates for segmental ROM of flexion-extension radiographs amongst two raters were reported as strong for measurements between raters using the same method (PCC = 0.738-0.929), with a small mean difference (-0.82º to -0.07º) (Cakir et al. 2006a). However, as observed before, the range of difference between two raters was large; yielding a wide 95%CI (-7.4º to 5.8º) (Cakir et al. (2006a). Similar findings were observed in their follow up study with a strong correlation between raters (PCC = 0.886-0.950) and small mean difference (-0.31º to 0.04º) but large confidence intervals (-3.0º to 3.1º) (Cakir et al. 2006b).

Inter-rater reliability of flexion-extension radiographs was further studied by Maigne et al. (2003) and Pearson et al. (2011). Estimates provided by Maigne et al. (2003) demonstrated mean differences between two raters measurements in extension, flexion and sitting was -0.38º to -1.05º. However, the 95%CI between raters was -3.1º to 4.8º for end-plate angle, suggesting wide variability in the differences (Maigne et al. 2003). Pearson et al. (2011) found inter-rater reliability estimates that were excellent for measurements made with the QMA method (ICC = 0.976) compared to the DMT that yielded moderate results (ICC = 0.693).

Haas et al. (1990) investigated tilt into side bending and rotation in standing and lateral bending positions using three examiners reporting a range of Kappa reliability estimates. For side bending, agreement between raters was reported as weak-to-moderate in neutral (Kappa = 0.17-0.56) for L1-L5 (Haas et al. 1990). This was also true for net segmental tilt in left lateral bending (LLB) (Kappa = -0.03-0.50) and right lateral bending (RLB) (Kappa = 0.00-0.27) excluding the measurement at L3 for LLB

that showed excellent reliability (Kappa = 1.00) (Haas et al. 1990). However, the inter-rater reliability estimates were better overall for rotation results which yielded moderate-to-good results in neutral radiographs (Kappa = 0.55-0.68) and weak-to-good results in LLB and RLB (Kappa 0.38-0.68) (Haas et al. 1990). Interestingly, reliability estimates at L5 were low across all three raters in neutral, LLB and RLB for tilt (Kappa = 0.16-0.19) and rotation (Kappa = 0.38-0.57) (Haas et al. 1990).

The SEM was also reported by Haas et al. (1990). They found the mean absolute discrepancy was <2º for tilt and <3º for rotation of neutral radiographs at all lumbar levels (Haas et al. 1990). This was less than half of the expected measurement error which was also true for net tilt (1.2º to 3.2º) and rotation (2.0º to 3.7º) in LLB and RLB (Haas et al. 1990). However, even though all measurement errors were reported as low, data were only presented from one rater pair (Haas et al. 1990).

Overall, results for radiography indicate that there is high intra-rater reliability between measurements made using the same method, and differing methods, in flexion-extension. This also appears true for inter-rater reliability in flexion-extension as well as, lateral bending radiographs. However, variability in the results suggest reliability could be affected by the selected method for measuring ROM from the radiographs. Moreover, the magnitude of the variability across 2 measurements of the same image should be considered when assessing the expected ROM alteration from interventions such as surgery.

Summary of findings – Table 6

Author	**Findings**
Haas et al. (1990)	Kappa (K): • Tilt (neutral) – L1: K = 0.47, L2: K = 0.56, L3: K = 0.46, L4: K = 0.22, L5 = 0.17. • Tilt (left lateral bending) – L1: K = 0.50, L2: K = -0.03, L3: K = 1.00, L4: K = 0.25, L5 = 0.19. • Tilt (right lateral bending) – L1: K = 0.24, L2: K = 0.25, L3: K = 0.00, L4: K = 0.27, L5: K = 0.16. • Rotation (neutral) – L1: K = 0.64, L2: K = 0.68, L3: K = 0.55, L4: K = 0.60, L5: K = 0.57. • Rotation (left lateral bending) – L1: K = 0.63, L2: K = 0.61, L3: K = 0.57, L4: K = 0.55, L5: K = 0.42.

	• Rotation (right lateral bending) – L1: K = 0.49, L2: K = 0.42, L3: K = 0.59, L4: K = 0.68, L5: K = 0.38. <u>SEM (one rater pair):</u> • Neutral radiographs – 1.0º - 1.9º (tilt); 2.1º - 2.6º (rotation). • Net tilt - 1.3º - 3.2º (left lateral bending); 1.2º - 2.8º (right lateral bending). • Net rotation – 2.0º - 3.7º (left lateral bending); 2.5º - 3.6º (right lateral bending).
Maigne et al. (2003)	<u>Angular motion:</u> • Extension to flexion = 13.9º ± 4.5º (patient group); 7.5º ± 4.3º (control group). • Extension to sitting = 10.0º ± 4.5º (patient group); 6.2º ± 4.0º (control group). <u>Intra-rater (mean difference between end plate angle measurements):</u> • 0.31º, 95%CI LoA -3.0º to 2.4º (extension); 0.04º, 95%CI LoA -3.0º to 3.0º (flexion); 0.03º, 95%CI LoA -3.0º to 3.0º (sitting). <u>Inter-rater (mean difference between end plate angle measurements):</u> • -0.38º, 95%CI LoA -3.1º to 3.9º (extension); -0.44º, 95%CI LoA -2.7º to 3.6º (flexion); -1.05º, 95%CI LoA -2.7º to 4.8º (sitting).
Breen et al. (2006)	• Side bending intra-subject variation 2.75º RMSE (observer 1). • Side bending intra-subject variation 2.91º RMSE (observer 2). • Raw data ranges not reported.
Cakir et al. (2006a)	<u>Intra-rater reliability:</u> • PCC (95%CI) = 0.902 (±4.2º), 0.782 (±6.8º), 0.916 (±4.0º), 0.881 (±4.7º). • Mean difference = -0.17º, 0.04º, -0.17º, -0.17º. <u>Inter-rater reliability:</u> • Range of PCC (95%CI) = 0.843, 0.809, 0.777, 0.738 (-7.4º/ +5.8º); 0.929, 0.913 (-4.5º/ +4.3º); 0.890, 0.861, 0.890, 0.891 (-4.9º/ +4.5º); 0.885, 0.888 (-5.0º/ +4.2º). • Mean difference between 2 measurement sets = -0.82º, -0.07º, -0.17º, -0.38º.
Cakir et al. (2006b)	<u>Intra-rater reliability:</u> • PCC (95%CI) = 0.962 (±2.1º), 0.903 (±3.3º), 0.955 (±2.0º), 0.916 (±3.0º). • Mean difference = 0.04º, 0.08º, -0.08º, -0.04º. <u>Inter-rater reliability:</u> • Range of PCC (95%CI) = 0.928, 0.903, 0.911, 0.917 (-3.0º/ +3.0º); 0.918, 0.905 (-2.9º/ +3.1º); 0.899, 0.930, 0.950, 0.950 (-2.4º/ +3.0º); 0.926, 0.886 (-2.8º/ +2.8º). • Mean difference between 2 measurement sets = 0.04º, -0.06º, -0.31º, -0.00º.
Landel et al. (2008)	• SEM ranged from 0.40º to 0.66º. • ICC 0.95- 0.99.
Pearson et al. (2011)	• Average change in intervertebral rotation = 5.1º (DMT), 5.7º (QMA).

	• Intra-rater – ICC = 0.870, SEM = 2.5º (DMT); ICC =0.997, SEM = 0.5º (QMA). • Inter-rater – ICC = 0.693 (DMT), 0.976 (QMA).
Sui et al. (2011)	• Did not report individual segmental ROM values. • Only reported ICC for the plastic spine model.
Chleboun et al. (2012)	<u>US:</u> • Spinous process distance in flexion = 25.6mm to 32.3mm. • Spinous process distance in extension = 21.5mm to 26.9mm. • Flexion ROM taken from neutral = 3.0mm to 4.4mm. • ICC 0.94, CI95% 0.85-0.97. • CoV = 1.8%. <u>MRI:</u> • Spinous process distance = 24.6mm to 35.6mm (flexion), 19.9mm to 29.4mm (extension). • Segmental ROM estimates from neutral = 1.8mm to 4.9mm (flexion), 0.9mm to 4.3mm (extension). • ICC 0.98, CI95% = 0.95-0.99. • CoV = 1.6%.
Mellor et al. (2014a)	<u>Segmental ROM:</u> • Flexion – L2/3 = 4.23º, L3/4 = 5.89º, L4/5 = 7.10º (Patients). • Flexion – L2/3 = 4.05º, L3/4 = 5.49º, L4/5 = 6.46º (Controls). • Extension – L2/3 = 5.04º, L3/4 = 4.15º, L4/5 = 4.78º (Patients). • Extension – L2/3 = 4.64º, L3/4 = 4.11º, L4/5 = 5.31º (Controls). <u>Intra-observer:</u> • SEM% = 3% (flexion); 8% (extension). • L2/3 = SEM 0.13º, ICC 0.98, CI95% 0.86-0.99 (flexion); SEM 0.35º, ICC 0.96, CI95% 0.85-0.99 (extension). • L3/4 = SEM 0.13º, ICC 0.98, CI95% 0.90-1.0 (flexion); SEM 0.24º, ICC 0.92, CI95% 0.72-0.98 (extension). • L4/5 = SEM 0.10º, ICC 1.0, CI95% 0.99-1.0 (flexion); SEM 0.19º, ICC 0.99, CI95% 0.97-1.0 (extension). • SEM 0.08-0.17º, ICC 0.99-1.0, CI95% 0.95-1.0, SEM % = 3% (lateral bending). <u>Inter-observer:</u> • ICC 0.74-0.99; CI95% 0.23-0.99. • SEM 0.17-0.31º; SEM% = 7% (flexion) • SEM 0.27-0.77º; SEM% = 19% (extension) • L2/3 = SEM 0.31º, ICC 0.91, CI95% 0.69-0.98 (flexion); SEM 0.77º, ICC 0.76, CI95% 0.27-0.94 (extension) • L3/4 = SEM 0.17º, ICC 0.98, CI95% 0.91-0.99 (flexion); SEM 0.41º, ICC 0.74, CI95% 0.23-0.93 (extension). • L4/5 = SEM 0.31º, ICC 0.97, CI95% 0.88-0.99 (flexion); SEM 0.27º, ICC 0.99, CI95% 0.96-1.0 (extension). • SEM 0.18-0.55º, ICC 0.85-0.99, CI95% 0.51-1.0 (lateral bending).
Yeager et al. (2014)	• Segmental ROM: 4.40º (flexion) and 2.00º (extension).

	• Intra-observer: SEM 0.10º, ICC 0.98, CI95% 0.98-0.99 (flexion/ extension, SEM% = 2% (flexion), 5% (extension). • Inter-observer: SEM 0.22º, ICC 0.96, CI95% 0.95-0.97 (flexion/ extension), SEM% = 5% (flexion), 11% (extension).
Cuestas-Vargas (2015)	• Segmental ROM was 15.5º +/- 2.04º (flexion), SEM = 0.54º. • Repeated measures: ICCs CI95% = 0.995-0.999 (within day), CI95% = 0.996-0.999 (between day). • MDC_{95} = 1.5º (or 10%).
Tozawa et al. (2015)	• Lumbar interspinous process distance ranged from 29.2mm to 30.1mm. <u>Intra-rater reliability:</u> • ICC CI95% 0.79-1.0. • Examiner A: ICC 0.97-0.98, CI95% 0.93-0.99. • Examiner B: ICC 0.96-0.98, CI95% 0.92-0.99. • Examiner C: ICC 0.97-0.98, CI95% 0.94-0.99. • Examiner D: ICC 0.97-0.99, CI95% 0.94-1.0. • Examiner E: ICC 0.90-0.99, CI95% 0.79-1.0. • MDC_{95} value of 0.29mm. <u>Inter-rater reliability:</u> • ICC 0.914, CI95% 0.80-0.97; ICC 0.725, CI95% 0.55-0.87.
du Rose and Breen (2016)	<u>Between day Intra-rater reliability:</u> • L2/3 = SEM 0.45º, ICC CI95% 0.92-1.0. • L3/4 = SEM 0.23º, ICC CI95% 0.96-1.0. • L4/5 = SEM 0.39º, ICC CI95% 0.97-1.0. • L5/S1 = SEM 0.54º, ICC CI95% 0.82-0.99.
Tozawa et al. (2016)	<u>Intra-rater reliability:</u> • (ICC 0.990-0.998, CI95% 0.963-0.999. • Measurer A: ICC 0.997, CI95% 0.993-0.999. • Measurer B: ICC 0.992, CI95% 0.981-0.998. • Measurer C: ICC 0,998, CI95% 0.996-0.999. • Measurer D: ICC 0.985, CI95% 0.963-0.996. • Measurer E: ICC 0.991, CI95% 0.978-0.997. • Measurer F: ICC 0.995, CI95% 0.987-0.998. • Measurer G: ICC 0.995, CI95% 0.987-0.999. • Measurer H: ICC 0.992, CI95% 0.980-0.998. • Measurer I: ICC 0.990, CI95% 0.977-0.997. • MDC_{95} values of 0.62-1.8mm. • Measurer D: MDC_{95} = 1.8mm. • Measurer F: MDC_{95} = 1.1mm. • Measurer H: MDC_{95} = 0.62mm. • Measurer I: MDC_{95} = 1.5mm. <u>Inter-rater reliability:</u> • ICC 0.969, CI95% 0.90-1.0.
Breen et al. (2018)	<u>Segmental ROM:</u> • Lying: 5.14º (flexion), 5.33º (extension).

	• Standing: 9º to 14º (flexion), 2.01º (extension). Repeated measures: • ICC 0.80, CI95% 0.74-0.85 (lying flexion and extension). • ICC 0.82-0.91, CI95% 0.76-0.93 (standing flexion and extension). • ICC 0.95, CI95% 0.92-0.96 (lying lateral bending). • ICC 0.90-0.92, CI95% 0.0.87-0.94 (standing lateral bending). • MDC_{95} 4.66º, MDC% 91% (flexion in lying). • MDC_{95} 5.19º, MDC% 97% (extension in lying). • MDC_{95} 3.3-3.7º, MDC% 60-69% (lateral bending in lying). • MDC_{95} 9.10º, MDC% 100% (flexion in standing). • MDC_{95} 5.53º, MDC% 176% (extension in standing). • MDC_{95} 4.5-4.7º, MDC% 97-98% (lateral bending in standing).
Mahato et al. (2019)	Segmental ROM: • 8.5º to 17.3º (right lateral bending). Repeated measures: • ICC 0.93-0.94 (right lateral bending between day). • CoV 14-15% (right lateral bending between day).

(K, kappa; L, lumbar; =, equals; SEM, standard error of measurement; º, degrees; 95%CI, ninety five percent confidence interval; LoA, limits of agreement; RMSE, root mean square error; PCC, pearson correlation coefficient; ±, plus or minus; ICC, intra-class correlation co-efficient; DMT, digitised manual technique; QMA, quantitative motion analysis; US, ultrasound; mm, millimetres; ROM, range of motion; CI95%, ninety five percent confidence interval; CoV, coefficient of variation; MRI, magnetic resonance imaging; L2/3, lumbar spine intervertebral level 2/3; L3/4, lumbar spine intervertebral level 3/4; L4/5, lumbar spine intervertebral level 4/5; SEM%, standard error of measurement percent; %, percent; L5/S1, spine intervertebral level lumbar 5/ sacral 1; MDC_{95}, minimal detectable change at 95% confidence level; MDC%, percentage magnitude of minimal detectable change at 95% confidence level).

1.1.4 Discussion

1.1.4.1 Modality evaluation

1.1.4.1.1 VF

VF provides a cost-effective, non-invasive (Manninen et al. 1988) method for segmental ROM assessment that can provide dynamic or static quantification of ROM and is often completed in a weight-bearing position. However, there is a tricky trade-off between radiation dose and image quality (Lam et al. 2009). Since low radiation dosage is used (Muggleton and Allen 1997); the contrast between the vertebrae and surrounding soft tissue is very low (Lam et al. 2009), making identification of anatomical landmarks difficult (Muggleton and Allen 1997). Furthermore, although radiation dose for VF of the lumbar spine compares favourably with exposures for a

single plain radiograph of the same region (Muggleton and Allen 1997; Harvey et al. 2016); the risks associated with radiation exposure (Kim et al. 2013) remain present.

As mentioned previously, manual mark-up of VF images remains necessary (Davis et al. 2015) but differences in mark-up practices exist (Harvey et al. 2016). Currently, there is no consensus as to which is the most effective (Harvey et al. 2016). What is more, it is a laborious and time consuming process (Muggleton and Allen 1997; Lam et al. 2009) that remains a source of error (Plocharski et al. 2018), and the choice of anatomical landmark identification can greatly influence the results (Harvey et al. 2016). Moreover, the optical distortion and out of plane motions (Harvey et al. 2016) are likely to pose significant challenge to the clarity of VF images and ultimately, its usefulness in quantifying segmental ROM.

1.1.4.1.2 US

US imaging is a safe, inexpensive modality (Heidari et al. 2015) which is portable, offering easy collection of static and dynamic images (Marshburn et al. 2014). Though there are no known deleterious effects of US it remains the domain of competent sonographers (The World Federation for Ultrasound in Medicine and Biology 2013).

Whilst it isn't commonplace to US image the spine, there is evidence that nearly all structures within the spine are visible with US (Ahmed et al. 2018). However, despite adequate visualisation of structures being outlined by Ahmed et al. (2018), the skill of completing US scanning largely remains operator dependent. For example, Margarido et al. (2010) showed 20 unsupervised trials plus teaching sessions were not enough for participants to achieve competence in different aspects of US assessment of the lumbar spine. Therefore, if US imaging was to become more routine for assessing segmental ROM of the spine; specific training may be necessary. Furthermore, as US machines evolve, enhancements in image quality are further likely to facilitate easier imaging of the spine (Ahmed et al. 2018).

In comparison to other modalities, field of vision is small with US and directly limited to the area beneath the US transducer (Hides et al. 1998). Also, distinct individual characteristics, such as BMI, are likely to affect the image quality; meaning this modality may not be universally appropriate (Hides et al. 1998). Despite this, real time

analysis, video capture and enhancements to the technology and its image processing are likely solutions.

In summary, US is an inexpensive, safe and accessible modality that is already used extensively in clinical practice for other purposes. Therefore, it affords great potential for regular monitoring of lumbar spinal ROM. Nevertheless, it requires a skilled operator to image the lumbar spine and resolution of images may vary between patients based on extraneous patient variables or sonographer expertise.

1.1.4.1.3 MRI

MRI uses non-ionising radiation (Wassenaar et al. 2012), is non-invasive (Sett and Crockard 1991) and is considered a safe technology (Hartwig et al. 2009). Furthermore, it offers real advantages in terms of image quality, resolution and consistency (Wassenaar et al. 2012; Ahmed et al. 2018). MRI has the ability to visualise the entire spine, spinal cord and surrounding structures in its entire length (Sett and Crockard 1991); providing further opportunities such as, the identification of structural changes. Moreover, MRI can produce sectional images of equivalent resolution, in any projection, without moving a patient (NHLBI ca.2020). This ability to obtain images in multiple planes adds to its versatility (NHLBI ca.2020).

Analysis of spinal ROM requires the use of open MRI which eliminates a patients feeling of claustrophobia, along with the associated implications of this effect, commonly seen with traditional closed MRI scanners (Tarantino et al. 2013; Michelini et al. 2018). However, it does have some disadvantages. This is represented mostly by the use of a low field magnet; resulting in low signal to noise ratio and leading to reduced image quality compared with the more common high field magnet (Michelini et al. 2018). Equally, patients with pacemakers and certain ferromagnetic appliances cannot be imaged with MRI (NHLBI ca.2020), and patient throughput is slow compared with other imaging modalities (Michelini et al. 2018; NHLBI ca.2020).

A further significant drawback to MRI is that the equipment is not only expensive to purchase, but also to maintain and operate (NHLBI ca.2020). Additionally, greater technological expertise is required for utilisation of MRI rather than most other imaging modalities (NHLBI ca.2020); highlighting important limitations.

Altogether, it is evident that MRI has good spinal visualising capabilities; coupled with consistency in image acquisition and interpretation. What is more, this modality does not pose a risk to most patients and offers the clinical advantage of looking at intervertebral disc deformation and soft tissue providing additional insights for patients with known pathologies. However, the substantial cost associated with using this technology indicates its lack of suitability for regular monitoring of lumbar segmental ROM.

1.1.4.1.4 Radiography

Radiography remains a cost effective spinal imaging method (Oakley and Harrison 2018) and the equipment is widely available (Janssen et al. 2011). Compared to other imaging modalities, like MRI, usually performed in the recumbent position; radiographs can also be taken in different anatomical positions (Oakley and Harrison 2018). Nonetheless, there are no established guidelines for imaging the thoraco-lumbar spine with radiographs (Leone et al. 2007; Janssen et al. 2011) and it is required to be performed in a specialised room (Janssen et al. 2011). There are also errors associated with distortion, magnification and positioning of individuals (Janssen et al. 2011; Mellor et al. 2014b). Furthermore, lots of heterogeneity exits in the methodology of radiographic segmental ROM measurements (Leone et al. 2007).

The most significant disadvantage to radiography though is its use of ionising radiation (Mellor et al. 2014b; Davey et al. 2014). This is a known mutagen that can increase the risk of diseases such as cancer (Logan et al. 2019). In addition, a higher beam energy is required due to the lumbar spines large x-ray attenuation and imaging of this area involves exposure to radiosensitive reproductive organs (Lai et al. 2020). These risks are an important consideration for repeated radiography examinations.

To summarise, however cost-effective radiography remains, the errors linked to image capture and variability in image analysis, coupled with the risks associated with ionising radiation exposure; makes this imaging modality unsuitable for frequent assessment of lumbar spinal ROM.

1.1.5 Conclusion

This review has provided a contemporary systematic analysis of the literature related to the reliability of VF, US, MRI and Radiography modalities currently used for non-invasive measurements of segmental ROM in the lumbar spine in-vivo. Excellent reliability is seen in all modalities. However, VF is limited by radiation exposure, as is radiography, and there is a high cost associated with MRI. Additionally, these modalities are not routinely available. Although US scanning is operator dependent and specific training may be required, it offers potential for routine clinical use due to its low cost and widespread availability. Therefore, US has the opportunity to provide a truly non-invasive and risk free method of measuring segmental ROM in individuals with LBP. Despite this, further research is necessary to determine whether US imaging yields truly consistent measurements of segmental ROM in the lumbar spine and whether this is also evident in within- and between-day repeated measures. If a method of segmental ROM assessment can be developed for routine clinical practice it could be a useful tool to evaluate abnormal segmental motion due to pain, spinal pathology or surgical intervention; signifying its potential value in the assessment, diagnosis and management of a variety of spinal related conditions.

1.1.6 Summary

This systematic review highlighted that further research into the quantification of segmental ROM measurements with US imaging is warranted. It also indicated that the application of such a technology to solve a clinical problem is lacking. Therefore, the provisional aim of this thesis was to develop a method using US imaging coupled with an electromagnetic tracking device to quantify spinal segmental ROM and curvature in-vivo and also, develop this application for individuals with altered spinal curvature i.e. scoliosis. Significant development and proof of concept work in this direction was completed. However, it became clear that due to the reliance on an ultrasonographer and the applicability of US imaging in routine physiotherapy clinical practice, it wasn't going to be possible. The reasons lay in extensive training required for a physiotherapist to appropriately operate the US machine as well as, ethical issues surrounding observation (or failed observation) of sinister pathology by coincidence on US images. Therefore, a slight change in the direction of this project occurred, whereby the focus shifted to exploring spinal and trunk shape using surface measures rather than imaging modalities. This has the potential to overcome limitations in current methods through the inclusion of trunk shape, enabling the method to be applied to clinical conditions with alterations to trunk shape i.e. scoliosis. To determine the current gap in the literature, there was a need for a review of research pertaining to this topic.

The following section presents a scoping review of the literature related to the reliability of surface curvature measurement devices for assessment of the thoracolumbar spine. Although this review focuses on reliability, each measurement system and its methods, through discussion of its advantages and limitations, are also evaluated.

1.2 The reliability of surface measurement methods for assessing spinal curvature in the thoracolumbar spine: A scoping review

1.2.1 Introduction

Spinal disorders remain one of the most common and costly complaints in clinical practice (Martin et al. 2008). Previous research has shown that trunk posture, kinematics and function is impacted by spinal disorders such as stenosis (Kuwahara et al. 2016), scoliosis (Schmid et al. 2015) and also by back pain (Christe et al. 2017; Chun et al. 2017). As a result, measurement of spinal curvature has become common place in the clinical assessment of the spine (Vrtovec et al. 2009). This can provide useful information about spinal function and can also be used as a clinical outcome measure to assess the impact of disease or to evaluate evidence-based treatments (Plaszewski et al. 2012).

Currently, static X-ray examination is the gold standard for measuring spinal curvature (Vrtovec et al. 2009; Ghandhari et al. 2020). Despite its confirmed validity, risks associated with repeated radiation exposure, availability and expense mean regular radiographic assessment is not recommended (Briggs et al. 2007). For that reason, surface curvature measurement devices and methods have been developed to overcome the limitations of X-ray. However, for consideration as a clinical alternative to X-ray, the reliability and validity of these measurement devices need to be established. Prior to establishing validity, reliability should be determined. Therefore, the aim of this scoping literature review is to: 1) explore existing devices and methods of spinal surface curvature measurements, 2) understand the advantages and limitations to these measurement methods, and 3) review reliability estimates of such methods.

1.2.2 Materials and Methods

The five-stage framework outlined by Arksey and O'Malley (2005) underpinned the approach to this scoping review. The focus of this review was exploring existing surface curvature measurement methods and their reliability in-vivo. The aims outlined above guided the search and facilitated the scoping capture of literature relating to this topic.

1.2.2.1 Search strategy

In November 2020, a systematic literature search of 16 electronic databases including: MEDLINE Complete, Academic Search Ultimate, CINAHL Complete, Complementary Index, SPORTDiscus and ScienceDirect, was conducted. Search terms, using Boolean logic (as seen in Table 7), were used and the search was limited to peer-reviewed articles published in the English Language. In addition, to identify citations within grey literature; reference lists of eligible articles identified through database searching were hand searched for further appropriate studies.

Table 7. Search strategy

Search Terms
Measure OR measurement OR measuring
AND spinal OR spine OR thoracolumbar
AND curvature OR shape
AND reliability
NOT radio* OR ultra* OR x-ray OR mri OR fluro*

(*, represents truncation)

Table 8. Inclusion and exclusion criteria

Inclusion criteria	**Exclusion criteria**
• Researching thoracic and/ or lumbar regions. • Curvature or shape spinal measurements. • Portable surface curvature methods. • In-vivo measurements. • Quantitative papers. • Any population type.	• Methods using imaging modalities including: ultrasonography, radiographic, magnetic resonance imaging and fluoroscopy. • Investigating only cervical spine region. • Not assessing reliability of measurement method. • Laboratory based/ non-portable methods. • In-vitro measurements. • Studies published after November 2020.

1.2.2.2 Inclusion and exclusion criteria

Firstly, articles were screened by title, abstract then full text; against inclusion and exclusion criteria (Table 8). A detailed flow chart of the search, including the number of citations yielded at each stage, is shown in Figure 2. Table 9 details the reasons for full-text article rejection.

Table 9. Reason for full-text article rejection

Reason for full-text citation rejection	Number of citations in this category
Not assessing reliability	4
Not researching in-vivo	1
Duplicate paper	3
Only abstract available	2
Laboratory-based/ non-portable method	3
Total number of full text citations rejected	13

1.2.2.3 Data charting and collation

For each citation, details relating to the study's author(s), year published, measurement device/ technology, participant, methodology and reliability results were collated into Table 10.

1.2.3 Results

A total of 20 studies were eligible for this review (Hart and Rose 1986; Lovell et al. 1989; Youdas et al. 1995; Hinman 2004b; Mannion et al. 2004; Dunleavy et al. 2010; Lewis and Valentine 2010; Singh et al. 2010; Williams et al. 2010; Czaprowski et al. 2012; de Oliveira et al. 2012; Williams et al. 2012; Gonzalez-Sanchez et al. 2014; MacIntyre et al. 2014; Topiladou et al. 2014; Livanelioglu et al. 2016; Sedrez et al. 2016; Was et al. 2016; Quek et al. 2017; Roghani et al. 2017).

All studies investigated the reliability of spinal curvature or shape measurements in-vivo using a specific surface measurement device. The flexicurve was investigated by seven articles (Hart and Rose 1986; Lovell et al. 1989; Youdas et al. 1995; Hinman 2004b; Dunleavy et al. 2010; de Oliveira et al. 2012; Sedrez et al. 2016), the spinal

mouse by four (Mannion et al. 2004; Topiladou et al. 2014; Livanelioglu et al. 2016; Roghani et al. 2017) and the inclinometer by another four citations (Lewis and Valentine 2010; Czaprowski et al. 2012; MacIntyre et al. 2014; Was et al. 2016). Williams et al. (2010) and Williams et al. (2012) investigated a fibre-optic measurement device; whilst Singh et al. (2010) and Gonzalez-Sanchez et al. (2014) used electromagnetic (EM) tracking sensors and lastly, Quek et al. (2017) looked into the reliability of a three dimensional (3-D) depth camera (Kinect) for curvature measurements.

'Normals' were investigated most by eligible studies (Hart and Rose 1986; Youdas et al. 1995; Mannion et al. 2004; Dunleavy et al. 2010; Singh et al. 2010; Williams et al. 2010; Czaprowski et al. 2012; Sedrez et al. 2016; Was et al. 2016; Quek et al. 2017). The next most common population studied was those with LBP (Lovell et al. 1989; Williams et al. 2012; Topiladou et al. 2014).

Articles measured a range of different positions and movements including standing (Hart and Rose 1986; Lovell et al. 1989; Youdas et al. 1995; Hinman 2004b; Mannion et al. 2004; Dunleavy et al. 2010; Lewis and Valentine 2010; Singh et al. 2010; Czaprowski et al. 2012; de Oliveira et al. 2012; Gonzalez-Sanchez et al. 2014; MacIntyre et al. 2014; Topiladou et al. 2014; Livanelioglu et al. 2016; Sedrez et al. 2016; Was et al. 2016; Quek et al. 2017; Roghani et al. 2017), sitting (Youdas et al.1995; Quek et al. 2017), flexion (Hart and Rose 1986; Youdas et al. 1995; Mannion et al. 2004; Williams et al. 2010; Williams et al. 2012; Topiladou et al. 2014; Roghani et al. 2017) and extension (Youdas et al. 1995; Mannion et al. 2004; Williams et al. 2012; Topiladou et al. 2014; Roghani et al. 2017).

Some articles measured only thoracic kyphosis (Lewis and Valentine 2010; Quek et al. 2017), others just lumbar lordosis (Hart and Rose 1986; Lovell et al. 1989; Youdas et al. 1995; Williams et al. 2010; Williams et al. 2012) and the rest; both thoracic and lumbar curvatures (Hinman 2004b; Mannion et al. 2004; Dunleavy et al. 2010; Singh et al. 2010; Czaprowski et al. 2012; de Oliveira et al. 2012; Gonzalez-Sanchez et al. 2014; MacIntyre et al. 2014; Topiladou et al. 2014; Livaneliogu et al. 2016; Sedrez et al. 2016; Was et al. 2016; Roghani et al. 2017).

Altogether, it appears that across the literature, many methods have been explored in quantifying curvature in both the thoracic and lumbar spine. Furthermore, research has investigated these methods in static positions i.e. sitting and standing but also during sagittal motion in individuals without pain or pathology and in those with back pain.

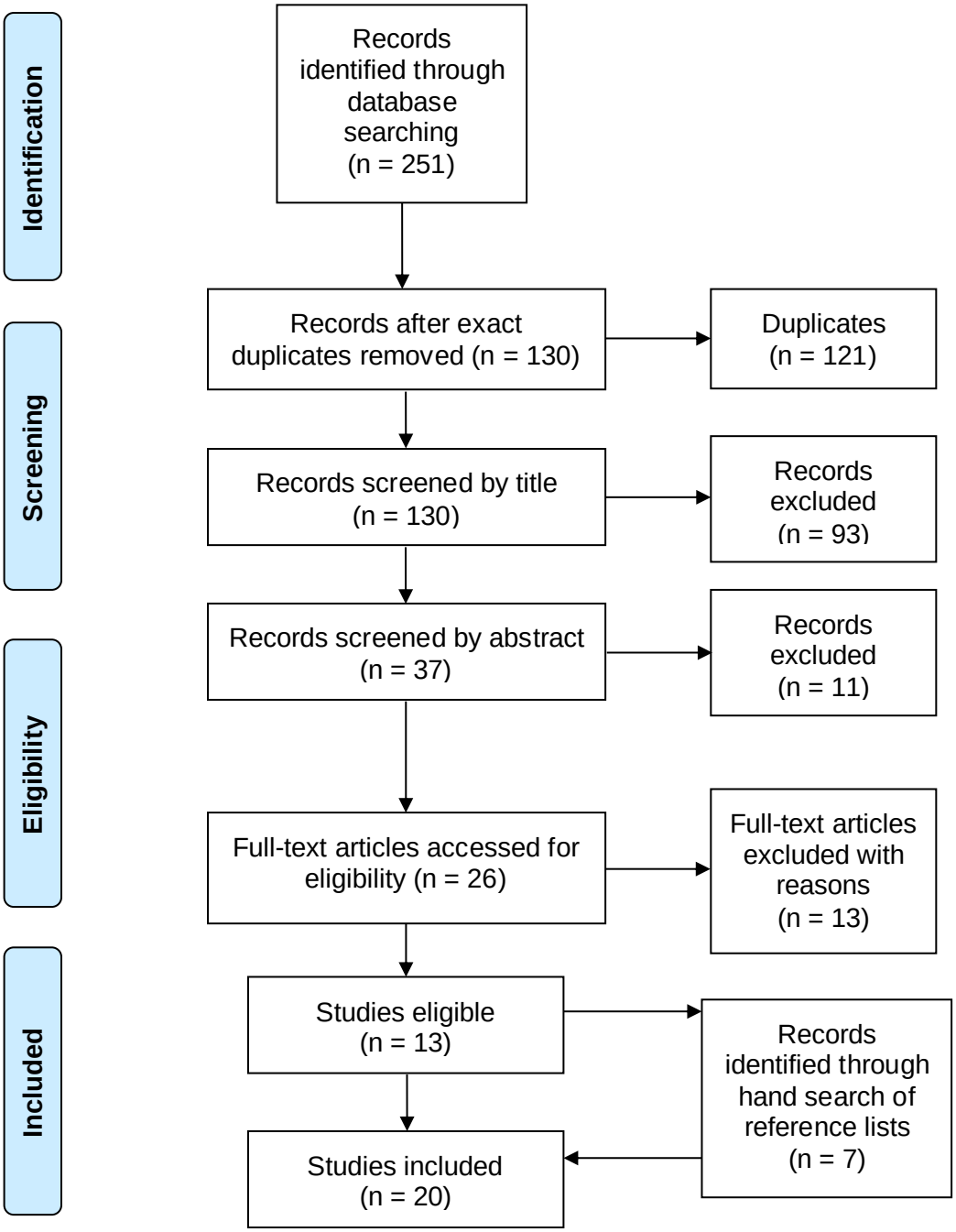


Figure 2. PRISMA flow diagram of search (n, number; =, equals)

Table 10. Data charting

Author(s)	Surface Measurement Device/ Technology	Methods	Reliability Results
Hart and Rose (1986)	Flexicurve	23 healthy adult participants. Quantified shape of Lx spine. Measured L1-S2. 2 sagittal measurements made for each position studied. Positions: normal standing, extreme of forward bending, forward bending with trunk curl and forward bending with back straight. 23 pairs of standing measurements and 66 pairs of complete forward bending measurements. Degree of agreement between test-retest measurements of same measurer analysed.	Intra- test-retest reliability of standing and complete forward bending measurements ICC = 0.97.
Lovell et al. (1989)	Flexicurve	40 participants without LBP and 40 with LBP. Lx spine lordosis measured in 'relaxed' normal standing position. Two measurers. Measured L3-S2. 2 sagittal measurements taken within day by one tester with a 1-minute rest in between. Tester 2 completed same testing on participant 1-3 minutes after.	Intra-tester reliability: ICC = 0.84 NSLBP and 0.94 for SLBP (E1) and ICC = 0.73 NSLBP and 0.83 for SLBP (E2). Inter-tester reliability: ICC = 0.41 NSLBP and 0.50 SLBP (first measurements taken) and 0.54 NSLBP and 0.52 for SLBP (second measurements taken).
Youdas et al. (1995)	Flexicurve	10 healthy volunteers.	Tangent method -

		3 conditions: standing, sitting with maximum trunk forward bending and lying prone with maximum backward bending. 2 methods of sagittal curvature measurement: tangent or trigonometric.	Intra- tester reliability: ICC = 0.82 and 0.90 (standing lumbar lordosis), 0.90 and 0.95 (maximum lumbar flexion in sitting), 0.96 and 0.98 (maximum lumbar extension in prone). Trigonometric method – Intra- tester reliability: ICC = 0.87 and 0.93 (standing lumbar lordosis), 0.84 and 0.91 (maximum lumbar flexion in sitting), 0.97 and 0.98 (maximum lumbar extension in prone).
Hinman (2004b)	Flexicurve	25 pre- and 26 post- menopausal women included. 3 graduate students with no experience using flexicurve measured each participant. 2 positions: standing relaxed posture and maximum erect posture. Measured C7-S1. Index of kyphosis and lordosis calculated from sagittal width and length measurements of each curve.	Index of kyphosis inter-rater reliability: ICC = 0.94 (relaxed), 0.93 (erect). Index of lordosis inter-rater reliability: ICC = 0.60 (relaxed), 0.73 (erect).
Mannion et al. (2004)	Spinal Mouse	20 healthy participants. Measured C7- (approximately) S3. 3 test positions: Standing upright, maximal flexion, maximal extension. 2 examiners completed x 3 sagittal measurements of each position on all participants on two separate days.	Intra-rater reliability between day: Standing Tx curvature: ICC (95% CI) = 0.73 (0.43-0.89), SEM 4.2º E1; 0.88 (0.67-0.94), SEM 2.8º E2. Standing Lx curvature: ICC (95% CI) = 0.90 (0.75-0.96), SEM 2.5º E1; 0.92 (0.80-0.97), SEM 2.4º E2. Flexion Tx curvature: ICC (95% CI) = 0.67 (0.32-0.86), SEM 5.5º E1; 0.86 (0.67-0.94), SEM 3.9º E2. Flexion Lx curvature: ICC (95% CI) = 0.85 (0.66-0.94), SEM 3.2º E1; 0.91 (0.79-0.97), SEM 3.0º E2. Extension Tx curvature: ICC (95% CI) = 0.76 (0.47-0.90), SEM 6.2º E1; 0.78 (0.52-0.91), SEM 4.6º E2.

			Extension Lx curvature: ICC (95% CI) = 0.80 (0.53-0.92), SEM 4.5º E1; 0.78 (0.52-0.91), SEM 4.6º E2. Inter-rater reliability: Standing Tx curvature: ICC (95% CI) = 0.87 (0.70-0.95), SEM 2.7º D1; 0.83 (0.62-0.93), SEM 3.3º D2. Standing Lx curvature: ICC (95% CI) = 0.87 (0.69-0.95), SEM 2.8º D1; 0.93 (0.83-0.97), SEM 2.2º D2. Flexion Tx curvature: ICC (95% CI) = 0.83 (0.61-0.93), SEM 4.3º D1; 0.70 (0.38-0.87), SEM 5.2º D2. Flexion Lx curvature: ICC (95% CI) = 0.85 (0.65-0.94), SEM 3.7º D1; 0.87 (0.70-0.95), SEM 3.3º D2. Extension Tx curvature: ICC (95% CI) = 0.64 (0.27-0.84), SEM 7.0º D1; 0.79 (0.54-0.92), SEM 5.0º D2. Extension Lx curvature: ICC (95% CI) = 0.88 (0.72-0.95), SEM 3.1º D1; 0.90 (0.75-0.96), SEM 3.5º D2.
Dunleavy et al. (2010)	Flexicurve	22 healthy participants. One session of testing. 2 examiners. Measured C7-S1. Normal standing position tested 3 times for each participant. Measured total length, thoracic length, thoracic width, lumbar length and lumbar width in sagittal plane.	Intra-rater reliability: ICCs = 0.61-0.97. Inter-rater reliability: ICCs = 0.56- 0.72. Total length – SEM 0.58cm, MDC95 1.62cm. Thoracic length – SEM 2.00cm, MDC95 5.55cm. Thoracic width – SEM 0.48cm, MDC95 1.33cm. Lumbar length – SEM 1.93cm, MDC95 5.36cm. Lumbar width – SEM 0.35cm, MDC95 0.96cm.
Lewis and Valentine (2010)	Inclinometer	45 participants with no upper body symptoms and 45 with upper body symptoms. Measurements of thoracic kyphosis made in relaxed standing.	Intra-rater reliability Tx kyphosis: Average ICC (95% CI) = 0.97 (0.95-0.99), SEM from ICC average = 1º (pps without symptoms). Average ICC (95% CI) = 0.97 (0.94-0.98), SEM from ICC average = 1.7º (pps with symptoms).

		2 inclinometers used – 1 placed over 1st and 2nd Tx SPs and 1 placed over region of 12th Tx and 1st Lx SPs. Angle produced by each inclinometer was measured 3 times in succession in sagittal plane. Each set of 3 measurements made on two occasions by one rater.	
Singh et al. (2010)	Electromagnetic tracking	52 healthy participants. 'natural' spinal curvature measured in standing. Palpation and marking of spinous processes and PSISs prior to moving a plastic probe with EM sensor over spine to trace SPs between T1 and L3. Measurements taken 3 times consecutively by the same examiner.	Intra-tester reliability: Tx kyphosis ICC (95% CI) = 0.93 (0.88-0.95), SEM 1.57º; Lx lordosis ICC (95% CI) = 0.98 (0.96-0.98), SEM 1.51º. Lateral curvatures for thoracic and lumbar regions ICC = 0.75, SEM 1.04º and 0.85, SEM 0.56º respectively.
Williams et al. (2010)	Fibre-optic	13 healthy participants. Lx curvature during flexion and lifting movements measured. Curvature was derived from the angles between S1 and L1 or L3 tangents.	Excellent similarity for repeated measures of whole lumbar and lower lumbar spine curvatures. (CMC = 0.97-0.98)
Czaprowski et al. (2012)	Inclinometer	30 healthy participants. Measured anterior-posterior (sagittal) spinal curvatures in spontaneous standing position. Three investigators performed the measurement 3 times at each palpation identified level of interest (inter-observer repeatability). One investigator performed a measurement of sagittal plane spinal curvatures in every subject one week apart (intra-observer repeatability).	Measurements displayed good reliability. Cronbach's alpha coefficient = 0.87 for Lx lordosis angle 0.83 for Tx kyphosis angle. Intra-rater reliability: no statistically significance (P>0.05) differences between results in all parts of spine in first and second measurements by investigator I. Inter-rater reliability:

			No significant difference between raters for Tx kyphosis but significant difference revealed for lumbar lordosis. Measurement errors reported as 2.8-3.8º.
de Oliveira et al. (2012)	Flexicurve	47 participants undergoing prescribed x-ray examination. C7, T1, T12, L5 and S1 SPs identified through palpation and marked by one evaluator. Flexicurve measurement from C7-S1 taken by: three different evaluators on the same day and two different occasions by one evaluator. Standing position measured. Sagittal plane measurements.	Inter-rater reliability ICC (95% CI) = 0.94 (0.87-0.98) Tx and 0.83 (0.60-0.94) Lx curvature. Intra-rater reliability ICC (95% CI) = 0.83 (0.57-0.94) Tx and 0.78 (0.47-0.92) Lx curvature.
Williams et al. (2012)	Fibre-optic	20 acute and 20 chronic LBP participants. Curvature measured dynamically during flexion, lifting and extension movements. Measurements taken between S1 and L1 (whole lumbar spine) and S1 and L3 (lower lumbar spine). Three movement trials completed.	Fibre-optic method highly reliable in measuring both whole lumbar and lower lumbar curvatures (CMC = >0.81 and ICC >0.99.)
Gonzalez-Sanchez et al. (2014)	Electromagnetic tracking	36 participants. 2 groups: normal weight and obese. One tester. Upright relaxed position measured. Phase 1: examiner palpated and marked reference points T1, T8, L1, L5 and PSISs. These points were digitised. Same examiner measured reference points x3 consecutively.	Intra-rater reliability of spinal segment angles: Cronbach's alpha ranged from 0.90 to 0.92 in sagittal plane and 0.82 to 0.86 in coronal plane.

		Probe with EM attached was traced from L1 to L5 and repeated 3 times. Phase 2: global co-ordinate system transformed to local co-ordinate system. Intra-rater reliability examined.	
MacIntyre et al. (2014)	Inclinometer	36 participants from osteoporosis outpatient clinic. Anatomical landmarks of interest palpated and marked. One rater measured: lumbosacral angle, lumbar standing posture and thoracic standing posture in sagittal plane. Test-retest reliability examined between day measurements.	test-retest reliability: Lumbosacral angle ICC (95% CI) = 0.91 (082, 0.95), SEM 2.5º, MDC_{90} 5.8º. Lumbar angle ICC (95% CI) = 0.90 (0.82, 0.95), SEM 2.9º, MDC_{90} 6.8º. Thoracic angle ICC (95% CI) = 0.91 (0.84, 0.95), SEM 3.5º, MDC_{90} 8.2º.
Topiladou et al. (2014)	Spinal mouse	50 adult participants with back pain or LBP. C7 and PSISs palpated and identified with markers. 2 measurements performed 30 minutes apart. New markers placed before 2nd measurement. Sagittal plane measurements: upright position, full flexion and full extension. Frontal plane measurements: upright position, left lateral bending and right lateral bending. Tx curvature (T1-T12) and Lx curvature (L1-L5) measured in all positions in both planes.	Intra-rater reliability in sagittal plane: Upright position: ICC (95%CI) = 0.96 (0.92-0.98), SEM 1.08º (Tx curvature) and 0.99 (0.97-0.99), SEM 0.39º (Lx curvature). Full flexion: ICC (95%CI) = 0.88 (0.75-0.94), SEM 2.74º (Tx curvature) and 0.99 (0.98-1.0), SEM 0.32º (Lx curvature). Full extension: ICC (95%CI) = 0.97 (0.94-0.99), SEM 0.98º (Tx curvature) and 0.96 (0.92-0.98), SEM 1.0º (Lx curvature). Intra-rater reliability in frontal plane: Upright position: ICC (95%CI) = 0.86 (0.70-0.93), 1.63º (Tx curvature) and 0.75 (0.48-0.88), SEM 1.71º (Lx curvature).

			Left lateral bending: ICC (95%CI) = 0.90 (0.78-0.95), 2.84º (Tx curvature) and 0.93 (0.85-0.96), SEM 1.0º (Lx curvature). Right lateral bending: ICC (95%CI) = 0.87 (0.73-0.94), SEM 2.59º (Tx curvature) and 0.90 (0.80-0.94), SEM 1.56º (Lx curvature).
Livanelioglu et al. (2016)	Spinal mouse	51 patients with AIS. Standing upright position assessed. Frontal plane thoracolumbar curvatures evaluated with spinal mouse by two therapists within the same day.	Inter-observer reliability: ICC (95% CI) = 0.93 (0.87-0.96) Intra-observer reliability: ICC (95% CI) = 0.96 (0.93-0.98) (Observer 1) and 0.94 (0.90-0.97) (Observer 2).
Sedrez et al. (2016)	Flexicurve	40 children for test-retest reliability and inter-rater reliability. 38 children for intra-rater reliability who had undergone x-ray examination but had no previous spinal surgery or congenital deformity. Two independent evaluators modelled each spine using flexicurve which provided angles for Tx and Lx spine in the sagittal plane. Sagittal standing position with normal posture measured.	Test-retest reliability: Tx kyphosis ICC (95% CI) = 0.93 (0.87-0.96), SEM 2.5º, MDC 4.9º. Lx lordosis ICC (95% CI) = 0.80 (0.61-0.89), SEM 4.3º, MDC 8.4º. Intra-rater reliability: Tx kyphosis ICC (95% CI) = 0.82 (0.65-0.91), SEM 4.1º, MDC 8.1º. Lx lordosis ICC (95% CI) = 0.67 (0.36-0.83), SEM 5.7º, MDC 11.2º. Inter-rater reliability: Tx kyphosis ICC (95% CI) = 0.83 (0.68-0.91), SEM 4.1º, MDC 8.0º. Lx lordosis ICC (95% CI) = 0.72 (0.47-0.85), SEM 5.7º, MDC 11.2º.
Was et al. (2016)	Inclinometer	Pilot study. 20 healthy participants.	Test-retest ICC reliability ranged between 0.70 and 0.90. Lx lordosis ICC = 0.90.

		Sagittal spine curvatures including Lx lordosis and Tx kyphosis. Measured in freestanding position. Examinations performed twice with one week interval by one examiner.	Tx kyphosis ICC = 0.80.
Quek et al. (2017)	Kinect (3-D depth camera)	33 healthy participants. 29 returned within 1-7 days to be re-examined. Standing and sitting positions measured. C7, T12 and S2 SPs manually palpated by examiner. Kinect measurement consisted of recording 5 consecutive frames of depth and image data from sensors.	<u>Intra-rater reliability:</u> Standing Tx kyphosis angle ICC (95%CI) = 0.96 (0.92-0.98), SEM 0.69º. Sitting Tx kyphosis angle ICC (95%CI) = 0.81 (0.60-0.91), SEM 1.07º.
Roghani et al. (2017)	Spinal mouse	19 women with hyperkyphosis and 14 without hyperkyphosis. Sagittal curvature assessed during neutral standing, full spinal flexion and full spinal extension. Same examiner on 2 different days (intra-rater reliability). Measured from C7 to approximately S3. Curvature calculated by spinal mouse software in each of the 3 positions (Tx – T1-T12 and Lx – L1-S1).	<u>Intra-rater reliability:</u> Hyperkyphosis group: Tx kyphosis ICC (95%CI) = 0.94 (0.86-0.98), SEM 1.56º, MDC 4.33º. Lx lordosis ICC (95% CI) = 0.97 (0.92-0.98), SEM 1.7º, MDC 4.71º. Without hyperkyphosis: Tx kyphosis ICC (95%CI) = 0.89 (0.69-0.96), SEM 1.75º, MDC 4.86º. Lx lordosis ICC (95% CI) = 0.97 (0.93-0.99), SEM 1.41º, MDC 3.91º.

(Lx, lumbar; L, lumbar vertebrae; S, sacral vertebrae; ICC, intra-class correlation coefficient; LBP, lower back pain; NSLBP, nonsignificant lower back pain group; SLBP, significant lower back pain group; E1, examiner one; E2, examiner two; C, cervical vertebrae; Tx, thoracic; 95%CI, ninety five percent confidence interval; SEM, standard error of measurement; º, degrees; D1, day one; D2, day two; cm, centimetres; MDC95, minimal detectable change at the ninety five percent confidence level; SPs, spinous processes; pps, participants; PSISs, posterior superior iliac spines; EM, electromagnetic; T, thoracic vertebrae; CMC, coefficient of multiple correlation; P, probability value; >, more than; MDC_{90}, minimal detectable change at the ninety percent confidence level; AIS, adolescent idiopathic scoliosis; MDC, minimal detectable change.)

1.2.3.1 Surface measurement methods and reliability

In this section, with the help of external literature; the surface measurement devices and methods for measuring thoracolumbar curvature are described and critiqued through evaluation of their advantages and limitations. Following this, the reliability results of each measurement method is presented and discussed.

All ICC values were interpreted using the classification described by Koo and Li (2016) where values <0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability and values >0.90 indicate excellent reliability. Since ICC values alone do not reveal absolute differences between measurements; where available, SEM and MDC values are also presented. This facilitates the application of these reliability estimates to clinical or research practice.

1.2.3.1.1 Flexicurve

The flexicurve method was first described by Burton (1986). This measurement device consists of a flexible ruler (Burton 1986; de Oliveira et al. 2012) marked in mm and centimetre (cm) increments (Burton 1986; Dunleavy et al. 2010). The technique requires a tester to manually mould the device to the midline contour of the spine (Burton 1986; Youdas et al. 1995) replicating an individual's spinal curvature (de Oliveira et al. 2012). This curve is then carefully traced onto paper where tangents at known anatomical landmarks are drawn; from which angles can be measured (Burton 1986; Hinman 2004b). The flexicurve has been used in 'normals' (Hart and Rose 1986; Youdas et al. 1995; Dunleavy et al. 2010) and to investigate LBP (Lovell et al. 1989; Mirbagheri et al. 2015) but not in conditions affecting spinal or trunk shape (e.g. scoliosis).

There are no safety concerns related to use of this instrument and it is easy to use; requiring only a very basic level of understanding of angles (Israel 1959; de Oliveira et al. 2012). The method involves a multistage process i.e. moulding the curve, tracing onto paper and then computing the pen and paper calculations. Nevertheless, none of this is particularly time consuming (Hinman 2004a). The flexicurve is also inexpensive (Israel 1959; de Oliveira et al. 2012) and allows for the whole spinal region

to be directly measured. It allows for measurement in the sagittal, frontal and transverse plane providing different moulds are taken. However, literature mostly focuses on sagittal measurements (Teixeira and Carvalho 2007) and it is limited to measuring static postures only.

1.2.3.1.1.1 Test-retest reliability

Excellent test-retest reliability for the flexicurve in standing and forward bending positions of healthy adult participants has been demonstrated (ICC 0.97) (Hart and Rose 1986). Good-to-excellent results amongst children for standing thoracic kyphosis (ICC (95%CI) 0.80 (0.61-0.89), SEM 4.3º) and lumbar lordosis (ICC (95%CI) 0.93 (0.87-0.96), SEM 2.5º) measurements have also been found (Sedrez et al. 2016).

1.2.3.1.1.2 Intra-rater reliability

Intra-rater reliability estimates for flexicurve measurements of lumbar spine curvature in the standing position are reported as good (ICC (95%CI) 0.78 (0.47-0.92) (de Oliveira et al. 2012), good-to-excellent (ICC 0.82-0.93) (Youdas et al. 1995), moderate-to-excellent (ICC 0.73-0.94) (Lovell et al. 1989) (ICC 0.61-0.97) (Dunleavy et al. 2010) and moderate (ICC (95%CI) 0.67 (0.36-0.83), SEM 5.7º) (Sedrez et al. 2016). In the same position, flexicurve thoracic kyphosis measurements have findings of good intra-rater reliability (ICC (95%CI) 0.83 (0.57-0.94)) (de Oliveira et al. 2012), (ICC (95%CI) 0.82 (0.65-0.91), SEM 4.1º) (Sedrez et al. 2016).

For maximum lumbar flexion in the sitting position, good-to-excellent intra-rater flexicurve measurements have been demonstrated (ICC 0.84-0.95) (Youdas et al. 1995), whilst excellent reliability has been reported for maximum lumbar extension in the prone position (ICC 0.96-0.98) (Youdas et al. 1995).

1.2.3.1.1.3 Inter-rater reliability

Inter-rater reliability for standing lumbar curvature measurements made with the flexicurve have been reported as good (ICC (95%CI) 0.83 (0.60-0.94)) by de Oliveira et al. (2012). Other studies have found poor-to-moderate lumbar curve inter-rater reliability estimates (ICC 0.41-0.54) (Lovell et al. 1989), (ICC 0.60-0.73) (Hinman 2004b), (ICC (95%CI) 0.72 (0.47-0.85), SEM 5.7º) (Sedrez et al. 2016) and (ICC 0.56-

0.64) for alternative methods of quantifying curvature including lumbar width, lumbar length and thoracic length measurements (Dunleavy et al. 2010).

Conversely, inter-rater repeatability for flexicurve thoracic kyphosis measurements were found to be good (ICC (95%CI) 0.83 (0.68-0.91, SEM 4.1º) (Sedrez et al. 2016) and excellent (ICC (95%CI) 0.94 (0.87-0.98) (de Oliveira et al. 2012) in normal standing positions; as well as excellent in relaxed (ICC 0.94) and erect (ICC 0.93) standing postures (Hinman 2004b). In addition, good reliability is demonstrated in thoracic width (ICC (95%CI) 0.72 (0.33-0.87)) and total thoracolumbar length (ICC (95%CI) 0.71 (0.13-0.89)) flexicurve measurements (Dunleavy et al. 2010).

1.2.3.1.1.4 Measurement errors

Both Sedrez et al. (2016) and Dunleavy et al. (2010) reported flexicurve SEM estimates. Depending on the region and analysis conducted, measurement errors varied from 2.5º to 5.7º for Sedrez et al. (2016).

Comparatively, Dunleavy et al. (2010) calculated SEM for the examiner with the highest mean square error and expressed SEM in absolute units as well as the percentage of the mean of the variables. Measurement error was moderate-to-high for thoracic width (SEM 0.48cm, %SEM 14.8%), lumbar length (SEM 1.93cm, %SEM 17.3) and lumbar width (SEM 0.35cm, %SEM 24.7%) but low for total length (SEM 0.58, %SEM 1.2%) and thoracic length (SEM 2.00cm, %SEM 2.4%) (Dunleavy et al. 2010). Additionally, Dunleavy et al. (2010) presented the MDC95. These values were high for width and length measurements of both thoracic and lumbar regions (MDC95 17-68%) (Dunleavy et al. 2010). As an example, the MDC95 value for thoracic length in standing was 5.55cm (Dunleavy et al. 2010). This means that with 95% confidence, a change greater than 5.55cm constitutes real change beyond natural variation expected with repeated flexicurve measurements.

In summary, it seems that the flexicurve is a reliable method of measuring spinal curvature in flexion, extension and standing of both the lumbar and thoracic spine. However, slightly lower reliability was determined between raters, especially for lordosis; with SEM values ranging from 1% to 25%. Therefore, the reliability seems to

be dependent on the operator and specific method employed to complete the measurement.

1.2.3.1.2 Spinal Mouse

The spinal mouse is a wheeled accelerometer device that works by taking the angle of the acceleration axis, relative to gravity, to provide the inclination or tilt angle and, uses double integration of acceleration to provide vertical distance travelled down the spine. Whilst the operator rolls the device over the spinous processes of the back, data is sampled every 1.3mm (Mannion et al. 2004) and transferred from the spinal mouse to a computer (Livanelioglu et al. 2016). Therefore, at each moment, the computer software receives the change in location of the device vertically and the tilt angle. This data is then used to calculate the relative location of each spinous process and the angle between relative spinous processes as well as, the total angle in the frontal and sagittal planes (Livanelioglu et al. 2016). Such a device has been used in 'normals' (Mannion et al. 2004), in those with LBP (Topiladou et al. 2014) and also, in individuals with scoliosis (Livanelioglu et al. 2016).

The spinal mouse is completely safe and is of a moderate cost (around £7000). It is very simple to use, requiring the operator to slowly roll the device down the back; providing immediate quantification of curvature (Post and Leferink 2004). Furthermore, it is able to measure the whole spine directly in the frontal and sagittal planes (Mannion et al. 2004; Post and Leferink 2004). Nonetheless, it is unable to measure dynamic motion. Additionally, there are some well understood issues with using accelerometers to measure distance due to mathematical integration succumbing to drift (Luinge and Veltink 2005; Luczak et al. 2017; Chandel and Ghose 2018) which may affect the devices accuracy.

1.2.3.1.2.1 Intra-rater reliability

In the sagittal plane, spinal mouse intra-rater reliability for thoracic curvature measurements were reported in the standing (Mannion et al. 2004) or upright position (Topiladou et al. 2014) as moderate for healthy participants between-day (ICC (95%CI) 0.73 (0.43-0.89), SEM 4.2º examiner 1; 0.88 (0.67-0.94), SEM 2.8º examiner 2) (Mannion et al. 2004) and excellent within-day for participants with back pain (ICC

(95%CI) 0.96 (0.92-0.98), SEM 1.08º) (Topiladou et al. 2014). For the same positions, lumbar curvature measurements also yielded excellent repeatability between-day (ICC (95%CI) 0.90 (0.75-0.96), SEM 2.5º examiner 1; 0.92 (0.80-0.97), SEM 2.4º examiner 2) (Mannion et al. 2004) and within-day (ICC (95%CI) 0.99 (0.97-0.99), SEM 0.39º) (Topiladou et al. 2014).

Three studies investigated the intra-rater between-day reliability of sagittal curvatures in flexion and extension positions (Mannion et al. 2004; Topiladou et al. 2014; Roghani et al. 2017). Roghani et al. (2017) reported combined ICC values for sagittal plane neutral standing, flexion and extension postures with the spinal mouse. They found excellent intra-rater repeatability in thoracic (ICC (95%CI) 0.94 (0.86-0.98), SEM 1.6º, MDC 4.3º) and lumbar (ICC (95%CI) 0.97 (0.92-0.98), SEM 1.7º, MDC, 4.7º) measurements of hyperkyphotic women and in lumbar lordosis measurements of healthy women (ICC (95%CI) 0.97 (0.93-0.99), SEM 1.41º, MDC 3.91º) (Roghani et al. 2017). However, thoracic kyphosis measurements amongst the latter group yielded only good reliability (ICC (95%CI) 0.89 (0.69-0.96), SEM 1.75º, MDC 4.86) (Roghani et al. 2017).

For flexion positions, moderate (ICC (95%CI) 0.67 (0.32-0.86), SEM 5.5º examiner 1) (Mannion et al. 2004) and good (ICC (95%CI) 0.86 (0.67-0.94), SEM 3.9º examiner 2) (Mannion et al. 2004), (ICC (95%CI) 0.88 (0.75-0.94), SEM 2.74º) (Topiladou et al. 2014) intra-rater reliability was found for sagittal thoracic curvature measurements between- and within-day; whilst measurements of lumbar curvature yielded good (ICC (95%CI) 0.85 (0.66-0.94), SEM 3.2º examiner 1) (Mannion et al. 2004) and excellent reliability (ICC (95%CI) 0.91 (0.79-0.97), SEM 3.0º examiner 2), (ICC (95%CI) 0.99 (0.98-1.0), SEM 0.32º) (Topiladou et al. 2014).

In extension, Mannion et al. (2004) reported good intra-rater reliability of spinal mouse measurements between-day for both thoracic (ICC (95%CI) 0.76 (0.47-0.90), SEM 6.2º; 0.78 (0.52-0.91), SEM 4.6º) and lumbar (ICC (95%CI) 0.80 (0.53-0.92), SEM 4.5º; 0.78 (0.52-0.91), SEM 4.6º) curvature in healthy participants, whereas excellent repeatability within-day (ICC (95%CI) 0.97 (0.94-0.99), SEM 0.98º; 0.96 (0.92-0.98), SEM 1.0º) was demonstrated by Topiladou et al. (2014) in participants with back pain.

Intra-rater reliability of frontal plane spinal mouse curvature measurements were reported in two studies. Livanelioglu et al. (2016) demonstrated excellent within-day intra-rater reliability of thoracolumbar measurements in the standing upright position for adolescent idiopathic scoliosis (AIS) subjects ((ICC (95%CI) 0.96 (0.93-0.98); 0.94 (0.90-0.97)). Conversely, in participants with back pain, good repeatability was shown for thoracic curvature measurements in an upright position (ICC (95%CI) 0.86 (0.70-0.93), SEM 1.63º) and in LLB and RLB (ICC (95%CI) 0.90 (0.78-0.95), SEM 2.84º; 0.87 (0.73-0.94), SEM 2.59º) (Topiladou et al. 2014). For lumbar curvature in upright, RLB and LLB positions, moderate (ICC (95%CI) 0.75 (0.48-0.88), SEM 1.71º), good (ICC (95%CI) 0.90 (0.80-0.94), SEM 1.56º) and excellent (ICC (95%CI) 0.93 (0.85-0.96), SEM 1.0º) reliability estimates were respectively demonstrated (Topiladou et al. 2014).

1.2.3.1.2.2 Inter-rater reliability

One study explored sagittal plane inter-rater reliability of the spinal mouse in healthy participants during standing, flexion and extension postures (Mannion et al. 2004). Two examiners measured each position three times for all participants on two separate testing days and reported the reliability between raters on both day one and day two of testing (Mannion et al. 2004). On day one, good inter-rater repeatability was seen for both thoracic and lumbar curvature measurements in standing (ICC (95%CI) 0.87 (0.70-0.95), SEM 2.7º; 0.87 (0.69-0.95), SEM 2.8º) and flexion (ICC (95%CI) 0.83 (0.61-0.93), SEM 4.3º; 0.85 (0.65-0.94), SEM 3.7º) (Mannion et al. 2004). However, in extension good reliability was only seen in lumbar measurements (ICC (95%CI) 0.88 (0.72-0.95), SEM 3.1º); with thoracic curvature yielding moderate results (ICC (95%CI) 0.64 (0.27-0.84), SEM 7.0º) (Mannion et al. 2004). In comparison, spinal mouse inter-rater reliability on day two was more varied. Excellent repeatability was demonstrated in standing lumbar curvature measurements (ICC (95%CI) 0.93 (0.83-0.97), SEM 2.2º) whereas, standing and extension thoracic and flexion and extension lumbar curvatures saw good results (ICC (95%CI) 0.83 (0.62-0.93), SEM 3.3º; 0.79 (0.54-0.92), SEM 5.0º; 0.87 (0.70-0.95), SEM 3.3º; 0.90 (0.75-0.96), SEM 3.5º) and moderate reliability was reported for thoracic curvature in flexion (ICC (95%CI) 0.70 (0.38-0.87), SEM 5.2º) (Mannion et al. 2004).

Livanelioglu et al. (2016) was the only study to investigate inter-rater reliability of spinal mouse thoracolumbar measurements in the frontal plane. In the standing upright position of participants with AIS they found excellent reliability between measurements taken on the same day by two therapists (ICC (95%CI) 0.93 (0.87-0.96)) (Livanelioglu et al. 2016).

1.2.3.1.2.3 Measurement errors

According to the SEM values in Roghani et al. (2017), for both hyperkyphotic and non-hyperkyphotic women, a change in kyphosis angle of >2º measured with the spinal mouse would be beyond the SEM. However, their MDC values illustrate a change >5º to 6º would represent true change beyond natural variability (Roghani et al. 2017).

For Mannion et al. (2004) the intra-rater between-day SEM ranged between 2.4º to 6.2º whereas, the inter-rater SEM ranged from 2.2º to 7.0º. The highest SEM was reported for thoracic curvature measurements in extension (6.2º intra-rater SEM examiner 1; 7.0º inter-rater SEM day 1) (Mannion et al. 2004). Notably, intra- and inter-rater curvature SEMs made in standing were particularly low (2.2º to 4.2º) (Mannion et al. 2004), indicating that the spinal mouse represents a sensitive instrument for investigating changes in standing sagittal spinal profile (Mannion et al. 2004). For example, the MDC for standing lumbar curvature is 6.9º (examiner 1) and 6.7º (examiner 2). This means that if a change of >6.7º occurred after a given intervention, this would represent true change at the 95% confidence level.

Overall, it appears the spinal mouse demonstrates reliability in standing and upright postures as well as flexion and extension positions in both the sagittal and frontal plane. However, within-day measurements have slightly higher reliability than between-day. Additionally, between-raters, it seems sagittal thoracic curvature measurements in flexion and extension positions have the lowest reliability. Sagittal thoracic curvature in extension also yielded the highest SEM, whilst standing sagittal SEMs remained low between 2.2º and 4.2º. As a result, it suggests reliability of the spinal mouse is affected by the time interval between measurements as well as the position and region of interest measured. Therefore, individuals planning to utilise this method should seek out specific reliability estimates for the region and movement planned.

1.2.3.1.3 Inclinometer

Inclinometers have a flat base that measure inclination proportional to the angle between the flat base and the vertical (Adams et al. 1986). For manual devices, the operator places the feet of the inclinometer over spinous processes and records directly the tilt angle relative to the vertical (Lewis and Valentine 2010). The operator has the freedom to acquire the absolute angle (relative to vertical) or take two (or more) measures and compute the relative angle (resultant angle between two tilt angles). Such a device has been used in 'normals' (Czaprowski et al. 2012; Was et al. 2016) and on those with LBP (Kluszczynski et al. 2017) but not in individuals with alterations to spinal/trunk shape (e.g. scoliosis).

These devices afford cheap, quick and simple posture analysis (Barrett et al. 2018) and both manual and digital inclinometers are available. Nevertheless, dynamic measurements are not possible with this device and multiple readings are required to provide measures of curvature along the spine.

1.2.3.1.3.1 Test-retest reliability

Was et al. (2016) and MacIntyre et al. (2014) researched the between-day test-retest reliability of the inclinometer in standing. In 20 healthy participants reliability estimates were moderate-good (ICC 0.70-0.90) with lumbar lordosis measurements yielding better reliability (ICC 0.90) than thoracic kyphosis measurements (ICC 0.80) (Was et al. 2016).

Conversely, amongst a sample of individuals at risk of osteoporotic fractures, MacIntyre et al. (2014) found almost identical reliability for thoracic angle measurements (ICC (95%CI) 0.91 (0.84, 0.95)) and lumbar angle measurements (ICC (95%CI) 0.90 (0.82, 0.95)). Small SEM (2.9º lumbar, 3.5º thoracic) was also reported for the inclinometer when measures of standing posture were performed by the same rater (MacIntyre et al. 2014).

1.2.3.1.3.2 Intra-rater reliability

In Czaprowski et al. (2012) study, one investigator performed a spinal curvature measurement of 30 healthy participants in standing one week apart to assess

between-day intra-rater reliability. They found no statistically significant difference ($p>0.05$) between measurements made on the two separate occasions. Moreover, measurement errors performed by the same examiner were low (2.8º-3.8º) indicating good sensitivity of the inclinometer for spinal curvature measurements in the standing position.

Lewis and Valentine (2010) investigated within-day intra-rater reliability of thoracic kyphosis inclinometer measurements in participants with and without shoulder pain. One examiner took a set of three measurements in a relaxed standing position on two occasions on the same day. Reliability estimates produced were excellent (ICC (95%CI) 0.97 (0.95-0.99); 0.97 (0.94-0.98)) for both groups. What is more, the SEM results based on the ICC average data for the combined T1/2 and T12/L1 (kyphosis) measurement was 1.0º for participants without upper body symptoms and 1.7º for those with symptoms (Lewis and Valentine 2010). These SEM findings suggest there may be less error associated with the inclinometer method when the mean of three measurements is employed (Lewis and Valentine 2010).

1.2.3.1.3.3 Inter-rater reliability

Inter-rater reliability for the inclinometer method was investigated by Czaprowski et al. (2012). Three investigators performed curvature measurements three times within-day for healthy participants in standing. No significant difference was found between raters for thoracic kyphosis measurements ($p=0.37$) but a significant difference was revealed for lumbar lordosis ($p=0.02$) (Czaprowski et al. 2012). Reliability estimates were reported using cronbach's alpha where a value closer to 1 demonstrates greater similarity between raters. For angle measurements, Cronbach's alpha was 0.87 for lumbar lordosis and 0.83 for thoracic kyphosis (Czaprowski et al. 2012). This indicates that inclinometer measurements taken by more than one investigator have good reliability (Czaprowski et al. 2012).

Altogether, it suggests the inclinometer is a reliable method for measuring spinal curvature in standing when performed by the same rater. Moreover, it seems that higher intra-rater reliability and lower SEM is seen when the mean of three measurements is employed. Nevertheless, measurements may be less reliable between raters particularly for lumbar lordosis. Furthermore, information concerning

the reliability of the inclinometer in positions other than standing is not included; thus highlighting an important limitation to this review.

1.2.3.1.4 Electromagnetic tracking

EM tracking systems involve a source that emits electromagnetic energy, and a sensor that records position and orientation information (Fenster et al. 2001) relative to the source (Parent 2012; Preim and Botha 2014). It has the capacity to provide both 3-D coordinates (for location) and 3-D angles (orientation); making it a true six degrees of freedom (6DOF) system. There are two distinct modes of use in the literature. The first involves attaching multiple sensors to the skin from which orientation (or tilt) information is provided and used to calculate curvature (Parkinson et al. 2013). This method is commonly researched in 'normals' (Parkinson et al. 2013) and in individuals with LBP (Pourahmadi et al. 2018), but not people with alterations in spinal/trunk shape (e.g. scoliosis). The second method involves using the sensor as a 3-D pen from the coordinates of location relative to the source (spinal sweep method). This involves the operator slowly tracing the sensor along the spinous processes; gathering location data in 3-D (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). This approach has been used in 'normals' and in individuals classified as obese or elderly (Singh et al. 2010; Gonzalez-Sanchez et al. 2014).

An EM tracking system is non-invasive, safe and costs between £2000 and £6000 depending on the number of sensors required. Attachment of the sensors and computing of angles is relatively simple with single plane measurement, involving just simple angle subtraction. However, this is not done within the EM software, but it is easy to complete in other applications including microsoft excel for example. Furthermore, although tracing of spinal shape is quick to complete and very simple, the computations are more complex since no software is available with this 'built in'. Even so, a routine in MATLAB (Mathworks 2019), or similar, will produce a result almost immediately.

The EM spinal tracing method collects data for the whole spine directly but it does not allow for dynamic motion capture. Comparatively, the orientation method, with multiple sensors, provides information for the regions covered by the sensors. Therefore, the more sensors used, the better the coverage of direct measurement. With this set up,

truly dynamic curvature measurement is possible but a large number of sensors would be required to cover the whole thoracolumbar spine which adds expense.

In addition, several important limitations are documented in the literature and apply to both methods. Firstly, although the device is portable, it has an optimum operational zone (Franz et al. 2014) which is around 220mm^3 to 720mm^3 from the source (Milne et al. 1996; Bull and Amis 1997). Outside of this region, consequences in positional and orientation accuracy are noted (Schuler et al. 2005) unless increases in power output occur (Bull and Amis 1997). As a result, data collection is constrained and the systems use in more dynamic settings is limited. Moreover, the systems operating accuracy can be significantly affected by metallic interference (Milne et al. 1996; Burnett et al. 2004; Ng et al. 2009; Franz et al. 2014) highlighting another source of error.

1.2.3.1.4.1 Electromagnetic tracking (spinal sweep method)

1.2.3.1.4.1.1 Intra-rater reliability

Singh et al. (2010) investigated spinal shape using the sweep method in 52 healthy participants. They found intra-tester within-day repeatability of the EM sweep method was excellent for thoracic kyphosis and lumbar lordosis measurements in standing (ICC (95%CI) 0.93 (0.88-0.95) and 0.98 (0.96-0.98) respectively) (Singh et al. 2010). These measurements also had small SEMs (1.6º thoracic kyphosis and 1.5º lumbar lordosis) indicating there is a high precision of sagittal curvature measurements with EM tracking (Singh et al. 2010). In addition, reliability estimates for lateral curvatures at thoracic and lumbar regions were reported as good (ICC (95%CI) 0.75 (0.61-0.85), SEM 1.04º; 0.84 (0.85-0.91), SEM 0.56º) suggesting the instrument is still highly reliable for extremely small curvatures (Singh et al. 2010).

In a follow up study, Gonzalez-Sanchez et al. (2014) analysed thoracolumbar curvatures in normal weight and obese participants during standing. They chose to present reliability using cronbach's alpha. Findings showed spinal segment angles had cronbach's alpha coefficients ranging from 0.90-0.92 in the sagittal plane and from 0.82-0.86 in the coronal plane indicating high intra-rater reliability within-day (Gonzalez-Sanchez et al. 2014).

To summarise, the EM tracking system using the spinal sweep method has demonstrated intra-rater reliability in the sagittal and frontal plane during standing with very low SEMs indicating high absolute reliability. Nonetheless, this method has only been explored in standing. Consequently, the review is limited to this literature and it is not known as to the repeatability of this method between-day, between-raters or in different positions.

1.2.3.1.5 Fibre-optic

Fibre-optic devices utilise light as their method of measuring curvature. A string of sensors is placed along a ribbon core and the intensity of light received by a sensor is modulated by the degree of bend thus, producing a curve (Williams et al. 2010). The fibre-optic ribbon is attached to the skin over S2 and the ribbon lies along the spine up to and past T1. It has a number of sensors along its length, thereby enabling the participant to move and for the shape of the spine to be determined.

The fibre-optic ribbon is safe to use, non-invasive and radiation free. It is also relatively cheap (£5000) but its availability is currently highly limited (Williams et al. 2010). The challenges to its use include fixing the reference sensor and providing adequate attachment, to enable movement, whilst also keeping the ribbon flush to the spine. Once these are solved, it is simple to use (Cloud et al. 2014) and curvature measures are available directly from the software. As the ribbon spans the length of the spine, it allows capture of whole spine curvature and can provide dynamic measurements (Williams et al. 2010; Williams et al. 2012). Nonetheless, this system is limited to the sagittal plane (Cloud et al. 2014). Such a system has been used in 'normals' (Williams et al. 2010) and in individuals with LBP (Williams et al. 2012; Williams et al. 2013) but not people with altered spinal/trunk shape (e.g. scoliosis).

1.2.3.1.5.1 Repeated measures reliability

Williams et al. (2010) used a fibre-optic based system to measure sagittal curvature of the whole and lower lumbar spine in 13 healthy participants. They calculated the coefficient of multiple correlation (CMC) to determine the similarity between three repeated movement trials of flexion and lifting. When curves are similar, CMC values are closer to 1. In this study, the fibre-optic system showed excellent similarity for

repeated measures of whole lumbar spine and lower lumbar spine curves in flexion and lifting movements (mean CMC = 0.97-0.98) (Williams et al. 2010). Accuracy of the fibre-optic system, as measured by the absolute mean difference for the peak curvature, for the whole lumbar spine was 2.3º ± 2.3º and 2.5º ± 2.7º and for the lower lumbar spine 2.7º ± 2.2º and 2.3º ± 2.0º during flexion and lifting movements respectively.

In their follow up study, Williams et al. (2012) used a fibre-optic device to measure curvature through time during flexion, lifting and extension movements in acute and chronic LBP sufferers. The mean CMC values were found to be excellent (0.81-0.97) across all movements and regions of the lumbar spine with small variability (Williams et al. 2012). Additionally, ICC values were excellent for all repeated movement trials (0.99) and absolute mean differences demonstrated small differences between repeated measurements (<2º for flexion, lifting and extension) (Williams et al. 2012).

In summary, Williams et al. (2010) and Williams et al. (2012) indicate the high reliability of repeated measures of fibre-optic methods in measurements of sagittal lumbar curvature in both healthy and back pain populations during dynamic movement. However, this method has received little exploration in the literature for measurements of spinal curvature. Therefore, limited research exists resulting in eligible studies investigating only the repeated movements of lumbar curvature measurements.

1.2.3.1.6 3-D depth camera (Kinect)

This device is a portable video game accessory that combines a video, infrared-sensing camera and infrared-transmitter (Castro et al. 2017; Quek et al. 2017). It acquires a two-dimensional representation of shape with the addition of depth, to yield 3-D coordinates (Castro et al. 2017; Mentiplay et al. 2013). From these coordinates, a virtual spinal curve can be created, and tangents can be used to determine curvature at specific points along the curve.

This technology is safe, remains affordable (£100) and is readily accessible (Mentiplay et al. 2013; Quek et al. 2017). In addition, images obtained can be processed within seconds, with almost instant results (Quek et al. 2017). However, identifying landmarks and obtaining curvature measurements takes longer and requires a specific

programme to be developed (Castro et al. 2017). The whole spine is able to be captured with this technology, enabling direct measurement of any thoracolumbar spinal region. Even so, the system requires calibration before use or a designated assessment area; highlighting further limitations (Quek et al. 2017).

1.2.3.1.6.1 Intra-rater reliability

One study eligible for this review investigated the intra-rater reliability of this 3-D depth and image camera (Quek et al. 2017). Three measurements of 33 healthy participants were taken in the standing and sitting positions with 29 participants returned within one to seven days to be re-examined. Thoracic kyphosis angle measured in each position demonstrated good-to-excellent reliability (ICC (95%CI) 0.96 (0.92-0.98); 0.81 (0.60-0.91)) (Quek et al. 2017). SEMs close to 1º (0.69º and 1.07º) (Quek et al. 2017) were also reported indicating high absolute reliability of the Microsoft Kinect for thoracic curvature measurements in standing and sitting postures.

Altogether, this 3-D depth and image camera method has demonstrated intra-rater reliability for thoracic kyphosis measurements between-day. Nevertheless, this is a fairly novel method in relation to spinal curvature measurements thus, little research exists investigating its use in other spinal regions, populations and postures.

1.2.4 Conclusion

This scope of the literature reviewed the existing methods of spinal surface curvature measurements. The advantages and limitations of these various techniques are outlined, and their reliability discussed. It is evident that there are multiple reliable methods to measure thoracolumbar surface curvature available. However, it is also clear that each method has inherent limitations and not one method will suit all applications. Consequently, caution is advised when choosing a method to consider its limitations, its relative reliability and its error estimates in order to determine if the desired method is sensitive enough to meet the needs of the specific application.

Furthermore, from this scoping review it also apparent that whilst many devices and options are available for detecting spinal shape or curvature, few have extended these methods to include measurement of trunk shape or shape assessment in individuals with altered trunk shape, for instance those with scoliosis. Current practice for measuring trunk shape in individuals with such conditions is often the domain of radiography, where repeated radiation exposure is a significant limitation. Therefore, future work should look to extend the knowledge of spinal shape measurement, beyond that of just the spine, to include the shape of the trunk. Resultantly, offering a non-radiation exposure method of assessing trunk shape and importantly, changes in trunk shape.

1.2.5 Summary

What is known?

Various reliable surface measurement methods exist for assessment of spinal shape in thoracic and lumbar regions.

Each method has inherent limitations so not one method suits all applications.

What is not known?

The reliability of these measurement methods in populations with altered spinal shape.

Whether these methods can be used to measure trunk shape.

Future work:

Further work is required to extend these measurement methods beyond that of spinal shape to include trunk shape.

Investigation into the consistency of these methods for measuring trunk and spinal shape should be investigated.

Studies should also aim to apply these methods in populations with altered spinal/trunk shapes.

1.3 Statement of the problem

Spinal disorders are a leading source of disability worldwide and one of the most common and costly complaints in clinical practice (Martin et al. 2008; Raciborski et al. 2016). This class of diseases has a broad scope and many of these pathologies are reported as LBP irrespective of actual diagnosis or cause (Raciborski et al. 2016). The prevalence of LBP is widely documented in the literature with lifetime rates of 50-85% (Hoy et al. 2010). Whilst the epidemiology of specific spinal-related disorders are less frequently studied, prevalence estimates have been reported as 6% for spondylolysis (Kalichman et al. 2009), 10.3% for osteoporosis (Wright et al. 2014) and 11-38% for lumbar spine stenosis (Jensen et al. 2020). Furthermore, research has identified a prevalence of >8% for scoliosis related deformities in adults and this rises considerably up to 68% in populations over the age of 60 (Schwab et al. 2005; Konieczny et al. 2013). To this end, spinal-related pathologies are one of the most common reasons for an individual to seek medical attention, impacting the consumption of resources more than any other health problem (Haldeman et al. 2012; Ramdas and Jella 2018). In the United Kingdom alone, the direct cost of back related pain exceeds £1 billion per year whilst the indirect cost is estimated to be significantly higher (Maniadakis and Gray 2000; Maetzel and Li 2002).

Previous research has linked spinal-related disorders to impairments of posture, function and kinematics (Schmid et al. 2015; Kuwahara et al. 2016; Christe et al. 2017; Chun et al. 2017). As a result, clinical examinations of such pathologies routinely involve the measurement of spinal curvature and trunk posture (Vrtovec et al. 2009; Pazos et al. 2007; Fortin et al. 2011). This helps identify the disorder, facilitates the monitoring of disease evolution and provides the basis for evaluating interventions (de Oliveira et al. 2012; Plaszewski et al. 2012). Consequently, measurement techniques for spinal and trunk posture are important tools required for clinicians.

Nevertheless, current practice is mostly based on subjective impressions including simple visual analysis which is difficult to compare amongst clinicians (Pazos et al. 2007; Fortin et al. 2011 de Oliveira et al. 2012). Alternatively, for quantifiable measures, reliance remains on radiographic evaluation which has been the gold-standard since the 1930s (Fortin et al. 2011; de Oliveira et al. 2012). Despite its confirmed validity, radiographs are expensive, invasive, not easily accessible and

involve exposure to radiation (Fortin et al. 2011; de Oliveira et a. 2012). They are also mostly used to verify bony structures or spinal alignment (Fortin et al. 2011). Consequently, radiographic evaluation is not suitable for repeated measurements over-time and provides limited analysis. As highlighted in the scoping review of literature; low-cost, reliable instruments have also been developed but these surface methods are significantly limited by their inability to detect posture beyond spinal shape.

To assess, treat and truly understand disorders that effect both the spine and trunk, analysis of spinal shape on its own is insufficient. For this reason, it is well recognised that there is a clinical need for a measurement method that can be used to assess, and monitor changes to, an individual's spinal and trunk shape without involving exposure to radiation. This measurement method must be reliable, of an acceptable cost and relatively simple to use if it is to be appropriate for a clinical setting. Eventually, this could offer a reliable, non-ionizing method for quantifying spinal and trunk shape in-vivo that could be appropriate for routine clinical use.

1.4 Need for the study

In order for this development to occur, future research should work towards the application of surface curvature measurement methods in assessment of the whole trunk and determine the consistency of these measurements before exploring this in 'normals' and those with altered trunk shape. However, the limitations, relative reliability and error estimates documented in the literature for surface curvature measurement methods of thoracolumbar shape need to be considered for these future applications.

This project will attempt to extend the current capacity of curvature measurement to resolve the insufficiencies identified in the review of literature whilst also, meeting the clinical needs outlined above. The working title of this project is: 'Measuring spinal and trunk shape using an electromagnetic sensor.'

This project has the following objectives:

1. Develop a method for measuring spinal and trunk shape utilising an EM tracking device, building on previous work of Singh et al. (2010) and Gonzalez-Sanchez et al. (2014).
2. Provide proof of concept of this method in determining spinal and trunk shape.
3. Determine the validity and reliability of this method.
4. Explore the optimal data processing for this method and make recommendations based on these findings.
5. Determine the capacity of this method to detect alteration in spinal and trunk shape.

Chapter 2. Development of Methods

The following section outlines how some of the additional methods specific to this thesis were developed. The aims were as follows:

- Determine whether the testing environment was suitable for data collection.
- Describe the development of the probe and the results pertaining to correction of orientation data for the probe.
- Demonstrate the functionality of the equipment and method to trace shapes different to those documented in the literature, including the tracing of whole trunk shape.

2.1 Instrumentation

2.1.1 Electromagnetic tracking system

An EM tracking system (Patriot, Polhemus, Vermont, US) was selected to map the 3-D shape of the spine and trunk. This system was chosen based on two specific aspects. Firstly, the review above demonstrated that many surface measurement methods have critical limitations. The flexicurve method is limited to just the points measured by the instrument. Therefore, to 'cover' the trunk would require many measurements and be extremely time consuming. This is also true of the inclinometer method, ruling this out. Equally, the fibre-optic method and Kinect method remain in their infancy and without significant computing input; computation remains non-automated and experimental. Furthermore, although the spinal mouse is very promising, it is constrained by the operating software to work only as specifically designed. This means extending measurements to other regions of the trunk and across the trunk are currently beyond the measurement capacity of the device. In view of these limitations, the spinal sweep method utilising an EM sensor operating as a 3-D pen or probe, which has been developed in earlier work, provides the greatest opportunity to extend spinal shape measurements to include trunk shape.

The second reason an EM system was chosen is because it delivers 6DOF tracking thus, providing both positional and orientation information needed to complete the tracing of spinal and trunk shape. The Patriot sensor system (Polhemus, Vermont, US)

is one example of an EM system consisting of a systems electronics unit, one standard sensor and one source (Polhemus 2020a). There is also the option of adding one additional sensor to expand the systems tracking capabilities (Polhemus 2020a). The source is placed at a fixed location where it emits an EM field and tracks the sensors within this field; recording the position and orientation data of the sensor in real time (Parent 2012; Preim and Botha 2014; Polhemus 2020a).

The static accuracy of the system provided by the manufacturer is 1.5mm RMS for the x,y,z position and 0.40º RMS for sensor orientation (Polhemus 2020a). However, some drawbacks relating to the range and accuracy of the magnetic field are documented in the literature. As a result, a study was conducted to ensure the proposed testing environment was appropriate for data collection.

2.1.2 Electromagnetic interference study

2.1.2.1 Introduction

The use of EM sensors is associated with several limitations including metallic interference (Milne et al. 1996; Burnett et al. 2004; Ng et al. 2009; Franz et al. 2014) and optimum operational zones (Milne et al. 1996; Bull and Amis 1997; Schuler et al. 2005; Franz et al. 2014). These limitations are important considerations for data collection as they also represent key sources of error.

Research has shown that an EM systems operating accuracy is related to the presence of metal and metallic interference (Milne et al. 1996; Franz et al. 2014). This effect appears not to be universal across all metals, explaining the conflict of effect in the literature. Yet when evident, metallic interference can significantly affect measurements taken by the EM device (Milne et al. 1996; Burnett et al. 2004; Ng et al. 2009; Franz et al. 2014).

In addition, as an EM device's source emits EM energy, it creates an operational zone (Franz et al. 2014). This zone is a cubic area around $220mm^3$ to $720mm^3$ from the source that the EM tracking system is designed to operate within (Milne et al. 1996; Bull and Amis 1997). If data is measured outside of this region then increases in power output are necessary (Bull and Amis 1997) otherwise, consequences are noted in

positional and orientation accuracy (Schuler et al. 2005). This indicates that the usefulness of the EM system for more dynamic applications is limited, constraining data collection to within this optimum operational range.

In light of these limitations, a study investigating the metallic interference and optimum operational zone for the EM tracking sensors in the proposed testing environment was conducted.

2.1.2.2 Methods

Aim: ensure testing environment is suitable for EM tracking system data capture and identify key sources of error in EM testing equipment.

Equipment:

- Laptop with EM Tracker Config software downloaded.
- Electromagnetic tracking system: systems electronics unit, source and sensor.
- Double sided sticky tape.
- Polystyrene surface piece.
- Wooden stool.

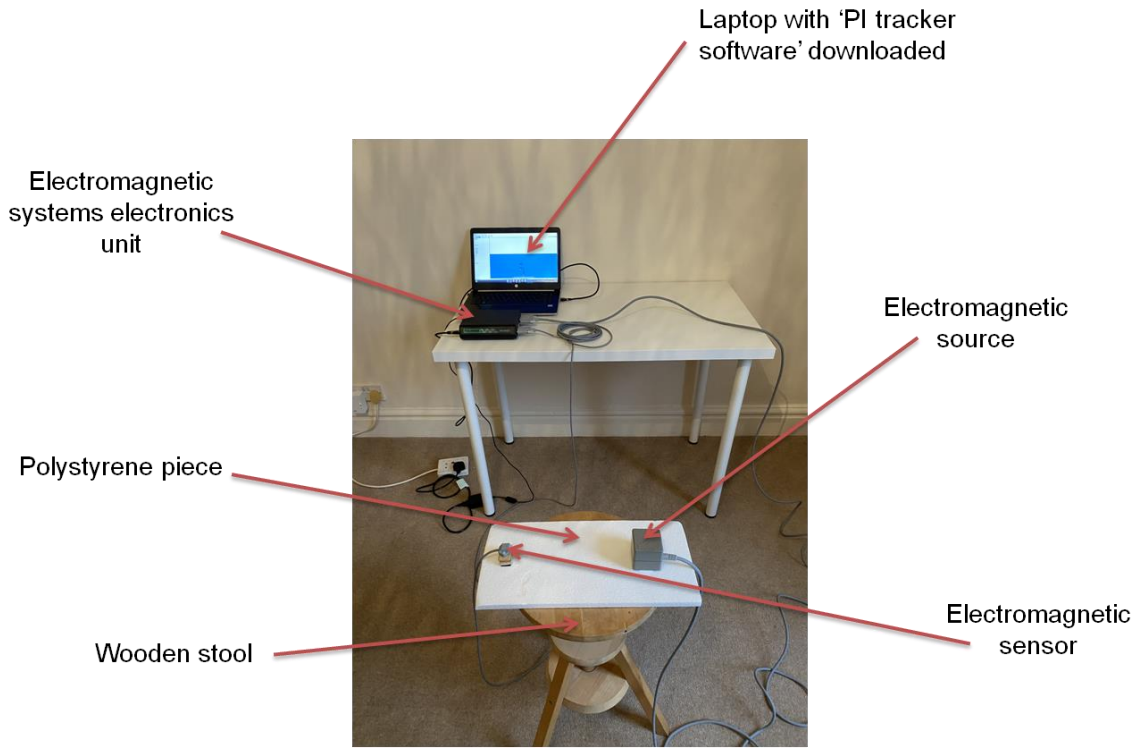


Figure 3 – Equipment set up and testing environment

The sensor and source were set up as shown in Figure 3. Data were captured pertaining to the location and orientation of the sensor relative to the source for 20-30 seconds and stored for analysis. The testing 'environment' was then 'manipulated' to explore the effect of various configuration on data stability. Manipulation can be seen in table 11 where condition 1 and 2 served as the control.

Table 11. Description of each test condition

Condition	Description
1	Source and sensor static on polystyrene surface on wooden stool.
2	Source and sensor static on polystyrene surface on wooden stool.
3	Source and sensor static on polystyrene surface on wooden stool. Wooden stool moved around room.
4	Source and sensor static on polystyrene surface. Polystyrene surface moved around room.
5	Source and sensor static on polystyrene surface. Polystyrene surface moving around wooden stool.

2.1.2.3 Results

Table 12 shows the average and mean (standard deviation (sd)) of data for each DOF measurement during conditions 1 to 5.

In condition 1, the sd was higher than expected for yaw therefore, this exact condition was repeated (condition 2). This repeat occurred 10 minutes later. The sd value for yaw in condition 1 was 173.0º whereas, in condition 2 it was 0.003º. The variation in yaw during condition 1 is depicted in Figure 4.

For conditions 2 to 5, the sd of all location measurements were <0.06inch (<1.5mm) and for orientation were <0.54º.

2.1.2.4 Discussion

Even though the source and sensor were both static in condition 1, there was high variation in yaw taken by the EM device as illustrated in figure 4. Evidence has suggested that ambient temperature can cause error in sensor and measurement

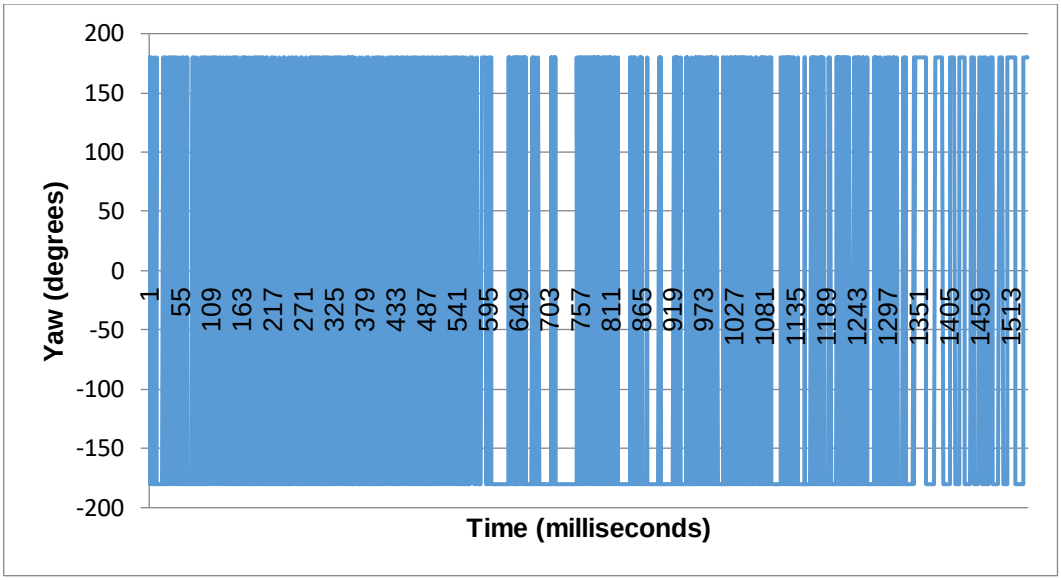


Figure 4. Variation in yaw during condition 1

Table 12. The mean and standard deviation measurements for each test condition

		Degree of freedom measurement					
		Distance (inches)			Orientation (degrees)		
Condition	**Statistic**	**X**	**Y**	**Z**	**Yaw**	**Pitch**	**Roll**
1	**Mean**	10.597	-0.314	0.242	-49.804	2.835	-1.430
	sd	0.001	0.002	0.002	173.009	0.029	0.002
2	**Mean**	10.598	-0.312	0.245	-179.981	2.862	-1.435
	sd	0.000	0.000	0.000	0.003	0.005	0.002
3	**Mean**	10.733	-0.160	0.109	178.657	4.425	-1.256
	sd	0.007	0.003	0.026	0.026	0.295	0.024
4	**Mean**	10.713	-0.202	0.028	178.635	3.617	-0.116
	sd	0.013	0.016	0.052	0.047	0.544	0.554
5	**Mean**	10.528	0.860	0.278	-177.676	2.076	-0.764
	sd	0.010	0.008	0.058	0.057	0.500	0.100

(sd, standard deviation)

system results (Parent 2012; Zhang et al. 2018). Since condition 1 was conducted almost immediately after the EM device had been switched on, it is possible that temperature changes explain these findings. Furthermore, the low sd seen in condition 2 for the same yaw measurement could be attributed to the fact that the EM device had been switched on for at least 10 minutes. Therefore, there was less error in

orientation because the device had already adjusted to the environmental temperature.

In conditions 2 to 5, the sd for all 6DOF measurements remained small. This indicates that the positional and orientation accuracy of the EM device is high in different positions of the proposed testing environment. Moreover, conditions 2 to 5 indicate that neither the stool nor testing environment demonstrates metallic interference.

Altogether, this study suggests that our testing occurs within the optimum operational zone of our EM device and thus, our proposed testing environment is appropriate for data collection.

2.1.2.5 Implications

From condition 1 we have inferred that error in results may occur within the first 10 minutes of the EM device being switched on, possibly as a result of temperature stabilisation of the device. This means that before any data collection occurs the EM device will be switched on at least 10 minutes before collecting any data.

Furthermore, the study also demonstrated that our proposed testing environment presents no metallic interference and is within the optimum operational zone of the EM device thus, no changes in power output are necessary.

2.2 Development of probe

In order to be able to trace the shape of both the spine and the trunk, a probe or stylus (with EM sensor attached) is required. There are two reasons for this. Firstly, the sensor's location for sensing occurs in the centre of the sensor, not at the edge. Therefore, any coordinates will relate to the sensor sensing centre not the casing edge. Secondly, a single centre point is needed in order to correct the data for orientation. For instance, if the sensor is rotated by 90 degrees, the centre of the sensor has now moved thus, the coordinates no longer accurately reflect the location of the sensor. So that the coordinates can be corrected for sensor location, a probe is critical if spinal and trunk shape is to be traced. Further details of this are described in section 2.2.2 below (correction of tip location).

2.2.1 Probe design

The probe design is shown in Figure 5. Due to the known limitation of metallic interference on EM sensors documented in the literature (Milne et al. 1996; Burnett et al. 2004; Ng et al. 2009; Franz et al. 2014), there was careful consideration as to the materials used. This resulted in the probe being made completely out of wood. In addition, plastic screws were used to fasten the EM sensor to this wooden probe to keep it stable.

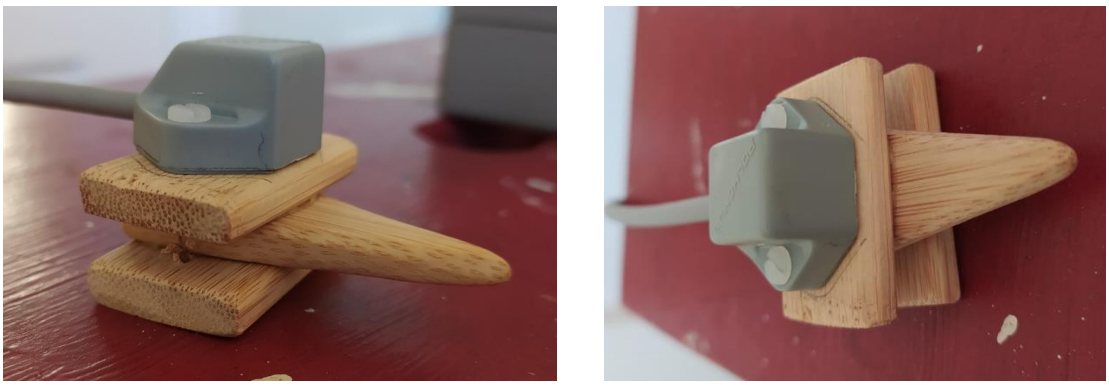


Figure 5. Wooden probe design

The distance of the probe tip to the sensor's measurement centre was measured using digital callipers (accuracy ± 0.2mm). This can be seen in Figure 6.

Distance from the sensor to the tip in X direction = 36.8mm (blue arrow)

Distance from the sensor to the tip in Z direction = 19.5mm (orange arrow)

Distance from the sensor to the tip in Y direction = 0mm

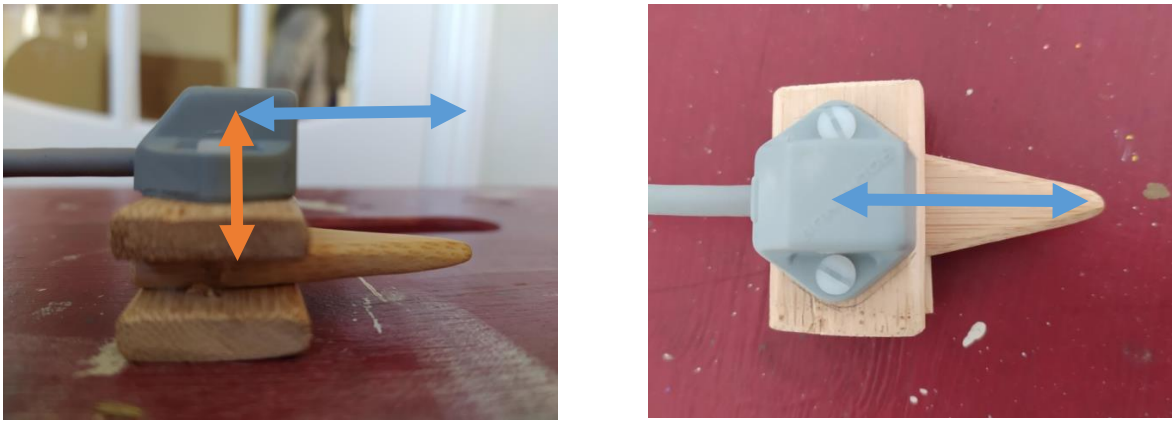


Figure 6. Distance of probe tip to sensors measurement centre in X,Y,Z direction

2.2.2 Correction of tip location

To demonstrate the effect of the coordinate transform (moving coordinates from sensor sensing centre to probe tip), the probe tip was held in place whilst the sensor was rotated. From this, 2 graphs were created: one prior to, and one following the coordinate transform. The transform requires the location of the tip relative to the measurement centre of the sensor, and the orientation of the sensor for every time point (i.e. 100Hz). The location of the probe is fixed therefore this is subtracted from the x,y,z coordinates to relocate them. Then a rotation matrix is created from the direction cosine obtained from the orientation of the sensor in pitch, roll and yaw. Following this, the coordinates are then reoriented about the axes of x, y and z to orient them correctly. This is completed for each sample (i.e. every time point).

The first graph plot of pitch, roll and yaw (Figure 7) demonstrates change in orientation of the sensor during the test of the coordinate transform. The tip is fixed, therefore it is physically being rotated about the probe tip.

Figure 8 shows the data comparison between coordinates with and without orientation correction. Figure 9 shows a close up of the Y coordinates.

The sd of the location reduced 58%, 66% and 53% following the correction algorithm. Therefore, with the sensor attached to the probe the x,y,z coordinates are now minimally

affected by orientation of the sensor. This means the probe with sensor attached now functions similarly to a 3-D pen.

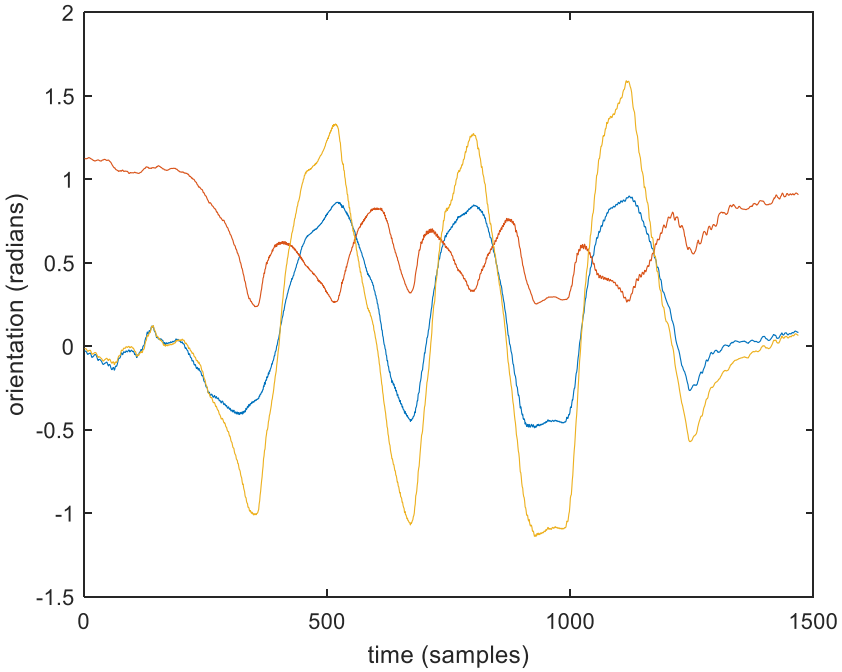


Figure 7. Pitch, roll and yaw for orientation test

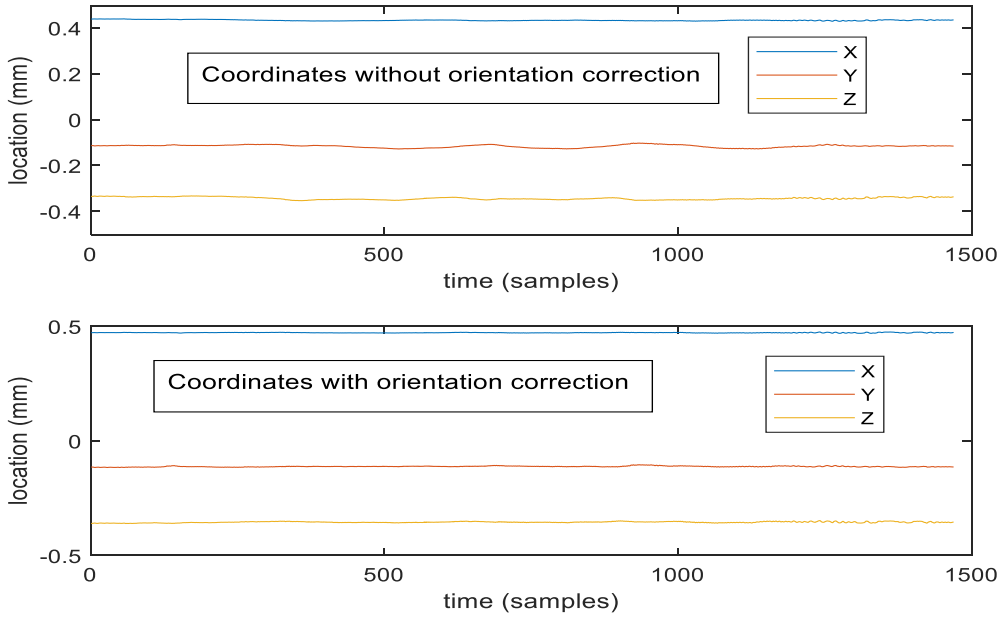


Figure 8. Comparison of X,Y,Z coordinates with and without orientation correction

Close up of Y coordinates:

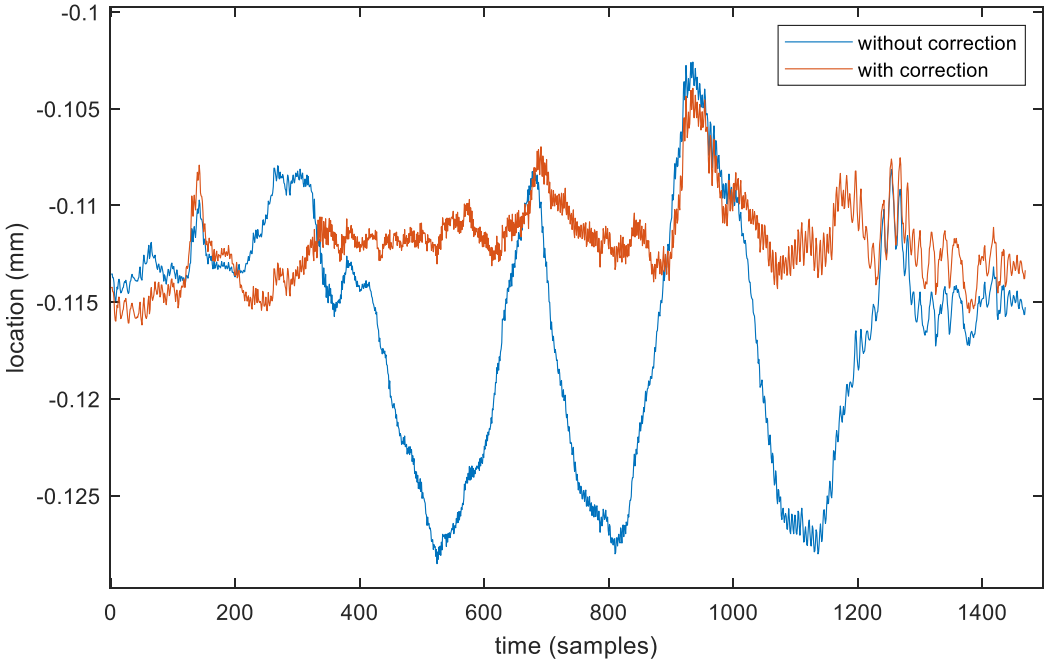


Figure 9. close up of Y coordinates with and without orientation correction

2.3 Computational methods

2.3.1 Spinal shape

Previous studies have demonstrated the capacity of the EM system to operate as a 3-D pen allowing the shape of the spine to be traced (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). From this, tangents can be drawn for curvature measures to be taken. However, to date, this has only been applied in standing and only for the shape of the spine itself (measurement of spinous processes profile). To achieve spinal shape measurement the following steps are required:

A. Digitize the bony landmarks of the spinous processes of T1, T8, L1, L5 and left and right posterior superior iliac spine (PSIS).
B. Sweep the probe along the spine.
C. Complete a coordinate transform from the EM source (global coordinate system) to the pelvis (local coordinate system).
D. Reorient the coordinates to map onto the local coordinate system.
E. Fit a polynomial smoothing function to the spine trace.
F. Create tangents from which relative angles (between two tangents) can quantify the curvature.

These steps were taken directly from the literature (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). To complete these steps, a specific MATLAB (Mathworks 2019) algorithm was created.

To determine if the curve fitting computation method would work for spinal shapes different from the standing human, a series of shapes were drawn with the probe and steps A-E were completed and plotted.

Resultant graphs using the 5^{th} order polynomial curve fitting (taken from the literature) can be seen in Figures 10 to 13 for each condition.

On visual inspection of these figures, they demonstrate the capacity of the EM system, and the MATLAB (Mathworks 2019) algorithm, to trace the shape of a simulated spine beyond that of the method previously published in the literature (Singh et al. 2010; Gonzalez-Sanchez et al. 2014).

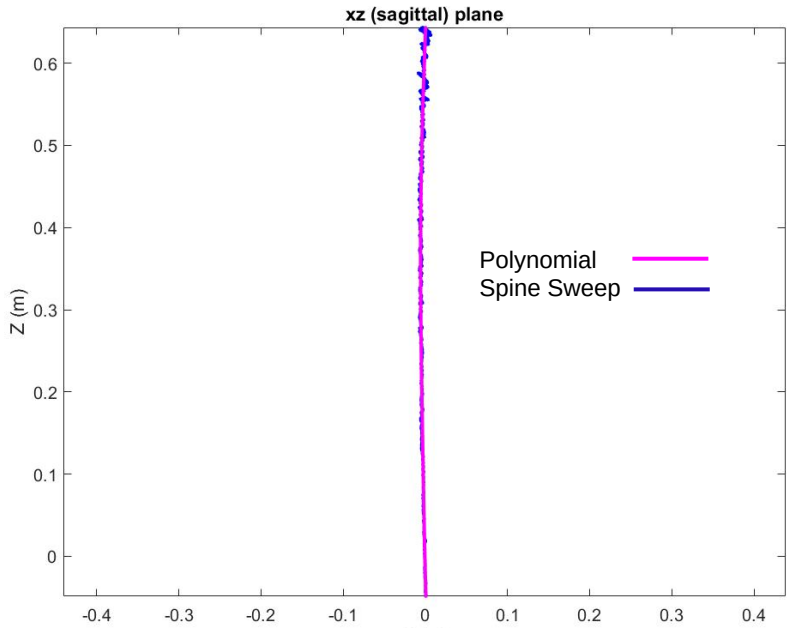


Figure 10. Condition 1 spinal sweep and 5th order polynomial smoothing

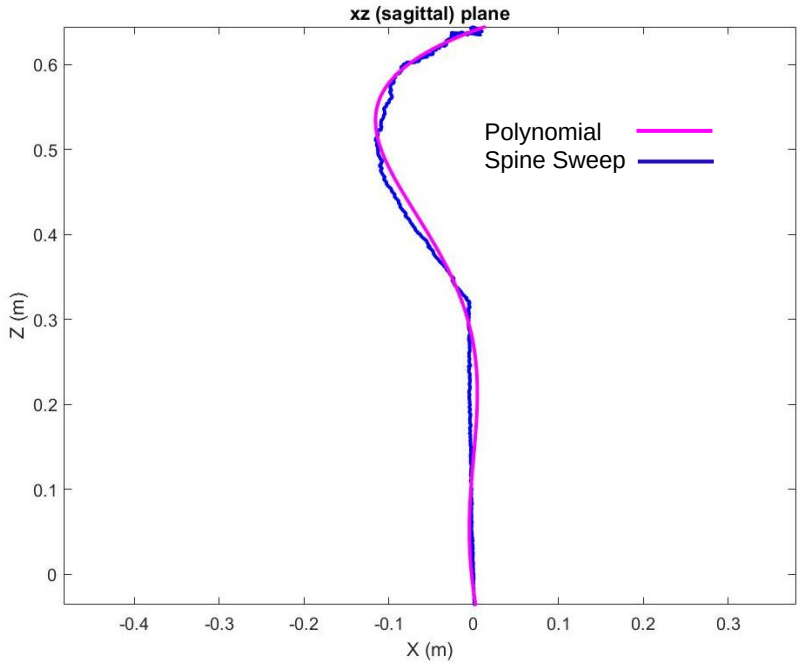


Figure 11. Condition 2 spinal sweep and 5th order polynomial smoothing

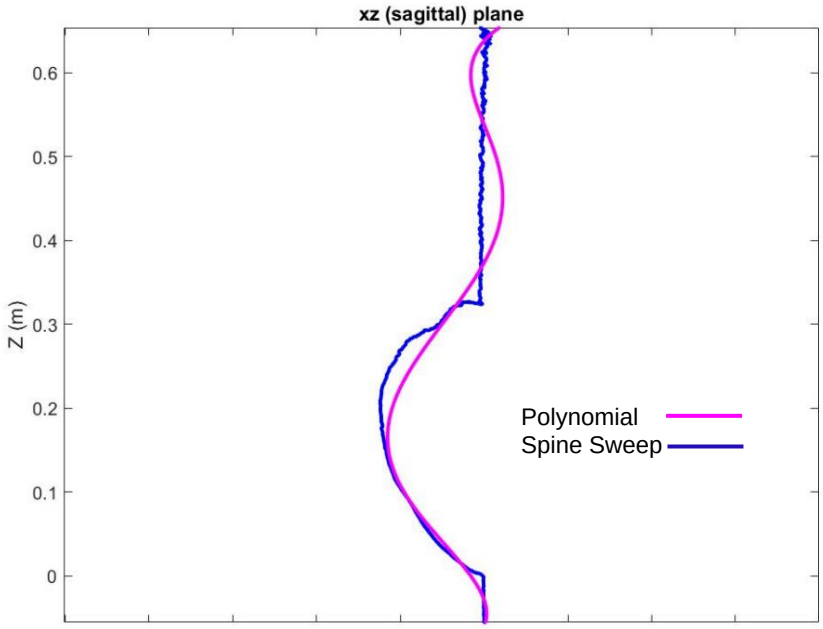


Figure 12. Condition 3 spinal sweep and 5th order polynomial smoothing

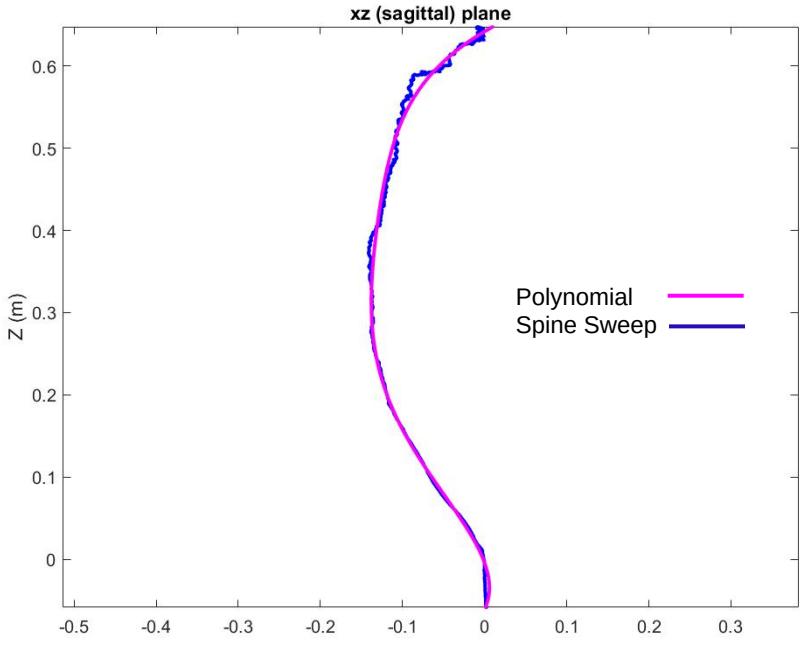


Figure 13. Condition 4 spinal sweep and 5th order polynomial smoothing

However, it was evident that a number of assumptions have been made in this processing method:

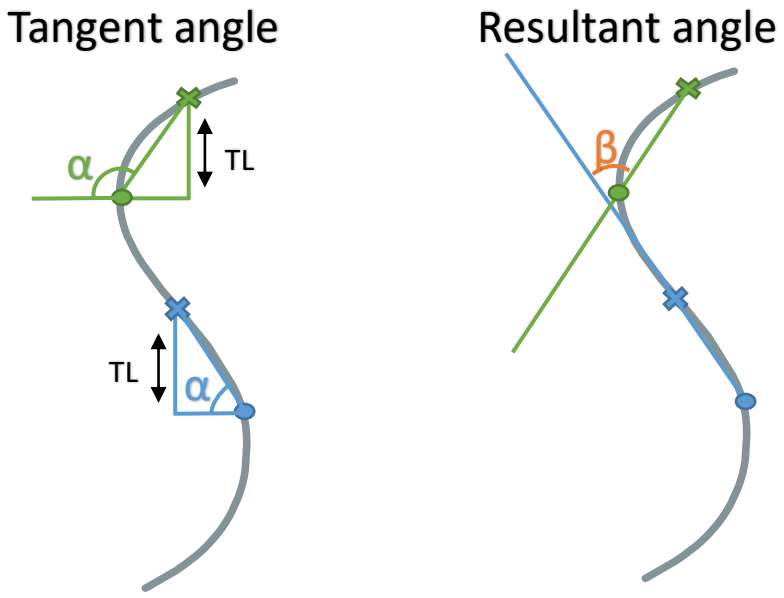
1. The ‘optimal’ mathematical fitting function is the 5th order polynomial.

This may be the case for the previously published work. However, looking at condition 3 (Figure 12), the polynomial fit (pink line) is a poor representation of the actual data.

2. The ‘optimal’ tangent length is 0.05m.

This is questionable depending on the area of interest and where the tangent spans. For example, in condition 3 (Figure 12) at the coordinate (0.0, 0.3) it is clear that if a tangent were to ‘cross’ this region, then a loss of the specific shape could result.

The drawing of tangents requires two points along the curve to be identified between which a straight line or tangent is drawn. The angle of this tangent is then calculated relative to the horizontal (Figure 14). The intersection of two tangents yields a resultant angle for each specific anatomical region i.e. L5 tangent and L1 tangent intersection results in the lumbar (Lx) angle. Tangent and resultant angle methods are shown in Figure 14.



(**×**, anatomical landmark; **o**, third point of triangle generated by mathematical function; **TL**, desired tangent length; **α**, tangent angle; **β**, resultant angle)

Figure 14. Spinal shape tangent and resultant angle method

Therefore, it is clear that a more systematic approach to justification of these aspects of the method is required and thus, will form part of this thesis.

2.3.2 Trunk shape

In addition, to further extend the novelty of this thesis, analysis was extended to include quantification of trunk shape. To achieve this, some additional steps are required in the capture and processing phase:

A. Digitize points as above.
B. Sweep the probe across the whole trunk surface.
C. Complete a coordinate transform from the EM source (global coordinate system) to the pelvis (local coordinate system).
D. Reorient the coordinates to map onto the local coordinate system.
E. Fit a surface smoothing function to the coordinates to fit a surface to the trunk shape trace.
F. Create tangents from which relative angles (between two tangents) can quantify the curvature in the transverse plane.

A specific MATLAB (Mathworks 2019) algorithm was also created to complete these additional steps. Then a flat surface was used as a proof of concept segregate. The landmarks of T1, T8, L1, L5, LPSIS and RPSIS were digitized on the flat surface prior to sweeping the probe across the surface. This was then plotted (Figure 15). The bony landmarks are represented as red circles and the sweeps of the probe as blue lines. Following this, the coordinates were plotted as a point cloud and a surface fitting function was applied in MATLAB (Mathworks 2019). Examples are seen in Figures 16 and 17.

At this stage, the trunk surface data is represented by a mathematical function enabling the retrieval of any coordinates, even if they were not traced by the probe. From here, the next step is to calculate the tangents for the angle in the transverse plane. This will tell us the protrusion of the trunk at various anatomical locations. This is calculated relative to the flat horizontal plane (Figure 18).

As mentioned for the spinal shape fitting model, a number of ‘decisions’ are required to determine some of the methods. Specifically, this relates to the optimal surface fit function to best map the trunk shape. Therefore, part of this thesis will seek to determine the optimal trunk shape fitting model as well as, explore the effect of tangent length.

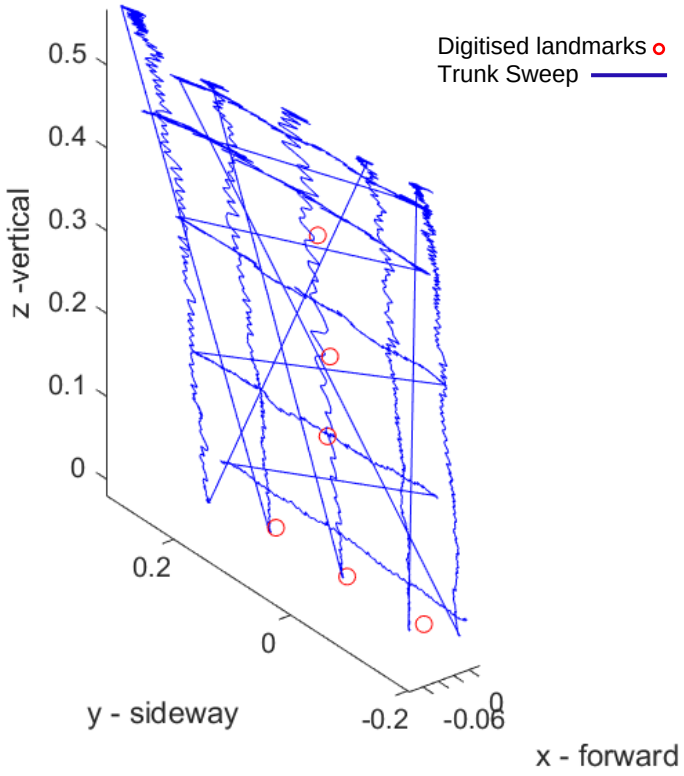


Figure 15. Plotted landmark digitisation on the flat surface

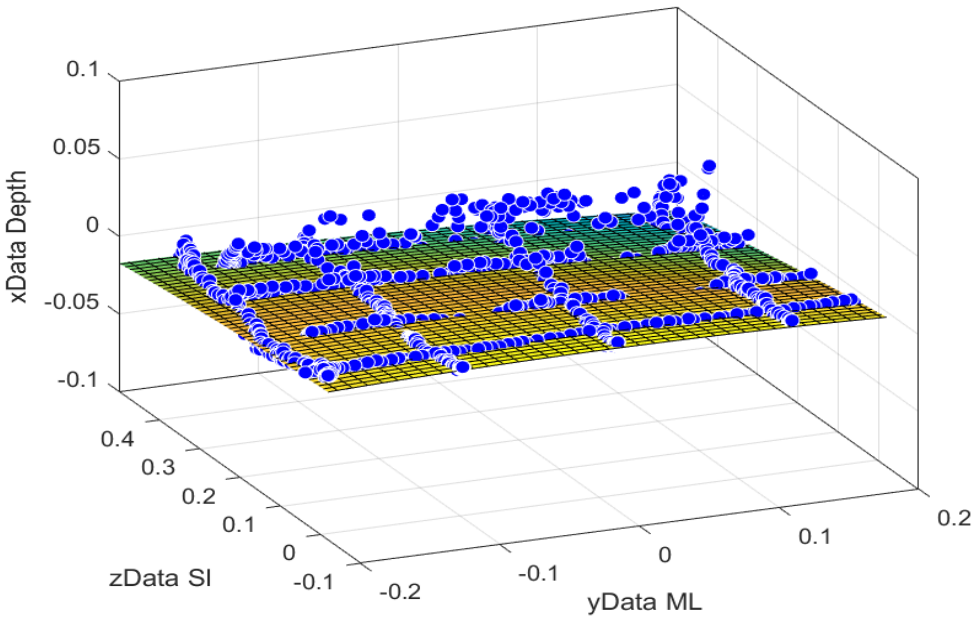


Figure 16. Surface fit example 1

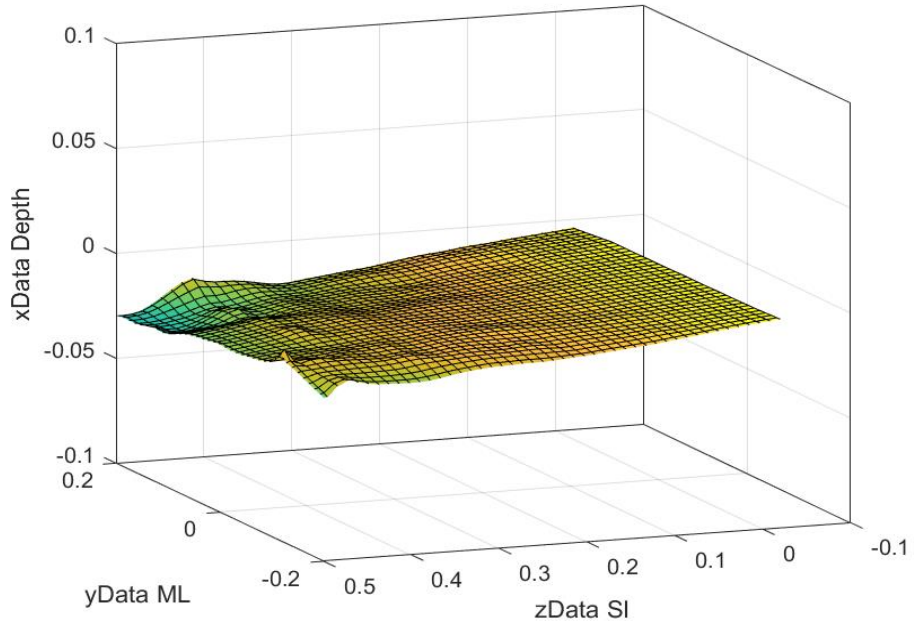
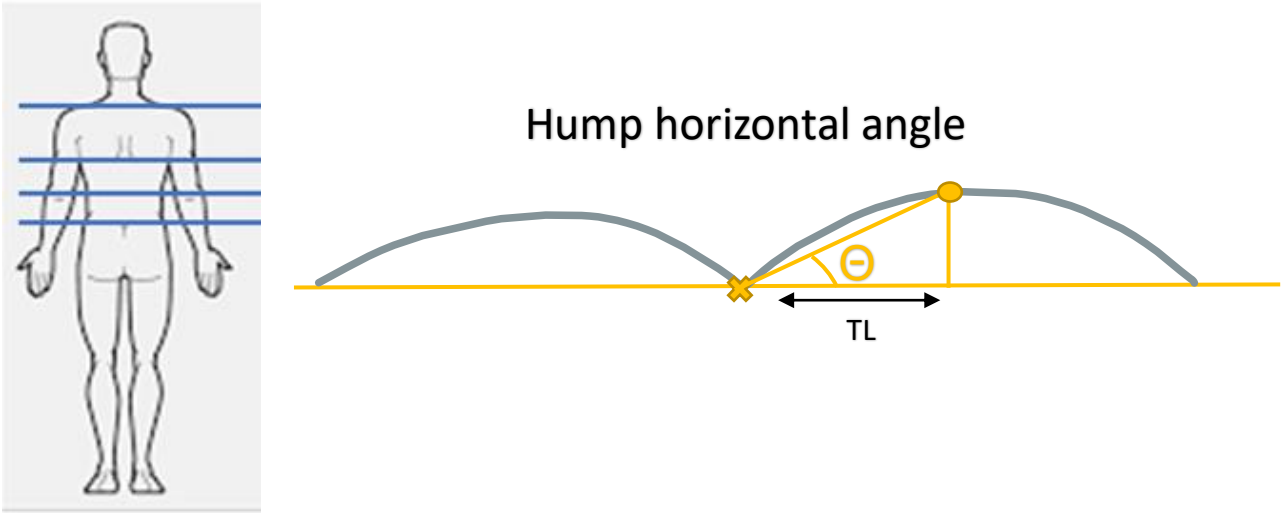


Figure 17. Surface fit example 2



(**×**, anatomical landmark; **o**, third point of triangle generated by mathematical function; **TL**, desired tangent length; **Θ**, hump horizontal angle)

Figure 18. Trunk shape hump horizontal angle method

2.4 Summary

In summary, we have attached the EM sensor to a wooden probe and corrected the tip location of the sensor; enabling the probe to now function like a 3-D pen. Furthermore, we have developed a data processing technique to measure both spinal and trunk shape and have established proof of concept of this method through tests on simulated spines/ shapes. A preliminary study has demonstrated that our proposed testing environment does not present metallic interference and is within the optimum operational zone of the EM system. Lastly, important implications have arisen whilst developing these methods which has helped inform the project's main methods. These include questions concerning what is the optimal mathematical function to fit the data for both spinal shape and trunk shape? and what is the influence of varying the tangent length? These questions will be carried forward into the main study.

Chapter 3. Main Methods

3.1 Study aims and experimental design

The purpose of this part of the study was to:

1. Determine the validity and reliability of this method.
2. Explore the optimal data processing for this method and make recommendations based on these findings.
3. Determine the capacity of this method to detect alteration in spinal and trunk shape and determine the validity and reliability of these measurements.

A repeated measures design was employed to achieve these aims.

3.2 Instrumentation

Spinal and trunk shape were measured using the EM device (Patriot, Polhemus) as described previously (section 2.1.1).

The EM sensor was attached to a wooden probe in order to trace spine and trunk shape. During the preliminary work above, a MATLAB (Mathworks 2019) programme was written to process the data.

3.3 Phantom models

In order to remove the natural human variability when exploring this novel method, phantom models were used. To create the phantom models, measurements between spine and trunk anatomical landmarks were firstly taken on a human model in three different positions: neutral standing, forward flexion and in extension. A qualified physiotherapist, using the techniques described in Field's anatomy, palpation & surface markings text (Field and Hutchinson 2013), identified bony landmarks through palpation and marked them with an eyeliner pencil. These included the spinous processes of C7, T1, T8, L1, L5, S2, the LPSIS and RPSIS, the left and right posterior iliac crest and the posterior-lateral corner of the right and left acromion. The distance (in cm) between these anatomical landmarks were then recorded using the flexicurve

instrument that could be moulded to the curvature of their back. The full list of measurements taken are listed in Table 13.

Table 13. Measurements taken between anatomical landmarks of the human model

Measurements
• Distance between C7 and T1. • Distance between T1 and T8. • Distance between T8 and L1. • Distance between L1 and L5. • Distance between L5 and S2. • Distance between L5 and left posterior-lateral iliac crest. • Distance between L5 and right posterior-lateral iliac crest. • Distance between S2 and LPSIS. • Distance between S2 and RPSIS. • Distance from T1 to posterior-lateral corner of left acromion. • Distance from T1 to posterior-lateral corner of right acromion. • Length of spine from C7 to S2.

(LPSIS, left posterior superior iliac spine; RPSIS, right posterior superior iliac spine.)

Later, the flexicurve was moulded by the physiotherapist into three different shapes to imitate extension, flexion and neutral spinal curvatures. These curves were traced onto paper and used to construct the shape and surface of each phantom model. This meant that resultant phantoms matched a real person in both size and shape.

Figures 19 to 24 shows the construction process of the flexion phantom model as an example. The phantoms were made entirely out of cardboard and pieced together using glue. The measurements of spine and trunk anatomical landmarks, taken from the human model in each position (Table 14, 15 and 16) were then marked onto the corresponding shaped phantom. Figures 25 to 30 show the finished phantom models.

3.3.1 Pseudo-scoliosis phantom

In addition to the above phantoms, a pseudo scoliotic phantom was also constructed from the existing phantoms (Figure 31). This was done by rotating the flexion and extension phantoms 90 degrees and attaching them together to create a smooth surface that was convex (kyphotic) to the left and concave (lordotic) to the right. In keeping with the clinical presentation of scoliosis, the pelvis was kept flat by attaching a piece of polystyrene below the model which enabled digitisation of these bony landmarks.

Table 14. Anthropometric measurements of human model in extension

Measurement	Centimetres
C7 to T1	2.3
T1 to left acromion process	Superior 1.0, Lateral 20.2
T1 to right acromion process	Superior 1.0, Lateral 20.0
T1 to T8	19.7
T8 to L1	14.4
L1 to L5	7.5
L5 to left iliac crest	Superior 1.7, lateral 6.1
L5 to right iliac crest	Superior 1.8, lateral 7.3
L5 to S2	2.6
S2 to left posterior superior iliac spine	Superior 0.6, lateral 4.1
S2 to right posterior superior iliac spine	Superior 0.6, lateral 4.3
C7 to S2	45.5

Table 15. Anthropometric measurements of human model in flexion

Measurement	Centimetres
C7 to T1	3.1
T1 to left acromion process	Superior 1.1, lateral 19.6
T1 to right acromion process	Superior 0.9, lateral 20.3
T1 to T8	19.9
T8 to L1	17.3
L1 to L5	12.6
L5 to left iliac crest	Superior 2.1, lateral 6.1
L5 to right iliac crest	Superior 2.1, lateral 7.0
L5 to S2	5.2
S2 to left posterior superior iliac spine	Superior 0.4, lateral 3.8
S2 to right posterior superior iliac spine	Superior 0.4, lateral 4.3
C7 to S2	58.1

Table 16. Anthropometric measurements of human model in neutral standing

Measurement	Centimetres
C7 to T1	2.4
T1 to left acromion process	Superior 0.9, lateral 19.4
T1 to right acromion process	Superior 0.6, lateral 19.8
T1 to T8	19.7
T8 to L1	16.6
L1 to L5	8.5
L5 to left iliac crest	Superior 1.9, lateral 6.2
L5 to right iliac crest	Superior 2.0, lateral 7.0
L5 to S2	4.3
S2 to left posterior superior iliac spine	Superior 0.6, lateral 4.0
S2 to right posterior superior iliac spine	Superior 0.6, lateral 4.3
C7 to S2	51.5

Figure 19. Construction process

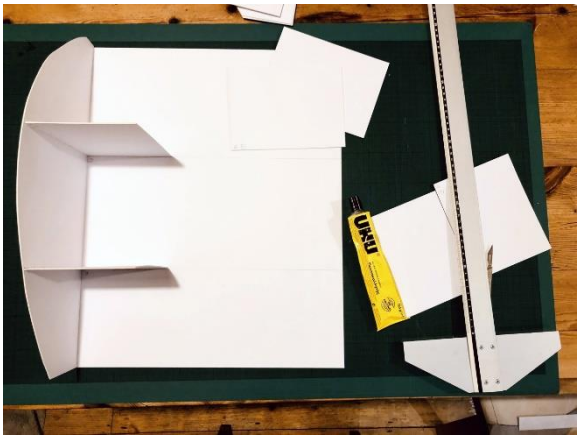


Figure 20. Construction process

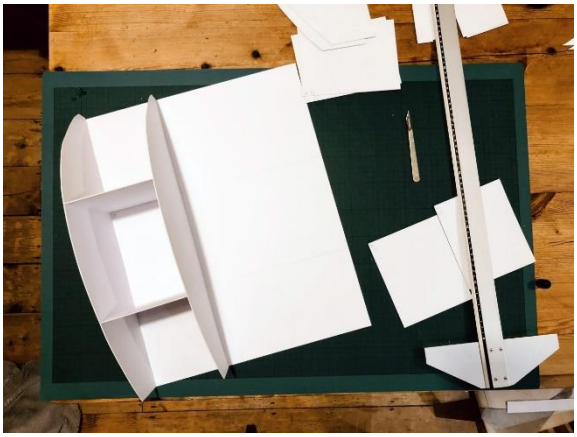


Figure 21. Construction process

Figure 22. Construction process

Figure 23. Construction process

Figure 24. Construction process

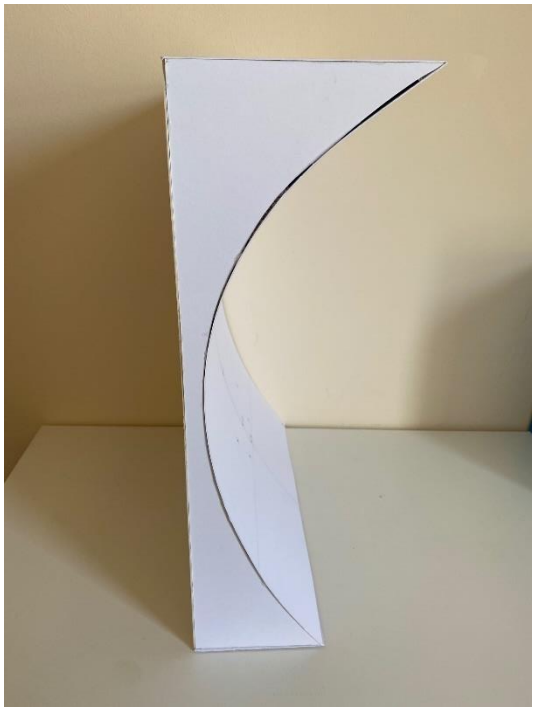


Figure 25. Final extension phantom

Figure 26. Final extension phantom

Figure 27. Final flexion phantom

Figure 28. Final flexion phantom

Figure 29. Final neutral phantom

Figure 30. Final neutral phantom

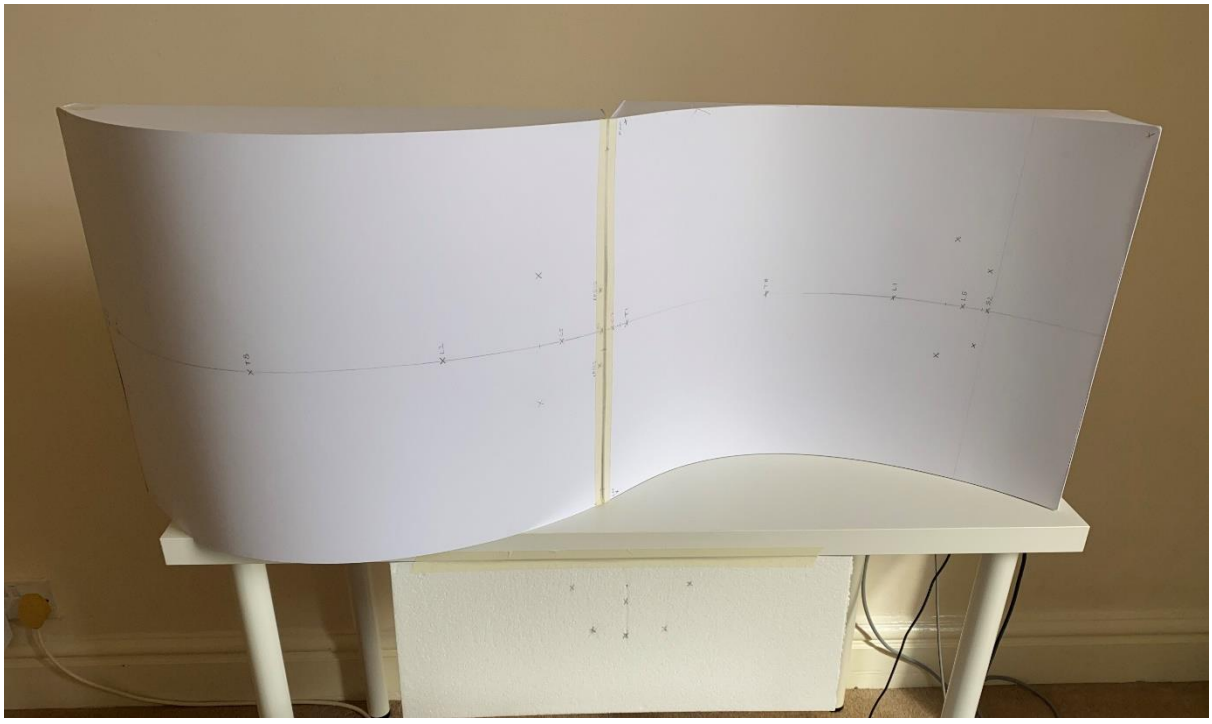


Figure 31. Final pseudo-scoliosis phantom

3.4 Method

Spinal and trunk shape data were collected on the neutral, flexion and extension phantoms in a vertical position. Firstly, the bony landmarks were digitised by collecting data pertaining to their location in 3-D space relative to the EM source. The tip of the wooden probe was then moved inferiorly down the phantom overlying the spine between T1 and S2 (spinal shape measurement) or from T1 to S2 with inferior and superior vertical and horizontal sweeps across the phantom surface (trunk shape measurement). The probe stayed in contact with the phantom at all times.

3.4.1 Reliability

Two sets of three repeats were conducted for both spinal and trunk shape. One data set was done consecutively for each model without moving or changing the position of the phantom during the 3 repeats (i.e. neutral x3, flexion x3, extension x3). The other set of trials matched a more clinical scenario whereby, one data trial was completed for each shaped model (i.e. extension x1, flexion x1, neutral x1) and then this was repeated 3 times.

Six data trials were also completed consecutively on the constructed ‘scoliosis’ model. This also occurred with the model in a vertical position. However, only measurements of trunk shape were taken and, due to the surface boundaries of the model, sweeps did not extend down to S2 or the pelvis. Instead, trunk sweeps were limited to just below L1, so angle variables analysed did not include L5. Even so, digitisation of bony landmarks did include the pelvis and L5. This meant coordinates could still be transformed from the global coordinate system to the local coordinate system.

3.4.2 Validity

Angles were calculated using the flexicurve method (Burton 1986; Hinman 2004b; de Oliveira et al. 2012) to establish validity of the computed angles with the EM system. In order to complete these pen and paper calculations, it was first necessary to understand the computed angle calculation method.

3.4.2.1 Computed tangent angle method for spinal shape

An example of this method for spinal shape is depicted in Figure 32. Firstly, the EM system provides the location of the anatomical landmark (e.g. L1) in 3-D space from which the z,x co-ordinates are extracted. This location represents one point of the triangle. The specified tangent length (e.g. 0.05 m) is then subtracted from this value to form the second location of the triangle. From this second point, the x coordinate is provided from the spinal shape polynomial fit data. This completes the triangle. The value calculated is then the angle relative to the left hand horizontal (represented by 'α' in Figure 32). This is the computed method for tangent angles of T1, T8 and L1.

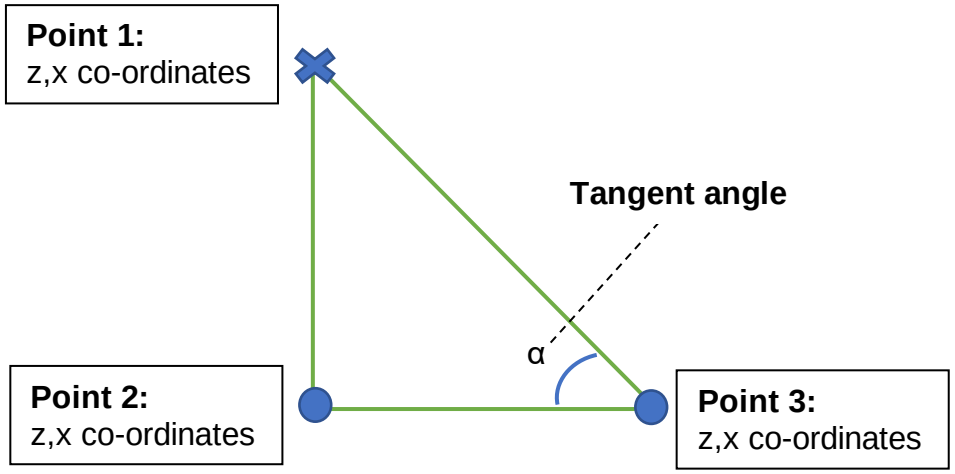


Figure 32. Spinal shape T1, T8 and L1 computed tangent angle calculation

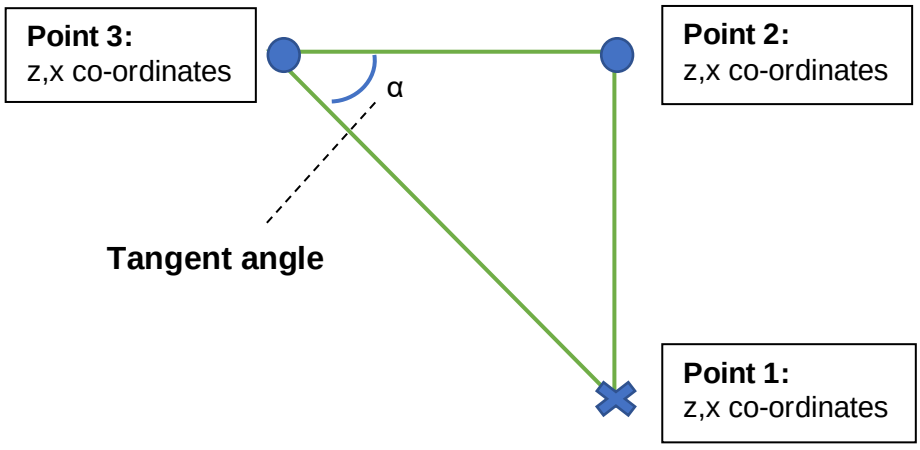
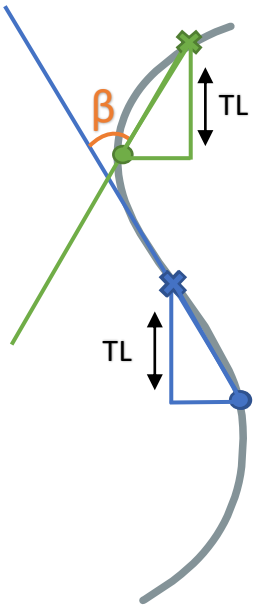


Figure 33. Spinal shape L5 computed tangent angle calculation

Conversely, this method is inverted when calculating L5 (Figure 33). During data collection, spinal sweeps ended at S2 meaning there is a lack of 'real' data values beyond L5 for the mathematical fitting function to draw from. As a result, if the tangent length was subtracted from the z,x coordinates for L5, the reliance on 'predicted values' at the edge of the fitted function could be prone to larger calculated angle errors. Instead, the tangent is added to the z,x coordinates of L5's 3-D location and the angle is computed relative to the right hand horizontal (represented by 'α' in Figure 33).

3.4.2.2 Computed resultant angle method for spinal shape

For resultant angles, the side of the triangle connecting point 1 and point 3 are extended. The angle created by the intersection of two tangents is then calculated. This is represented by 'β' in Figure 34.



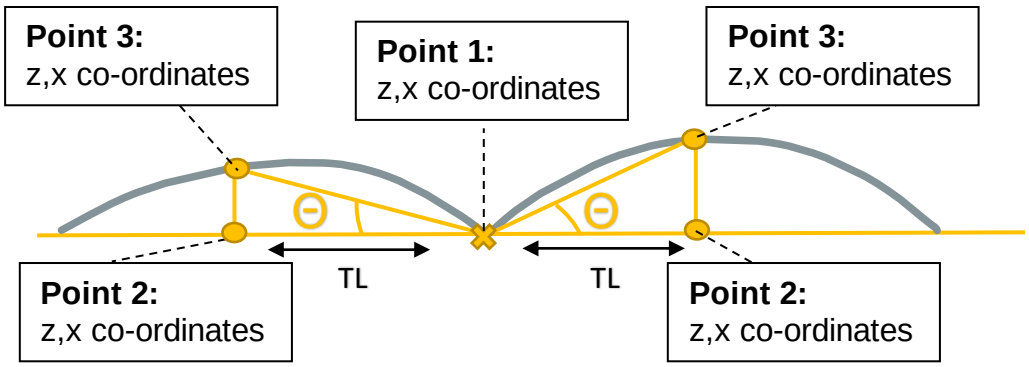
(**×**, anatomical landmark; **o**, third point of triangle generated by mathematical function; **TL**, desired tangent length; **β**, resultant angle)

Figure 34. Spinal shape computed resultant angle calculation

3.4.2.3 Computed hump horizontal angle method for trunk shape

For trunk shape, the computed method for the left and right hump horizontal angle calculations is depicted in Figure 35. Firstly, the EM system provides the location of the anatomical landmark in 3-D space from which the co-ordinates are extracted (point

1). The specified tangent length is then added or subtracted from this value to form the second point of this triangle (point 2). From this second location, the coordinates are provided from the surface fitting function, completing the triangle (point 3). The angle calculated is represented by 'Θ' in Figure 35.



(**×**, anatomical landmark; **o**, third point of triangle generated by mathematical function; **TL**, desired tangent length; **Θ**, hump horizontal angle)

Figure 35. Trunk shape computed hump horizontal angle calculation

3.4.2.4 Flexicurve method for spinal shape measurements

Understanding these computational processes enabled this to be copied during the flexicurve method of calculating angles. Figures 36 to 40 demonstrate the process of calculating spinal shape angles with the flexicurve.

Tangent angle calculations:

1. The flexicurve was moulded to the shape of the phantom model and the curve was traced onto paper (Figure 36 and 37).
2. The anatomical landmarks were plotted onto the curve, forming the first point of each triangle (Figure 38).
3. The specified tangent length was then drawn inferior (T1, T8 & L1) or superior (L5) from each anatomical landmark, forming the second point of the triangle.
4. A straight horizontal line was drawn to connect point two with the spinal curve, forming the third point of the triangle.
5. A straight line was drawn to connect point one and point three, completing the triangle.

6. A protractor was used to measure the angle relative to the left horizontal (‘α’) or right horizontal (‘β’) depending on the anatomical landmark being measured.

Resultant angle calculations:

7. The side of each triangle connecting point one and point three were extended superiorly and/or inferiorly.
8. The angle between the tangents where one line intersected with another were then measured using a protractor.

Tangent (Figure 39) and resultant (Figure 40) flexicurve angle calculations for the neutral model using the 0.05m tangent length are illustrated below as an example.

3.4.2.5 Flexicurve method for trunk shape measurements

To calculate trunk shape using the flexicurve, steps 1-6 mentioned above were also completed but in the transverse plane thus, yielding hump horizontal angles at each specific anatomical region. Examples can be seen in the results section (Figure 43 and 44).

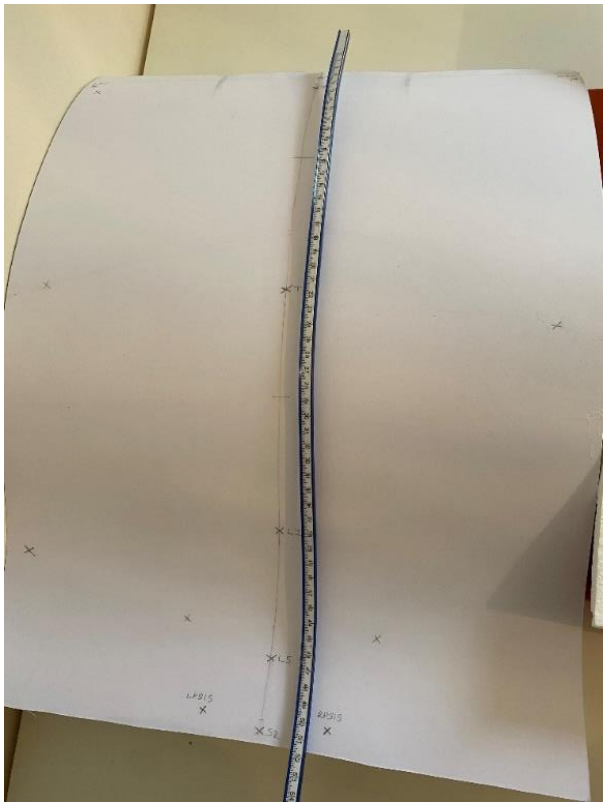


Figure 36. Moulding flexicurve

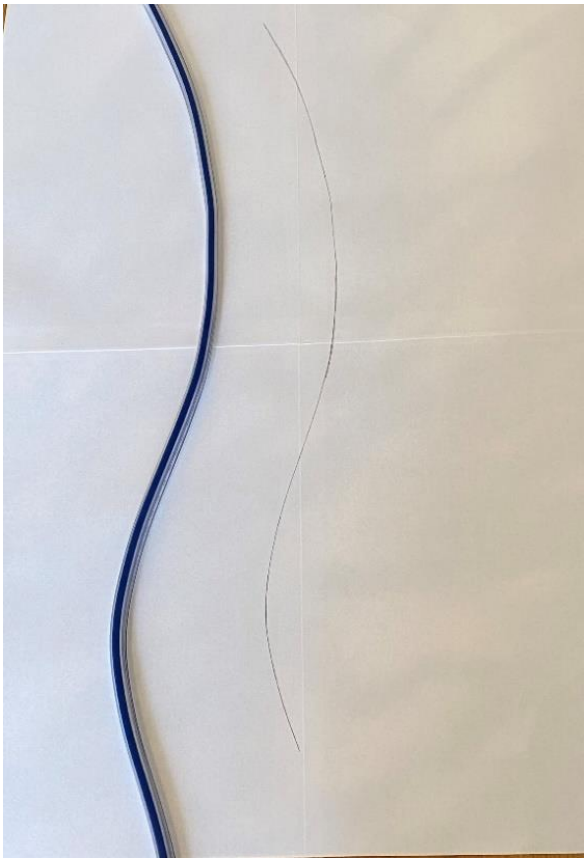


Figure 37. Tracing curve onto paper

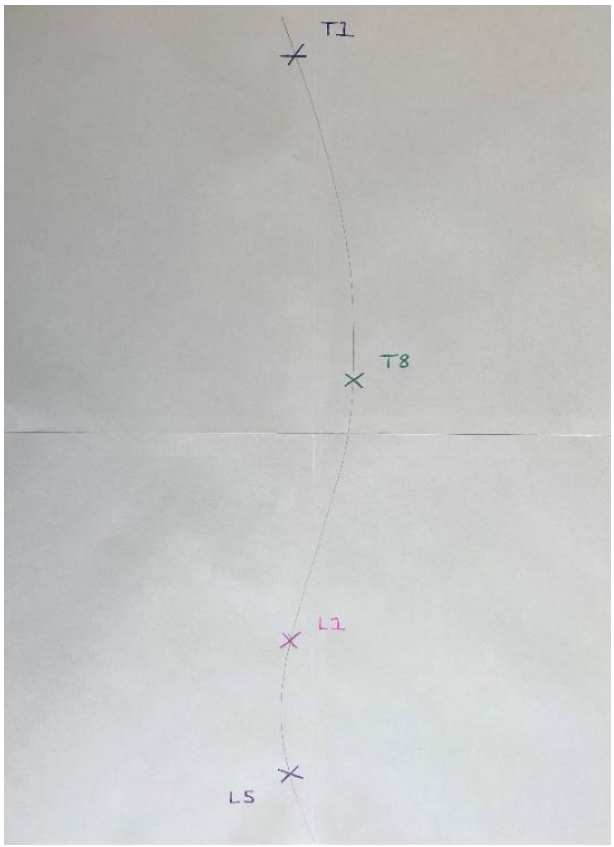
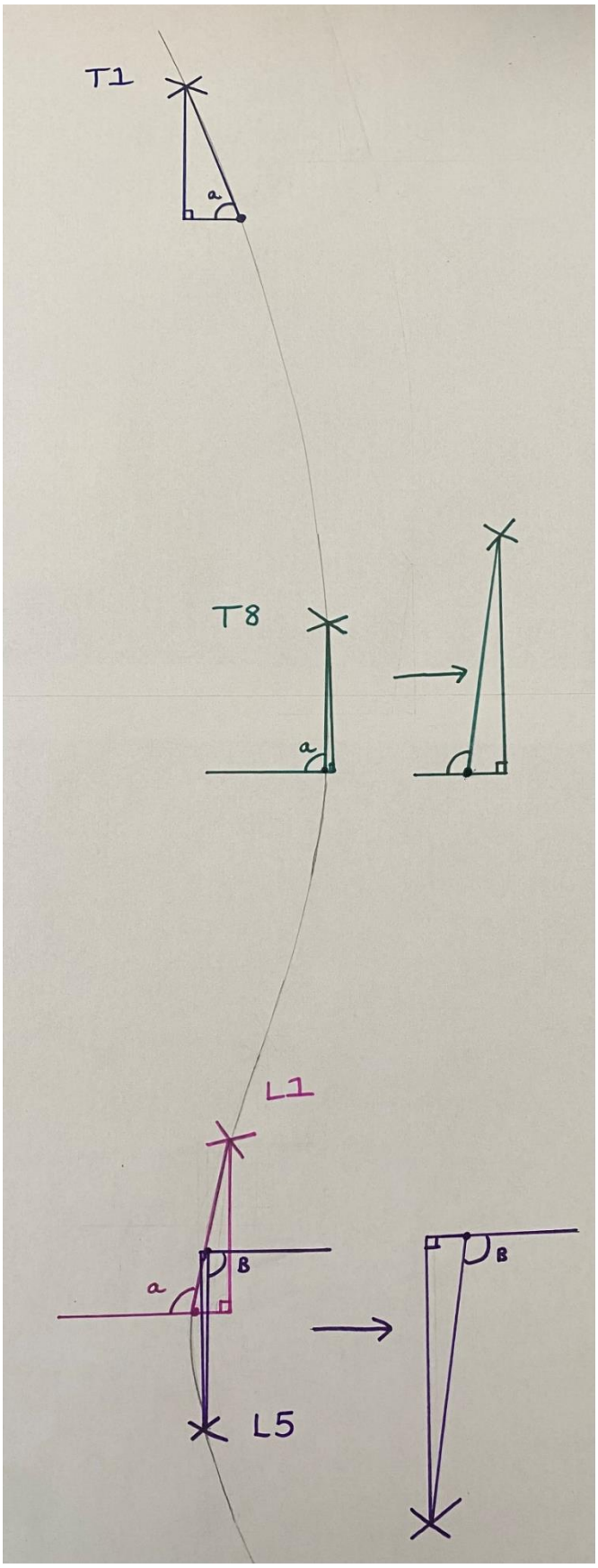
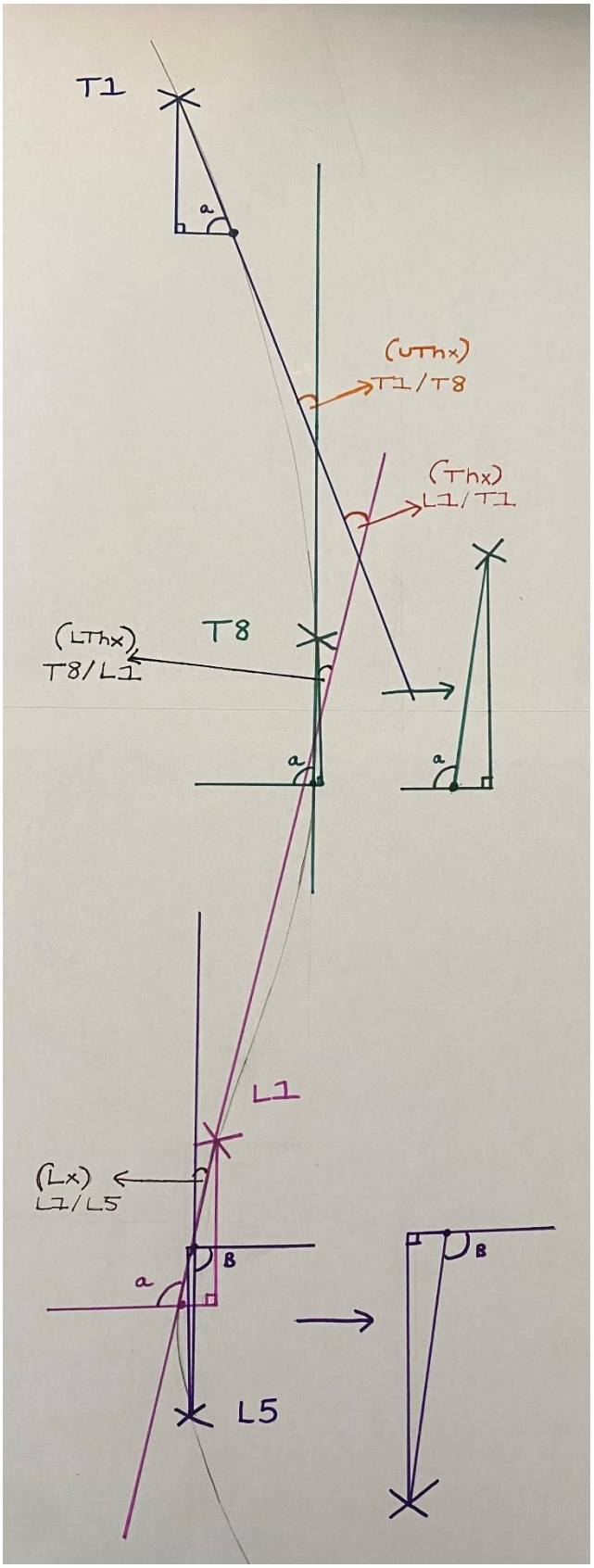


Figure 38. Plotting landmarks



(**×**, anatomical landmark; **o**, third point of triangle generated by mathematical function; **α**, tangent angle relative to left hand horizontal; **β**, tangent angle relative to right hand horizontal)

Figure 39. spinal shape flexicurve tangent angle calculations



(**×**, anatomical landmark; **o**, third point of triangle generated by mathematical function; **α**, tangent angle relative to left hand horizontal; **β**, tangent angle relative to right hand horizontal; UThx; upper thoracic; Thx, thoracic; Lthx, lower thoracic; Lx, lumbar)

Figure 40. Spinal shape flexicurve resultant angle calculations

3.5 Data processing

Data were imported into MATLAB (MathWorks 2019b) where all data processing was performed. To complete the steps outlined above, a specific MATLAB (Mathworks 2019) algorithm was written to yield either spinal shape (as probe sweep coordinates) or trunk shape (as trunk sweep coordinates). For spinal shape, the probe sweep coordinates were fitted with different polynomial functions (1st order to 8th order). From these fits, tangents in the sagittal plane were determined from which angles were derived. For trunk shape, as described in the development of methods section, the 3-D coordinates were plotted as a point cloud from which 3-D polynomials (2nd order to 5th order) and the Lowess surface functions were fitted to the data. From these surface fits, the tangent angles in the transverse plane were quantified to represent the trunk shape at specific anatomical locations. Analysis of the values yielded by different fits and functions and their effect on the spinal shape and trunk shape angles calculated were then determined.

The MATLAB (Mathworks 2019) algorithm was applied to each data trial for four different tangent lengths. The tangents were 0.025m, 0.05m, 0.075m and 0.1m (spinal shape) or 0.075m, 0.1m, 0.125m and 0.15m (trunk shape). This resulted in each repeat having 4 different sets of data. Examination of values yielded by different tangent lengths and their effect on spinal and trunk shape angles were also determined.

3.6 Data analysis

In order to determine how to select the best mathematical fitting function, the following steps were completed.

1. Determine the goodness of the fit.

The goodness of each mathematical fit for spinal and trunk shape were determined from the r^2 and RMSE results. This formed the foundation of analysis for the mathematical fit of each function and served to determine which mathematical function(s) best represented the data.

2. Determine the effect of fit on angles calculated.

Spinal and trunk angles (calculated from the intersection of two tangents) for specific fits were plotted with their 95% confidence intervals (95%CIs). This enabled the determination of any differences in the tangent and resultant angles calculated from the different fits.

3. Determine the intra-rater reliability of repeated angle measurement for the different fits.

To evaluate the consistency of spinal and trunk shape measurements with this method, the intraclass correlation coefficient (ICC 3,1) were calculated for all the trials across different tangent lengths. ICC was chosen as the reliability metric for this study as it was the most common statistic used in the literature of surface measurement methods (section 1.2). The ICC (3,1) form was selected based on the 2-way mixed effects model, mean of 3 measurements type and absolute agreement definition. ICC values were interpreted using the classification described by Koo and Li (2016) where values <0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability and values >0.90 indicate excellent reliability.

Since ICC values alone do not reveal absolute differences between measurements; the SEM and MDC were also calculated. SEM (sd x $\sqrt{}$(1-ICC)) is the determination of the amount of measurement error present in an instrument whilst MDC (SEM x 1.96 x $\sqrt{}$2) is a metric that represents true change beyond that of measurement error. Presenting SEM and MDC for this EM method facilitated the application of its reliability estimates to clinical practice.

The CoV (relative dispersion of data points around the mean) was also calculated (sd ÷ mean x 100) for the separate spinal shape trials only to investigate the reliability between the consecutive and non-consecutive data-capture methods.

4. Validity of the angles calculated for the different fits.

To evaluate how accurately the EM method measured spinal and trunk shape, angles calculated with the EM tracking system were compared against those calculated using the flexicurve method (Burton 1986; Hinman 2004b; de-Oliveira et al. 2012) to

determine validity. Visual inspection of the data provided answers as to the pattern of angle change as the tangent length increased.

From these steps conclusions over the optimal fit were made.

Chapter 4. Results

4.1 Spinal Shape

4.1.1 Mathematical fitting function

Table 17 shows the mean r^2 values for the 1st to 8th order polynomial fits of the extension, flexion and neutral models in both the consecutive and non-consecutive trials. The mean r^2 value for the 1st order polynomial fit is very low (r^2 =<0.2) across all the trials. However, for the 2nd and 3rd order polynomials there is some variation in the mean r^2 values (r^2 = 0.69-1.0) compared to the results for the 4th to 8th order polynomial fits which are more consistent (r^2 =0.99-1.0). As a result, it is recommended that any polynomial fits containing r^2 values <0.99 (i.e. the 1st, 2nd and 3rd order polynomials) should not be used for spinal shape measurements.

Table 17. Spinal shape r^2 results

		Spinal shape r2 results							
		P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**
Spinal Shape Trial	**ExtCon**	0.0089	0.9952	0.9998	0.9999	1.0000	1.0000	1.0000	1.0000
	FlexCon	0.0086	0.9773	0.9970	0.9997	0.9999	0.9999	0.9999	0.9999
	NeuCon	0.1647	0.6857	0.9349	0.9940	0.9948	0.9995	0.9998	0.9998
	ExtNC	0.0213	0.9943	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999
	FlexNC	0.0039	0.9761	0.9966	0.9996	0.9998	0.9999	0.9999	0.9999
	NeuNC	0.1968	0.6962	0.9431	0.9945	0.9948	0.9996	0.9998	0.9998

(r^2, r squared; ext, extension; flex, flexion; neu, neutral; con, consecutive trials; NC, non-consecutive trials; P1, 1st order polynomial; P2, 2nd order polynomial; P3, 3rd order polynomial; P4, 4th order polynomial; P5, 5th order polynomial; P6, 6th order polynomial; P7, 7th order polynomial; P8, 8th order polynomial)

Since there were little differences in the r^2 value for the 4th, 5th, 6th, 7th and 8th order polynomial fits (r^2 = 0.994-1.0), the mean and 95%CIs of angles computed for these fits were subsequently studied. Analysis was conducted on the 0.025m tangent length non-consecutive trials for the T1, LThx, L1 and L5 angle variables. An example of this analysis is shown in Figure 41 for the L1 tangent angle data of the neutral model. This figure is representative of the most common pattern seen in the data for the angle variables analysed. This figure clearly shows the 95%CIs for the 4th and 5th order

polynomials do not overlap which is also true of the 5th and 6th order polynomial fits. This indicates difference (at the 95% confidence level) between these angles computed. However, the 95%CIs of the 6th, 7th and 8th order polynomial fits are all overlapping, suggesting no differences between these computed angles.

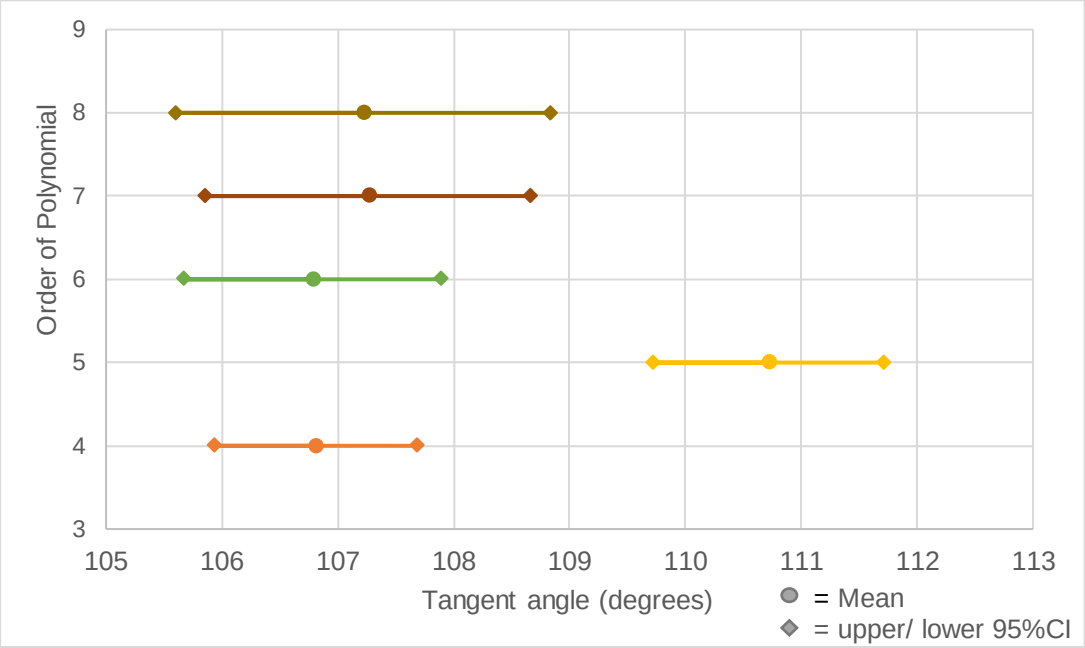


Figure 41. L1 tangent angle mean and 95% confidence intervals for the 4th to 8th order polynomial fits of the neutral spine shape with the 0.025m tangent length

Neutral spinal shape tangent and resultant angles of LThx, L1 and L5 showed differences between the angle calculated for the 4th and 5th as well as, the 5th and 6th order polynomials whereas, for the calculated angle at T1 only the 5th order polynomial appeared different to 6th. For extension, differences were seen between the 4th and 5th and 5th and 6th order polynomial data for the L1 tangent angle only, whilst flexion also demonstrated a difference between the 5th and 6th order polynomial tangent angle values at L1. Overall, the data seemed to detect that, across the angle variables analysed, there were consistently no differences between the values calculated for the 6th, 7th and 8th order polynomial fits. However, there were some differences seen between the 4th and 5th or 5th and 6th order polynomial fits, depending on the spinal shape model and angle variable studied.

Based on this, the 6th order polynomial fitting function can be recommended for use for spinal shape measurements with an EM system. This can be concluded since there were some differences seen between angles calculated for the 4th and 5th and 5th and

6th order polynomial fits, but no differences detected between the 6th, 7th or 8th order polynomial fits. Therefore, since the 7th and 8th order polynomial fitting functions require additional mathematical processes but do not yield significantly different results to the 6th order polynomial, the 6th order polynomial should be used.

4.1.2 Tangent length

As the tangent length increases, the data shows the computed angle changes. The pattern of change shows either an increase or decrease in the angle calculated depending on the variable analysed. Although angle change is evident, the differences between the angles calculated at different tangent lengths is low. For example, the mean of the sd results for angles calculated for the 6th order polynomial in the three sets of trials across all four tangent lengths for the consecutive extension data is 0.01º. Therefore, although it cannot be recommended to use tangent lengths interchangeably, the tangent length chosen should not significantly affect spinal shape measurements if one tangent length is used consistently.

4.1.3 Reliability

Two different measurement methods were used for each set of the three spinal shape trials. The 'consecutive' trials consisted of completing the three spinal shape measurements of one phantom model without moving the equipment or phantom. The other set of 3 repeats occurred 'non-consecutively' where spinal shape was measured once on each phantom model and then repeated.

To assess the reliability between these methods, the CoV were calculated for both sets of trials and then compared. Due to the recommendations mentioned above, the CoV data of computed angles with the 6th polynomial fit for one tangent length (0.05m) were examined. These results are shown in Table 18.

Overall, all CoV values are moderate to low (<15%) with most angle variables across all the trials yielding a CoV <1%. The higher CoV values are seen in the non-consecutive trials particularly for the neutral phantom model. However, since the variation around the mean for both the consecutive and non-consecutive trials is low, this indicates that the consecutive method does not produce significantly more reliable

results from the non-consecutive method. As a result, one set of data collection can be recommended and since the non-consecutive method is more reflective of clinical applications, this is an appropriate approach for spinal shape measurements.

Table 18. Spinal shape CoV results

		CoV results (%)							
		T1	**T8**	**L1**	**L5**	**Uthx**	**Thx**	**LThx**	**Lx**
Spinal Shape Trial	**ExtCon**	0.204	0.204	0.395	0.255	0.675	0.419	0.288	1.918
	ExtNC	0.431	0.242	0.283	0.253	1.687	0.719	1.056	2.627
	FlexCon	0.332	0.172	0.062	0.139	-0.050	-0.139	-0.429	-1.092
	FlexNC	1.277	0.214	0.107	0.357	-1.505	-0.880	-0.258	-4.326
	NeuCon	0.148	0.031	0.135	0.248	-0.388	-0.350	-1.648	0.606
	NeuNC	2.529	0.391	0.473	0.204	-4.105	-5.150	-13.415	1.785

(CoV, co-efficient of variation; %, percent; L, lumbar; LThx, lower thoracic; Lx, lumbar; T, thoracic; Thx, thoracic; Uthx, upper thoracic; ext, extension; flex, flexion; neu, neutral; con, consecutive trials; NC, non-consecutive trials; -, minus)

The ICC, SEM and MDC results for spinal shape measurements of the consecutive and non-consecutive trials of extension, flexion and neutral models are shown in Table 19. The ICC's were extremely high (>0.999) regardless of tangent length. Furthermore, extremely small values were seen for the MDC (<0.018º) and SEM (<0.007º). This shows that angles calculated with the EM method were highly reliable for spinal shape measurements.

Table 19. Spinal shape ICC, MDC and SEM results for 6th order polynomial fit

	Tangent Length (m)			
Reliability	**0.025**	**0.05**	**0.075**	**0.1**
ICC	0.9998	0.9998	0.9998	0.9997
MDC (º)	0.0138	0.0126	0.0118	0.0168
SEM (º)	0.0050	0.0045	0.0043	0.0061

(ICC, intra-class correlation coefficient; MDC, minimal detectable change; SEM, standard error of measurement; º, degrees; m, metres)

4.1.4 Validity

Tangent and resultant angles from one consecutive 0.05m tangent spinal shape trial (SSE1/ SSF1/ SSN1) for each model were compared against corresponding angles calculated with the flexicurve method (Table 20 and 21). Based on conclusions above, only the computed angles for the 6th order polynomial fit were used.

The results showed that the angles calculated with the flexicurve method were no more than 6.0º different to the computed angles calculated by the EM system for spinal shape. Consequently, it can be concluded that the EM tracking method is valid for measurements of spinal shape. Potential sources of difference are discussed later in chapter 5 (section 5.5). The results also showed that resultant angles calculated with the EM system for flexion (kyphotic) curves will result in a negative value and for extension (lordotic) curves, the value is defined as positive.

Table 20. Comparison of tangent angles calculated between EM method and flexicurve for spinal shape

Model and tangent angle	Flexicurve (º)	6th order polynomial fit (º)	Difference (º)
Extension			
T1	133	129.621	3.379
T8	87	85.784	1.216
L1	64	59.937	4.063
L5	63	57.704	5.296
Flexion			
T1	47	41.422	5.578
T8	87	86.249	0.751
L1	116	114.225	1.775
L5	118	122.527	4.527
Neutral			
T1	69	66.041	2.959
T8	91	96.069	5.069
L1	102	103.650	1.650
L5	86	91.343	5.343

(º, degrees; T, thoracic; L, lumbar)

Table 21. Comparison of resultant angles calculated between EM method and flexicurve for spinal shape

Model and resultant angle	Flexicurve (º)	6th order polynomial fit (º)	Difference (º)
Extension			
T1/T8 (UThx)	42.5	43.837	1.337
T1/L1 (Thx)	66	69.685	3.685
T8/ L1 (LThx)	29	25.848	3.152
L1/ L5 (Lx)	1	2.233	1.233
Flexion			
T1/T8 (Uthx)	-39	-44.827	-5.827
T1/L1 (Thx)	-67	-72.803	-5.803
T8/ L1 (LThx)	-29	-27.976	-1.024
L1/ L5 (Lx)	-7	-8.302	-1.302
Neutral			
T1/T8 (UThx)	-25	-30.028	-5.028
T1/L1 (Thx)	-32	-37.609	-5.609
T8/ L1 (LThx)	-9	-7.581	-1.419
L1/ L5 (Lx)	13	12.306	0.694

(º, degrees; T, thoracic; /, slash; L, lumbar; UThx, upper thoracic; Thx, thoracic; LThx, lower thoracic; Lx, lumbar; -, minus)

4.1.5 Spinal Shape - summary of recommendations and conclusions

Mathematical fitting function

Polynomial fits 1, 2 and 3 that have data containing an r^2 value <0.99 should not be used for spinal shape measurements.

There were consistently no differences seen between the computed angles calculated for the 6th, 7th and 8th order polynomial fits but some differences were seen between the 4th and 5th or 5th and 6th order polynomial fits.

Therefore, the 6th order polynomial fit is recommended for spinal shape measurements with an EM system.

Tangent length

The computed angle changed as the tangent length increased.

The pattern of change was different depending on the angle variable analysed.

It is recommended that tangent lengths should not be used interchangeably.

However, the tangent length chosen should not significantly affect spinal shape measurements with the EM method if used consistently.

Validity

Differences between angles calculated with the EM method and flexicurve method were <6.0º.

Resultant angles were positive for extension or lordotic curves and negative for flexion or kyphotic curves.

Spinal shape angles calculated with the EM method were valid.

Reliability

The EM method was highly reliable (ICC = >0.999, CoV = <15%) and highly accurate (MDC = <0.018º, SEM = <0.007º).

The consecutive data collection technique did not produce significantly more reliable results than the non-consecutive technique.

One set of data can be recommended for collection of spinal shape measurements with the EM method.

4.2 Trunk Shape

4.2.1 Mathematical fitting function

The mean r^2 and RMSE data of each mathematical function for the extension, flexion and neutral trunk shape trials are seen in Table 22 and 23. The results show more variability in the mean r^2 values for polynomial 22 (r^2 = 0.69-0.99) and 33 (r^2 = 0.94-1.0) functions compared to polynomial 44, 55 and Lowess fits which have more consistent results (r^2= 0.98-1.0).

Table 22. Trunk shape r^2 results

		r2 results for each mathematical function				
		L	**P22**	**P33**	**P44**	**P55**
Trunk shape trial	**ExtCon**	0.9887	0.9870	0.9907	0.9914	0.9920
	FlexCon	0.9931	0.9744	0.9957	0.9980	0.9981
	NeuCon	0.9767	0.7073	0.9420	0.9832	0.9864
	ExtNC	0.9928	0.9920	0.9959	0.9961	0.9964
	FlexNC	0.9936	0.9752	0.9962	0.9986	0.9987
	NeuNC	0.9761	0.6878	0.9447	0.9819	0.9867

(r^2, r squared; ext, extension; flex, flexion; neu, neutral; con, consecutive trials; NC, non-consecutive trials; L, lowess function; P22, polynomial 22 function; P33, polynomial 33 function; P44, polynomial 44 function; P55, polynomial 55 function)

Table 23. Trunk shape RMSE results

		RMSE results for each mathematical function (º)				
		L	**P22**	**P33**	**P44**	**P55**
Trunk shape trial	**ExtCon**	0.0038	0.0040	0.0034	0.0032	0.0031
	FlexCon	0.0040	0.0077	0.0031	0.0021	0.0020
	NeuCon	0.0035	0.0124	0.0055	0.0029	0.0026
	ExtNC	0.0031	0.0032	0.0022	0.0021	0.0020
	FlexNC	0.0038	0.0075	0.0029	0.0018	0.0017
	NeuNC	0.0035	0.0126	0.0053	0.0030	0.0026

(RMSE, root mean square error; º, degrees; ext, extension; flex, flexion; neu, neutral; con, consecutive trials; NC, non-consecutive trials; L, lowess function; P22, polynomial 22 function; P33, polynomial 33 function; P44, polynomial 44 function; P55, polynomial 55 function)

This pattern is also observed in the RMSE results where polynomial 44, 55 and Lowess functions have similar outcomes (RMSE = 0.002º to 0.004º) but higher values and more variation is seen for polynomial 22 (RMSE = 0.003º to 0.013º) and polynomial 33 (RMSE = 0.002º to 0.006º) fits. As a result, it is recommended that any surface smoothing functions containing r^2 values <0.97 and RMSE values >0.005º (i.e. polynomial 22 and 33 functions) should not be used for trunk shape measurements.

The mean and 95%CIs of T1, T8, L1 and L5 angle variables calculated with the 0.075m tangent length in the non-consecutive trunk shape trials were subsequently analysed. On visual inspection, there were consistently no differences seen between the computed angles calculated for polynomial 44, 55 and Lowess functions except for the left hump angle at L5 for the neutral phantom model. An example of this analysis is shown in Figure 42 for the right T8 angle data of the neutral model. This figure is representative of the most common pattern seen in the data for the angle variables analysed.

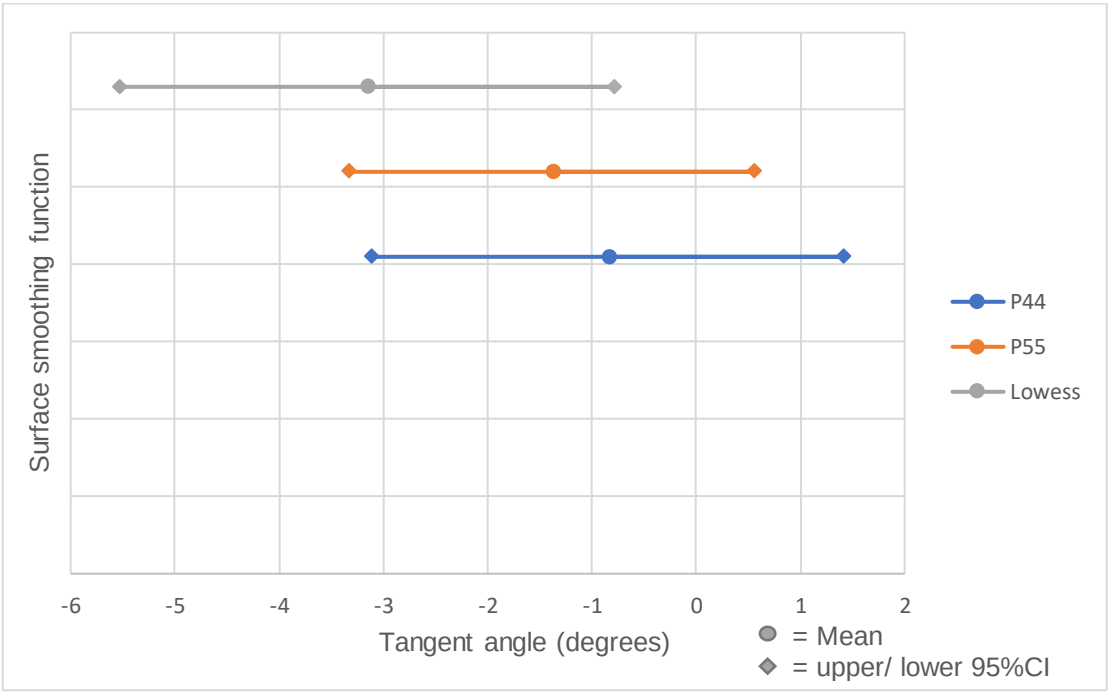


Figure 42. Right hump horizontal T8 angle mean and 95% confidence intervals for the polynomial 44, 55 and Lowess fits of the neutral spine shape with 0.075m tangent length

This graph demonstrates that the 95%CIs for all three surface fitting functions overlap, indicating no difference between the computed angles. Since this is a common pattern

across this data, no further conclusions can be drawn as to which mathematical function for trunk shape data is best from this analysis alone.

In the scoliosis trials, the results also showed high r^2 (0.98-0.99) and low RMSE (0.01º) values for polynomial 44, 55 and Lowess functions (Table 24 and 25). However, analysis of the computed angles calculated between the functions found differences between polynomial 44 and 55 for some angle variables. There were consistently no differences seen between calculated angles from polynomial 55 and Lowess fits. Therefore, from these results it can be concluded that the polynomial 44 function should also not be used for trunk shape measurements.

Table 24. Scoliosis trunk shape r^2 results

		r^2 result for each mathematical function		
		Lowess	**Polynomial44**	**Polynomial55**
Scoliosis trial	**1**	0.9817	0.9789	0.9874
	2	0.9792	0.9768	0.9864
	3	0.9830	0.9803	0.9870
	4	0.9815	0.9776	0.9852
	5	0.9852	0.9766	0.9848
	6	0.9840	0.9763	0.9858

(r^2; r squared)

Table 25. Scoliosis trunk shape RMSE results

		RMSE for each mathematical function (º)		
		Lowess	**Polynomial44**	**Polynomial55**
Scoliosis trial	**1**	0.0106	0.0113	0.0088
	2	0.0109	0.0115	0.0088
	3	0.0104	0.0112	0.0091
	4	0.0101	0.0111	0.0090
	5	0.0097	0.0122	0.0099
	6	0.0104	0.0126	0.0098

(RMSE, root mean square error; º, degrees)

In summary, polynomial 22 and 33 functions are not recommended for trunk shape measurements due to the poorer results and higher variability demonstrated in the r^2 and RMSE results of the extension, flexion and neutral trials. For the same trials there were commonly no differences seen between the angles calculated for polynomial 44, 55 and Lowess fits. However, in the scoliosis trials differences were seen between angles calculated for polynomial 44 and 55 functions. As a result, a further recommendation of not using the polynomial 44 function for trunk shape measurements can be made.

4.2.2 Validity

Using the non-consecutive data in the extension, flexion and neutral trials, the left and right horizontal kyphosis angles calculated for T8 at each tangent length were compared against corresponding angles obtained using the flexicurve method (Figure 43). These results are seen in Table 26 for the polynomial 55 function and Table 27 for the Lowess function. Computed angles, for both mathematical functions, were no more than 3.0º different to angles calculated with the flexicurve. This demonstrates the validity of trunk shape measurements with the EM method.

However, it should be noted that the trunk shape does not change as you move horizontally across the flexion, extension or neutral phantom surfaces therefore, these hump angles should be relatively flat. Based on this notion, angles computed by the EM system and those measured with the flexicurve should be close to zero but, looking at Tables 26 and 27 this was not the case. Therefore, errors were evident in both methods for measuring angles of a flat surface. Examining the mean angles calculated for T1, T8, L1 and L5 with the 0.075m tangent length, it appears that angles calculated for the Lowess function were most commonly the furthest from zero. This indicates that the polynomial 55 function may be the best representation of known truth.

On the other hand, results from the scoliosis trials may be more indicative of the validity of angles calculated by each mathematical function. Tables 28 and 29 compare the mean values computed by the Lowess and polynomial 55 functions against those calculated with the flexicurve (Figure 44). Here the results show better validity for the Lowess function (largest difference = 4.8º) compared to the polynomial 55 fit (largest difference = 7.4º).

Overall, it can be concluded that the EM tracking method is valid for measurements of trunk shape. However, the polynomial 55 function was likely the better representation associated with known truth for ‘flat hump’ surfaces. Nevertheless, the Lowess function had higher validity for measurements of an altered trunk surface when compared against the flexicurve.

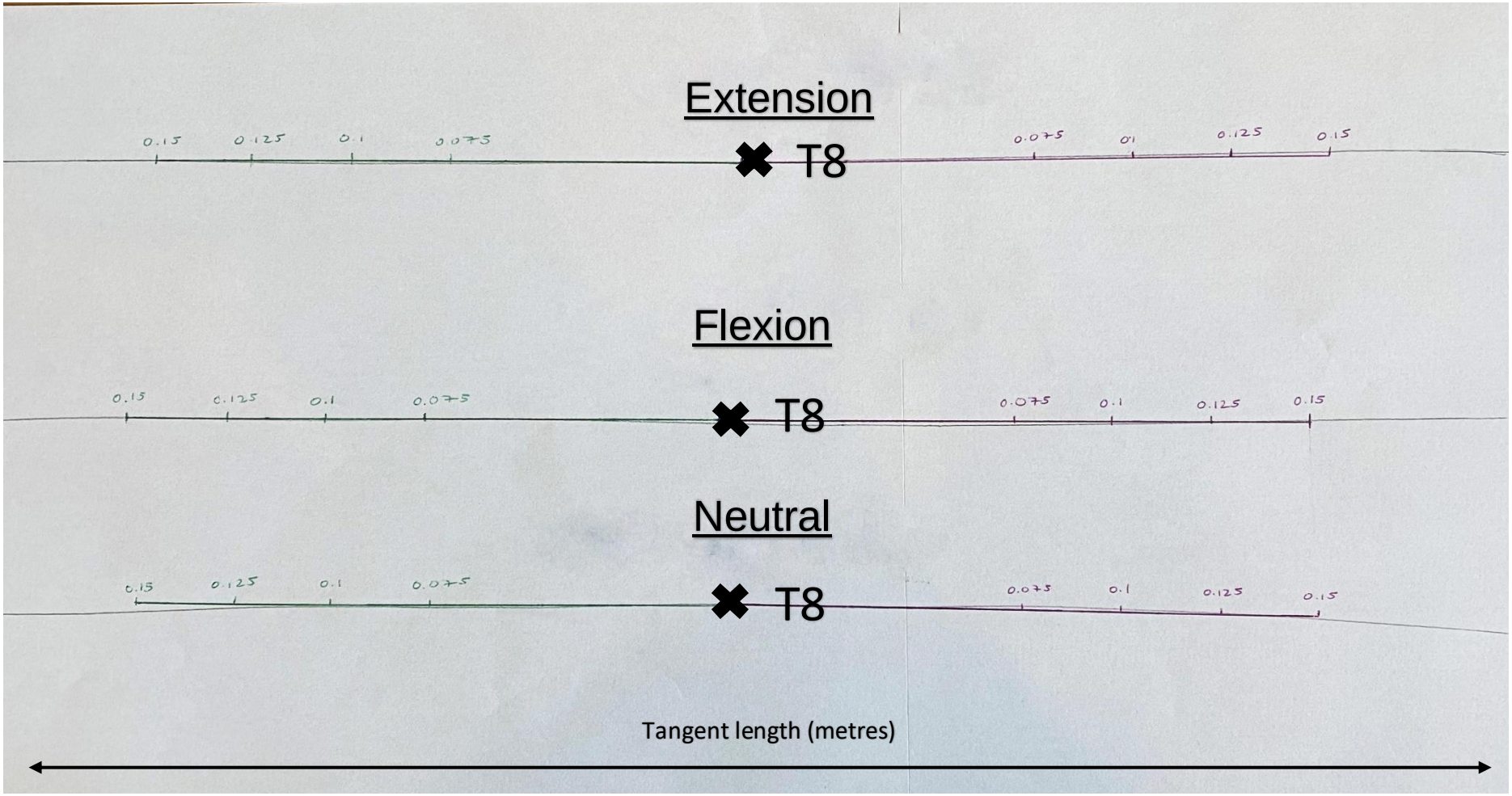


Figure 43. Extension, flexion and neutral T8 hump horizontal flexicurve angle calculations

Table 26. Comparison of T8 hump horizontal angles calculated between flexicurve and EM method using P55 fit for trunk shape of extension, flexion and neutral phantoms

	Left T8 hump horizontal angle					Right T8 hump horizontal angle			
Flexion	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**	**Flexion**	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**
	0.075	-1	-0.419	-0.581		0.075	0	0.336	0.336
	0.1	-1	-0.308	-0.692		0.1	0	0.546	0.546
	0.125	0	-0.256	0.256		0.125	0	0.701	0.701
	0.15	0	-0.244	0.244		0.15	0	0.804	0.804
Extension	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**	**Extension**	Tangent (m)	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**
	0.075	1	1.481	0.481		0.075	-0.5	-1.364	0.864
	0.1	1	1.441	0.441		0.1	-0.5	-1.400	0.900
	0.125	1	1.429	0.429		0.125	0	-1.463	1.463
	0.15	1.5	1.461	0.039		0.15	0	-1.581	1.581
Neutral	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**	**Neutral**	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**
	0.075	-1	-0.303	-0.697		0.075	-0.5	-1.378	0.878
	0.1	0.5	0.121	0.379		0.1	0	-1.264	1.264
	0.125	0	0.487	0.487		0.125	-0.5	-1.251	0.751
	0.15	0.5	0.827	0.327		0.15	-1	-1.273	0.273

(T, thoracic; m, metres; º, degrees; P55, polynomial 55 function; -, minus)

Table 27. Comparison of T8 hump horizontal angles calculated between flexicurve and EM method using Lowess fit for trunk shape of extension, flexion and neutral phantoms

	Left T8 hump horizontal angle					Right T8 hump horizontal angle			
Flexion	**Tangent (m)**	**Flexicurve (º)**	**Lowess (º)**	**Difference (º)**	**Flexion**	**Tangent (m)**	**Flexicurve (º)**	**Lowess (º)**	**Difference (º)**
	0.075	-1	-3.710	2.710		0.075	0	-3.053	3.053
	0.1	-1	-3.299	2.299		0.1	0	-2.726	2.726
	0.125	0	-2.302	2.302		0.125	0	-1.564	1.564
	0.15	0	-1.883	1.883		0.15	0	-1.090	1.090
Extension	**Tangent (m)**	**Flexicurve (º)**	**Lowess (º)**	**Difference (º)**	**Extension**	**Tangent (m)**	**Flexicurve (º)**	**Lowess (º)**	**Difference (º)**
	0.075	1	3.917	2.917		0.075	0.5	1.284	0.784
	0.1	1	3.557	2.557		0.1	0.5	0.811	0.311
	0.125	1	2.837	1.837		0.125	0	0.016	0.016
	0.15	1.5	2.560	1.060		0.15	0	-0.476	-0.476
Neutral	**Tangent (m)**	**Flexicurve (º)**	**Lowess (º)**	**Difference (º)**	**Neutral**	**Tangent (m)**	**Flexicurve (º)**	**Lowess (º)**	**Difference (º)**
	0.075	-1	-1.814	0.814		0.075	-0.5	-3.152	2.652
	0.1	-0.5	-1.223	0.723		0.1	0	-3.008	3.008
	0.125	0	-0.158	0.158		0.125	-0.5	-2.357	1.857
	0.15	0.5	0.321	0.179		0.15	-1	-2.298	1.298

(T,thoracic; m, metres; º, degrees; Lowess, Lowess function; -, minus)

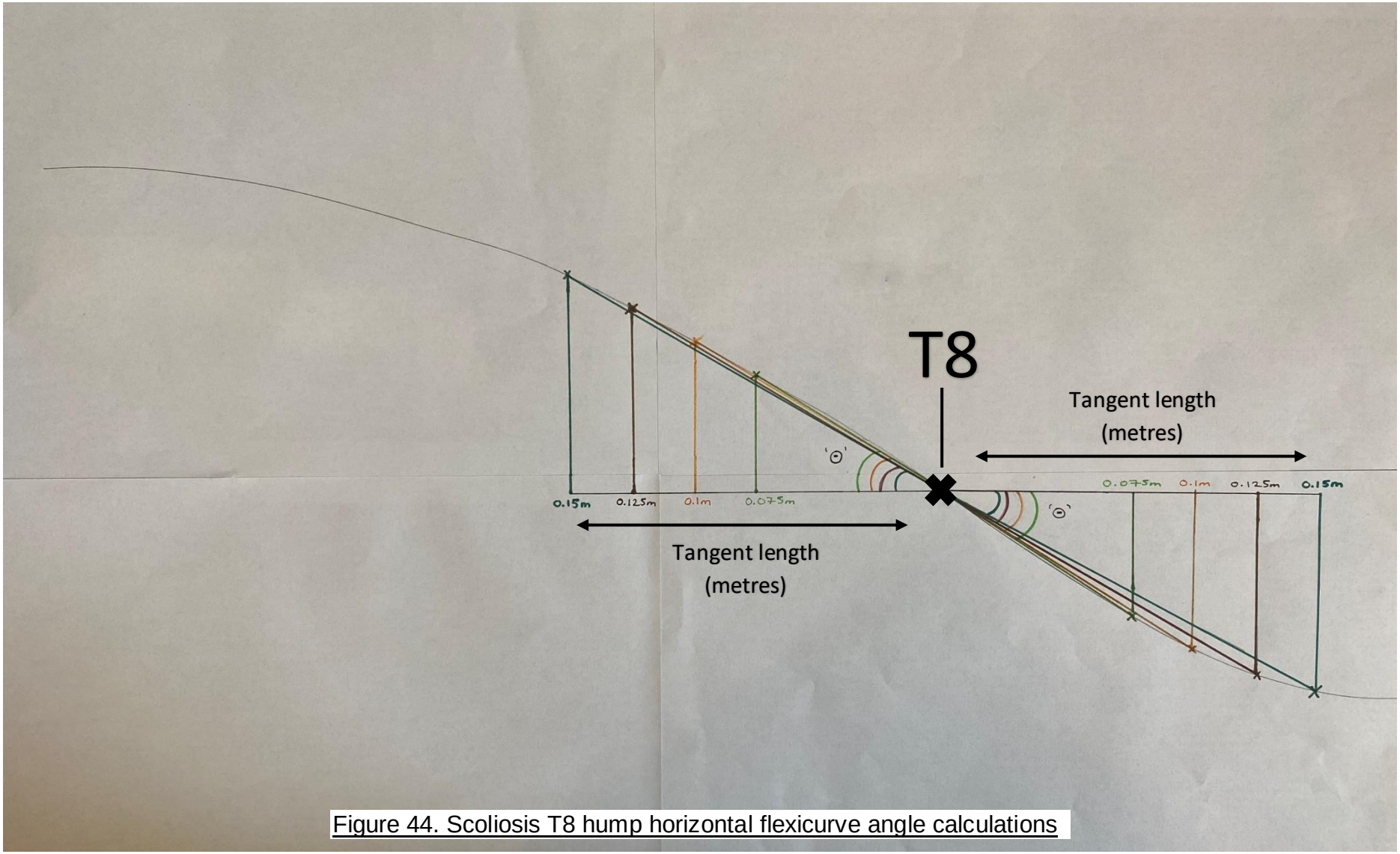


Figure 44. Scoliosis T8 hump horizontal flexicurve angle calculations

Table 28. Comparison of T1, T8 and L1 angles calculated between flexicurve and EM method using P55 fit for trunk shape of scoliosis phantom

	Left hump horizontal angle					Right hump horizontal angle			
T1	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**	**T1**	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**
	0.075	29	31.867	2.867		0.075	-43	-42.994	-0.006
	0.1	27	33.319	6.319		0.1	-42	-40.602	-1.398
	0.125	26	33.085	7.085		0.125	-40	-39.147	-0.853
	0.15	25	32.375	7.375		0.15	-36	-37.686	-1.686
T8	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**	**T8**	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**
	0.075	33	33.697	0.697		0.075	-36	-41.819	-5.819
	0.1	32	34.110	2.110		0.1	-35	-40.127	-5.127
	0.125	30	33.674	3.674		0.125	-34	-38.833	-4.833
	0.15	28	32.822	4.822		0.15	-31	-37.494	-6.494
L1	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**	**L1**	**Tangent (m)**	**Flexicurve (º)**	**P55 (º)**	**Difference (º)**
	0.075	34	32.351	1.649		0.075	-41	-41.976	-0.976
	0.1	31	32.992	1.992		0.1	-39	-40.076	-1.076
	0.125	29	32.690	3.690		0.125	-37	-38.685	-1.685
	0.15	28	31.934	3.934		0.15	-34	-37.278	-3.278

(T,thoracic; L, lumbar; m, metres; º, degrees; P55, polynomial 55 function; -, minus)

Table 29. Comparison of T1, T8 and L1 angles calculated between flexicurve and EM method using Lowess fit for trunk shape of scoliosis phantom

Left hump horizontal angle					Right hump horizontal angle				
	Tangent (m)	Flexicurve (º)	Lowess (º)	Difference (º)		Tangent (m)	Flexicurve (º)	Lowess (º)	Difference (º)
T1	0.075	29	32.329	3.329	**T1**	0.075	-43	-46.915	-3.915
	0.1	27	31.764	4.764		0.1	-42	-44.132	-2.132
	0.125	26	30.278	4.278		0.125	-40	-41.840	-1.840
	0.15	25	28.723	3.723		0.15	-36	-39.306	-3.306
	Tangent (m)	Flexicurve (º)	Lowess (º)	Difference (º)		Tangent (m)	Flexicurve (º)	Lowess (º)	Difference (º)
T8	0.075	33	31.656	1.344	**T8**	0.075	-36	-38.232	-2.232
	0.1	32	30.623	1.377		0.1	-35	-37.489	-2.489
	0.125	30	29.198	0.802		0.125	-34	-36.658	-2.658
	0.15	28	27.831	0.169		0.15	-31	-35.623	-4.623
	Tangent (m)	Flexicurve (º)	Lowess (º)	Difference (º)		Tangent (m)	Flexicurve (º)	Lowess (º)	Difference (º)
L1	0.075	34	31.857	2.143	**L1**	0.075	-41	-42.685	-1.685
	0.1	31	30.963	0.037		0.1	-39	-40.433	-1.433
	0.125	29	29.690	0.690		0.125	-37	-38.483	-1.483
	0.15	28	28.348	0.348		0.15	-34	-36.086	-2.086

(T, thoracic; L, lumbar; m, metres; º, degrees; Lowess, Lowess function; -, minus)

4.2.3 Tangent length

The extension, flexion and neutral trials indicated that slightly different angles are calculated for different tangent lengths. However, the variability was low suggesting that the tangent length does not have a significant effect on trunk shape measurements. A good example of this is shown in Figure 45 for the left T1 hump horizontal angle of an extension trial (TSE3) whereby, the angles calculated with the polynomial 55 function across the 4 different tangent lengths yield a sd of 0.515º. However, this conclusion is in relation to models with a 'flat hump' surface. As a result, analysis of data from the scoliosis trials will provide better conclusions for the effect of tangent length on trunk shape measurements made with the EM method.

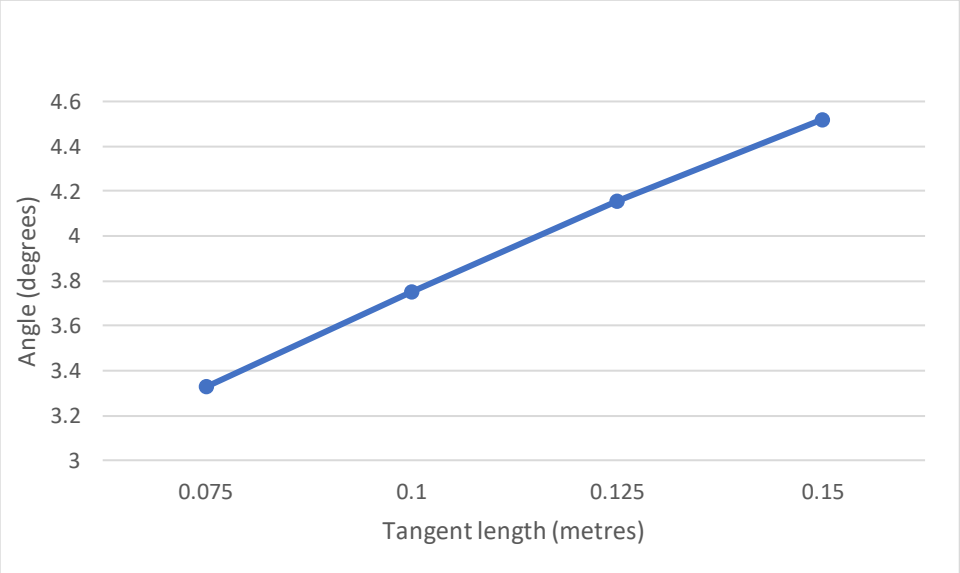


Figure 45. Left T1 hump horizontal angles calculated across the different tangent lengths for one extension trunk shape trial using the polynomial 55 function

Based on recommendations above, investigation of the mean 'scoliosis' angles calculated with the Lowess and polynomial 55 functions for different tangent lengths were conducted. Most commonly, the trend indicated that the hump horizontal angle decreased as the tangent length increased. However, there was frequently low variability seen in the angles calculated for each tangent length. This is demonstrated by the mean sd of polynomial 55 angles calculated at each tangent length being <1.5º. Thus, it can be concluded that the tangent length chosen for trunk shape

measurements should not affect results as long as the measurer uses the tangent length consistently.

To summarise, the sd of angles calculated with different tangent lengths was low for the extension, flexion and neutral trials. Nonetheless, these models had a 'flat hump' surface so, conclusions from the scoliosis trials are likely to be better. In the scoliosis trials there was generally a pattern of decrease in the hump horizontal angle as the tangent length increased. Even so, low variability between the angles at different tangent lengths was found suggesting that the tangent length chosen shouldn't significantly affect trunk shape measurements as long as it is used consistently.

4.2.4 Reliability

The ICC's for polynomial 55 and Lowess functions were calculated for the extension, flexion and neutral trials at each tangent length and then used to also calculate the MDC and SEM (Table 30). ICCs were shown to be excellent for the Lowess function (0.994-0.996) but poor for polynomial 55 (0.464-0.498). Smaller MDC and SEM values were also seen for Lowess (MDC = 0.18º to 0.2º, SEM = 0.06º to 0.07º) compared to the polynomial 55 function (MDC = 2.1º to 2.2º, SEM = 0.7º to 0.8º). Although errors were low for both fits showing trunk shape angles calculated with the EM system were highly consistent, the reliability estimates and absolute reliability were better for the Lowess function.

Table 31 shows the ICC, MDC and SEM for polynomial 55 and Lowess mathematical functions in the scoliosis trials. Good-to-excellent ICC results (0.809-0.999) were found for both mathematical functions with excellent results seen for the 0.075m tangent length. SEM results were similar for both functions (0.07º to 1.4º) but MDC values slightly lower for the polynomial 55 fit (MDC = 0.2º to 1.6º) compared to the Lowess function (MDC = 0.3º to 4.0º). However, both functions demonstrated similar MDCs for the 0.075m tangent and generally, reliability results worsened as the tangent length increased. This demonstrates that the EM system is capable of measuring an altered trunk shape and is able to detect changes to trunk shape.

Altogether, the results indicated that for trunk surfaces the Lowess function is the most reliable. For an altered trunk shape both mathematical functions had good-to-excellent

reliability and low SEM however, lower reliability estimates may be yielded by bigger tangent lengths.

Table 30. Trunk shape ICC, MDC and SEM results for polynomial 55 and Lowess fit of extension, flexion and neutral phantoms

			Tangent Length (m)			
		Reliability	**0.075**	**0.1**	**0.125**	**0.15**
Mathematical smoothing function	**P55**	**ICC**	0.4967	0.4636	0.4663	0.4979
		MDC (º)	2.2083	2.1387	2.1134	2.0693
		SEM (º)	0.7967	0.7716	0.7625	0.7465
	Lowess	**ICC**	0.9959	0.9956	0.9951	0.9944
		MDC (º)	0.1760	0.1752	0.1830	0.1981
		SEM (º)	0.0635	0.0632	0.0660	0.0715

(P55, polynomial 55 function; ICC, intra-class correlation coefficient; MDC, minimal detectable change; SEM, standard error of measurement; º, degrees; m, metres)

Table 31. Trunk shape ICC, MDC and SEM results for polynomial 55 and Lowess fit of scoliosis phantom

			Tangent Length (m)			
		Reliability	**0.075**	**0.1**	**0.125**	**0.15**
Mathematical smoothing function	**P55**	**ICC**	0.9981	0.8434	0.8443	0.8446
		MDC (º)	0.1989	1.5525	1.4424	1.3875
		SEM (º)	0.0718	0.5601	0.5204	0.5006
	Lowess	**ICC**	0.9985	0.8350	0.8236	0.8093
		MDC (º)	0.2632	2.4918	3.3124	3.9571
		SEM (º)	0.0950	0.8990	1.1950	1.4276

(P55, polynomial 55 function; ICC, intra-class correlation coefficient; MDC, minimal detectable change; SEM, standard error of measurement; º, degrees; m, metres)

4.2.5 Trunk Shape - summary of results

Mathematical fitting function

Extension, flexion and neutral trials

Polynomial 44, 55 and Lowess functions had similar r^2 and RMSE values.

Polynomial 22 and 33 functions had more variable r^2 and RMSE results.

No differences were seen for angles calculated between the polynomial 44, 55 and Lowess functions.

Scoliosis trials

Polynomial 44, 55 and Lowess functions had similar r^2 and RMSE results.

No differences were seen between the computed angles of polynomial 55 and Lowess fits.

Differences were observed between some angles computed for the polynomial 44 and 55 functions.

Tangent length

Across all trials, the variability in angles calculated at different tangent lengths was low.

There was a general pattern of decrease in angles calculated in the scoliosis trials as the tangent length increased.

Validity

Extension, flexion and neutral trials

Angles calculated using the polynomial 55 and Lowess function were $<3^{o}$ different from angles calculated with a flexicurve.

Angles calculated with the polynomial 55 function were commonly the closest to zero (known truth).

Scoliosis trials

The largest difference seen between the flexicurve and EM system was 4.8º for the Lowess function and 7.4º for the polynomial 55 function.

Reliability

Extension, flexion and neutral trials

ICCs were excellent for the Lowess function and poor for the polynomial 55 function.

The Lowess function had smaller MDC and SEM results compared to the polynomial 55 function.

Scoliosis trials

ICCs were good-to-excellent for the polynomial 55 and Lowess functions.

SEM results were similar, but MDC results were better for the polynomial 55 function.

As the tangent length increased, reliability results worsened for both functions.

4.2.6 Trunk Shape - summary of recommendations and conclusions

Mathematical fitting function

Polynomial 22 and 33 functions containing r^2 values <0.97 and RMSE values >0.005 should not be used for trunk shape measurements.

The polynomial 44 function should also not be used as differences were seen between angles calculated with this and the polynomial 55 function in the scoliosis trials.

Tangent length

It is recommended that tangent lengths should not be used interchangeably but if used consistently, the tangent length chosen should not significantly affect trunk shape measurements with an EM system.

Validity

The EM method is valid for measurements of trunk shape and altered trunk shape.

The polynomial 55 function may have higher validity for 'flat hump' surfaces.

For altered trunk surfaces, higher validity was seen for the Lowess function.

Reliability

The Lowess function had better reliability results than the polynomial 55 function for measurements of 'flat hump' surfaces.

Both functions were reliable for measurements of an altered trunk surface and had low SEM.

Overall

It is recommended that the Lowess function should be used for trunk shape measurements with an EM system because:

It yielded high r^2 and RMSE results across all trials,

It didn't calculate different angles from the polynomial 55 function,

It had higher validity for measurements of an altered trunk surface,

It yielded good-to-excellent ICC results across all trials,

It had the best MDC and SEM in measurements of ‘flat hump’ surfaces and

It had low SEM in measurements of an altered trunk surface.

Chapter 5. Discussion

This thesis aimed to build upon previous research by Singh et al. (2010) and Gonzalez-Sanchez et al. (2014) to develop a method for measuring spinal and trunk shape utilising an EM device. This was achieved through the development of a data capture and processing method that extended spinal shape analysis to also include quantification of trunk shape. To our knowledge, this study is the first to measure spinal and trunk shape using an EM system.

Data collected on extension, flexion and neutral phantom models provided proof of concept of this method for measuring both spinal and trunk shape. Furthermore, the capacity of this method to detect alteration in trunk shape was determined through measurements of a pseudo-scoliosis phantom. ICC statistics from trial repeats facilitated measurement of this method's reliability, whilst validity was established through comparison of angles calculated with the flexicurve. In addition, the effect of different tangent lengths, mathematical functions and measurement techniques on results were examined. These findings, combined with the reliability and validity outcomes, enabled specific conclusions and optimal recommendations concerning the method's use in measurements of spinal and trunk shape to be made.

5.1 Reliability of spinal shape measurements

Reliability results of spinal curvature are compared to those reported in earlier studies (Table 32). In this thesis, spinal shape measurements using an EM method exhibited excellent repeatability (ICC = >0.999). This degree of reliability is consistent with previous research that used a similar methodology (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). In these studies, excellent (ICC = 0.93, 0.98) (Singh et al. 2010) and high (Cronbach's alpha = 0.82-0.92) intra-rater reliability within-day was demonstrated in 'healthy' (Singh et al. 2010), normal weight and obese participants (Gonzalez-Sanchez et al. 2014). However, Singh et al. (2010) and Gonzalez-Sanchez et al. (2014) only investigated standing, whereas this study explored extension, flexion and neutral shapes. Therefore, the reliability of the EM method for spinal curvature measurements can be extended to different shapes based on this research.

Table 32. Sagittal spinal curvature reliability results of the present study compared with those in the literature

Study	Method	Reliability investigated	Sagittal reliability results	SEM
Present	EM	Intra-rater	ICC = 0.9997-0.9998	0.004-0.006º
Hart and Rose (1986)	Flexicurve	Test-retest	ICC = 0.97	Not reported
Lovell et al. (1989)	Flexicurve	Intra-rater Inter-rater	ICC = 0.73-0.94 ICC = 0.41-0.54	Not reported
Youdas et al. (1995)	Flexicurve	Intra-rater Inter-rater	ICC = 0.82-0.98 ICC = 0.84-0.98	Not reported
Hinman (2004b)	Flexicurve	Inter-rater	ICC = 0.60-0.94	Not reported
Mannion et al. (2004)	Spinal mouse	Intra-rater Inter-rater	ICC = 0.67-0.92 ICC = 0.64-0.93	2.4-6.2º 2.2-7.0º
Dunleavy et al. (2010)	Flexicurve	Intra-rater Inter-rater	ICC = 0.61-0.97 ICC = 0.56-0.72	0.35-2.0º
Lewis and Valentine (2010)	Inclinometer	Intra-rater	ICC = 0.97	1º, 1.7º
Singh et al. (2010)	EM	Intra-rater	ICC = 0.93,0.98	1.51º, 1.57º
Williams et al. (2010)	Fibre-optic	Repeated measures	CMC = 0.97-0.98	Not reported
Czaprowski et al. (2012)	Inclinometer	Intra-rater	Cronbach's alpha = 0.83 and 0.87	Not reported
de Oliveira et al. (2012)	Flexicurve	Intra-rater Inter-rater	ICC = 0.83, 0.78 ICC = 0.83, 0.94	Not reported
Williams et al. (2012)	Fibre-optic	Repeated measures	CMC = >0.81, ICC = >0.99	Not reported
Gonzalez-Sanchez et al. (2014)	EM	Intra-rater	Cronbach's alpha = 0.90-0.92	Not reported
MacIntyre et al. (2014)	Inclinometer	Test-retest	ICC = 0.90-0.91	2.5-3.5º
Topiladou et al. (2014)	Spinal mouse	Intra-rater	ICC = 0.88-0.99	0.32-2.74º
Sedrez et al. (2016)	Flexicurve	Test-retest Intra-rater Inter-rater	ICC = 0.80, 0.93 ICC = 0.82, 0.67 ICC = 0.83, 0.72	2.5º, 4.3º 4.1º, 5.7º 4.1º, 5.7º
Was et al. (2016)	Inclinometer	Test-retest	ICC = 0.70-0.90	Not reported
Quek et al. (2016)	Kinect (3-D depth camera)	Intra-rater	ICC = 0.96, 0.81	0.69º, 1.07º
Roghani et al. (2017)	Spinal mouse	Intra-rater	ICC = 0.89-0.97	1.41-1.75º

(EM, electromagnetic; SEM, standard error of measurement; ICC, intra-class correlation coefficient; CMC coefficient of multiple correlation; º, degrees; 3-D, 3 dimensional)

In comparison to previous studies with an EM device, other methods have investigated the reliability of spinal curvature measurements in positions different from standing. Similar to this study, Roghani et al. (2017) combined ICC values for sagittal neutral standing, flexion and extension postures, reporting good-to-excellent intra-rater repeatability for the spinal mouse method (ICC = 0.89-0.97) (Roghani et al. 2017). These ICC values are lower than those reported in this thesis which may be explained by the difficulty of ensuring participants adopt identical postures for each trial repeat. As Roghani et al. (2017) measured human populations and between-day, inconsistencies in the actual degree of curvature assessed may have arisen on the separate occasions. This would not have been a problem in the present study as the shape of the phantom models remained consistent, owing to the higher reliability seen.

It appears that repeatability of spinal shape is dependent on multiple factors including method, measurement protocol, spinal region and analysis conducted. Across the literature, wide variations are evident in reliability estimates of sagittal spinal curvature measurements with surface methods (Table 32). Previously, the best ICC values (>0.99) were reported by Williams et al. (2012) for a fibre-optic device. However, this study only investigated the lumbar spine. Conversely, excellent ICC values were shown in this research across thoracic and lumbar curvature measurements. These compare favourably with those reported in earlier works. One possible explanation for this is the absence of human variability or human error. Reliability is often built around the method of application which was explored in this thesis by using phantom models instead of human participants. Accordingly, the natural variability associated with research in-vivo was removed. This includes subtle postural shifts, issues with skin movement and interactions between investigator and participant. Elimination of these potential sources of error is likely to have enhanced this study's reliability to the extent that spinal shape results were statistically excellent.

SEM for spinal curvature with the EM method was shown to be 0.004 to 0.006º in the present study. This is smaller than the SEM reported in earlier research using a similar method (0.56º to 1.6º) (Singh et al. 2010). It is also considerably lower than intra-rater SEMs reported for the flexicurve (0.35º to 2.0º) (Dunleavy et al. 2010) (4.1º to 5.7º) (Sedrez et al. 2016), spinal mouse (2.4º to 6.2º) (Mannion et al. 2004) (0.32º to 2.74º) (Topiladou et al. 2014) (1.41º to 1.75º) (Roghani et al. 2017) and inclinometer (1º to

1.7º) (Lewis and Valentine 2010) (2.5º to 3.5º) (MacIntryre et al. 2014) (Table 32). Explanation could further be attributed to the absence of humans since, in this research, the method of studying phantoms means error contributions are limited to the measurement system and/ or operator only. Even so, the SEM results are extremely small (<0.007º) indicating a high precision of spinal curvature measurements with an EM device and thus, high absolute reliability of this method. This therefore seems to suggest that the error of variance due to operator and system were extremely low.

To estimate the value beyond which true change has occurred, the MDC was also presented in this study. This metric is a measure of how sensitive a device is to detecting change. The MDC values reported in this thesis for spinal curvature are very small (0.02 to 0.14º) (Table 19), further suggesting that the error induced by system and/or operator are minimal. Despite this, it is not clear if such small errors would occur with other operators because between operator comparisons were beyond the scope of this thesis. Furthermore, inter-rater reliability using an EM method is yet to be explored in the literature. Nevertheless, the skill of tracing spinal shape is not particularly onerous. It is also similar to that of the spinal mouse method where research has shown an increase of only 0.8º for the largest SEM reported between intra- and inter- rater investigations (Mannion et al. 2004). Based on these results, increased error of 0.8º would still keep the SEM of the EM method below that of other surface curvature devices. Nevertheless, research is warranted to determine whether low error and consistent measures of spinal curvature would result between operators with this method.

Overall, this study has shown spinal shape measurements with an EM system have excellent reliability and low SEM for extension, flexion and neutral spine shapes. However, the absence of human error may explain why the reliability is statistically excellent, higher than previous studies using a similar methodology and the highest reported in the literature for surface curvature methods. Future studies should investigate measurements between different operators to further explore sources of error thus, improving the understanding of error contributions with this EM method.

5.2 Critique of existing surface methods

Whilst many devices are available for detecting spinal shape or curvature, it is apparent that, prior to this study, not one method has been extended to include measurement of trunk shape or assessment of altered trunk shape. This lack of enhanced application may be explained by each method's inherent limitations. Given that both the flexicurve and inclinometer are limited to the points directly measured by the instrument, many measurements would be required to measure the whole trunk. This would be extremely time consuming. Equally, the fibre-optic method and Kinect method remain in their infancy. Hence, without significant computing input, computation remains non-automated and experimental. Research has highlighted the spinal mouse as promising. However, it is constrained by the operating software to work only as specifically designed. Consequently, measuring regions not directly representing spinal shape (i.e. spinous processes) are currently beyond the measurement capacity of this device.

In view of all these limitations, the EM method provided the greatest opportunity to extend spinal shape measurements to include trunk shape. As a result, in the initial phases of this research, a novel method to quantify trunk shape utilising an EM system was developed. Subsequently, the reliability of this method to measure flat and altered trunk surfaces were investigated. This overcame the insufficiencies identified in the literature.

5.3 Reliability of trunk shape measurements

This novel method was shown to have excellent reliability (ICC = 0.994-0.996) (Table 30) for trunk shape measurements of extension, flexion and neutral models. However, good-to-excellent results (ICC = 0.809-0.999) (Table 31) were seen for measurements of a pseudo-scoliosis phantom. With such promising results, future research should apply this method in humans as well as, populations with known deformity to further determine reliability of trunk shape measurements using this EM method.

Importantly, the challenge in evaluating trunk surface asymmetry comes from the precision and accuracy required by a method to quantify small but significant changes to trunk shape (Pazos et al. 2007). In this study, low SEM results were found for trunk shape measurements across the flexion, extension, neutral (0.06º to 0.07º) and scoliosis (0.09º to 1.4º) trials. This demonstrates high absolute reliability of the method

to measure different and altered trunk surfaces. Similarly, the MDC results are low (0.18 to 4.0º), highlighting the sensitivity of this method to detect changes in trunk shape.

Sensitivity of an instrument is important for the screening and monitoring of conditions, like scoliosis, where changes in spinal and trunk shape occur gradually over-time as deformity develops (Hawes and O'Brien 2006; Janicki and Alman 2007; Latalski et al. 2017; Dunn et al. 2018). It is also a significant factor associated with examining the effectiveness of treatments that provide progressive correction to such deformities (Marquez et al. 2012). Examples include surgical interventions, physiotherapy and orthotics (Leblanc et al. 1998; Pazos et al. 2005; Zabjek et al. 2005). Following surgery Liang et al. (2012) found a mean scoliotic curvature change of 13.5° at follow up whilst other studies reported mean changes of 10.9° (Pellios et al. 2016) and 13.41° (Aulisa et al. 2017) after years of brace application. This demonstrates that the proposed method is sensitive to investigate these spinal shape changes over-time but future research should also explore the method's use in monitoring trunk shape changes over-time.

To summarise, this research resolved challenges identified in previous literature by extending a surface curvature measurement method beyond spinal shape to also include quantification of trunk shape. Measurements of trunk shape with this novel method yielded excellent reliability for extension, flexion and neutral models, whilst good-to-excellent outcomes were found for measurements of a pseudo-scoliosis phantom. Moreover, SEM and MDC results highlighted the sensitivity of the EM method to measure, and detect changes to, trunk shape. This is important for the application of such devices in the screening and monitoring of disease and treatment outcomes. Future studies should now be employed to measure trunk shape in humans and individuals with known deformity. Additionally, research should investigate the method's use in monitoring trunk shape changes over-time resulting from either natural disease progression or treatment intervention.

5.4 Data capture technique

The mean of three repeats is often used in healthcare assessments because one measurement rarely provides accurate enough data sufficient for its application. In this

study, three data trials were conducted for both spinal and trunk shape measurements of the extension, flexion and neutral models using either a ‘consecutive’ or ‘non-consecutive’ data capture technique. It was hypothesized that the consecutive data trials would yield better reliability due to the nature of this technique and thus, the elimination of sources of error. However, the CoV results (Table 18) suggest otherwise.

Although higher CoV values were seen in the spinal shape non-consecutive trials, most trials had a CoV <1%. This suggests that the consecutive technique does not produce significantly more reliable results than the non-consecutive technique. In clinical assessments, it is unlikely that an individual would have to maintain one specific position for the prolonged period of time required to complete three data trials, particularly in the case of trunk shape tracing. Therefore, the non-consecutive technique is more reflective of real-world applications.

In summary, the consecutive data capture technique does not produce significantly more reliable results than the non-consecutive technique with this EM method. As the non-consecutive approach reflects clinical applications better, error should remain low, yielding reliable results for spinal and trunk shape.

5.5 Validity

Reliability is not the only important component when deciding whether a method can be implemented for use in a clinical environment. Confirming that the device yields valid results is also essential. This is of particular significance when the aim of such a method is to assess the extent of disease, deformity or injury as well as, examine the effectiveness of interventions.

Although the reliability of spinal shape measurements with an EM system has previously been established (Singh et al. 2010; Gonzalez-Sanchez et al. 2014), this study is the first to show the method to be valid. For this research, the concern was whether or not the angles computed with the EM system accurately represented the actual tangent or resultant angle of the anatomical region being investigated.

Typically, an established gold-standard comparator is used in order to make such an assessment of accuracy. However, this is difficult to do for spinal and trunk posture since this remains the domain of radiography where accessibility, cost and radiation exposure are significant restrictions. Nevertheless, previous research has validated the flexicurve technique against x-ray analysis of spinal curvature (Teixeira et al. 2007; de Oliveira et al. 2012; Grindle et al. 2020). Therefore, the flexicurve was identified as a suitable alternative comparator for assessment of validity.

Spinal shape measurements with the EM system were shown to be valid (≤6.0º different to flexicurve method). However, there may be several explanations for the differences in angles calculated between the two methods. Firstly, a protractor does not measure to any decimal places unlike the EM system. Secondly, multiple stages of the flexicurve process are associated with measurement error. These errors are widely documented in the literature and include loss of format of the flexicurve, its alignment on paper, differing thickness of drawn lines and the marking of co-ordinates (de Oliveira et al. 2012).

For trunk shape measurements of the extension, flexion and neutral shapes, differences between the flexicurve angles and EM system were ≤3.0º. However, it is unclear why each method failed to consistently yield angles more similar to 0º considering the 'hump' surface of these models remained flat in the transverse plane. Therefore, neither the flexicurve nor EM system may be particularly suited methods for measuring flat trunk surfaces. Given that the EM system 'models' surfaces as a series of curves, measurements of flat surfaces may be particularly challenging with this method. Despite this, the surface of an individual's trunk is rarely completely flat thus, this is unlikely to be a problem in a clinical environment.

Additionally, flat surfaces are not the best way of determining the validity of the EM system to measure trunk shape. Therefore, comparisons of computed angles measured from the pseudo-scoliosis phantom with the flexicurve 'truth' may be more useful. Between the two methods, calculated angles were no more than 4.6º different, demonstrating validity of this method in measurements of altered trunk surfaces. Provided the flexicurve angles are a good representation of the truth, this puts the error of the EM system below that previously accepted for radiographic evaluation (5º) (Morrissy et al. 1990).

Overall, the spinal and trunk shape angles calculated with the EM system were shown to be valid against corresponding angles calculated with the flexicurve. However, caution is advised in using the flexicurve as a gold-standard comparator due to its measurement error and, possibly, its limited validity in measuring flat surfaces. Nevertheless, lower validity is likely to be seen for measurements of 'flat hump' surfaces with the EM method, but this shouldn't to be a concern for clinical applications.

5.6 Study methodology

It became apparent that several assumptions concerning the data processing of spinal shape measurements with an EM device had been made in previous studies (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). This included the 5th order polynomial and 0.05m tangent length being the 'optimal'. However, during this thesis' development of methods (section 2.3.1), the 5th order polynomial was seen to be a poor representation of data when the spinal shape was unusual (condition 3 (Figure 12)). Furthermore, depending on the area of interest and where the tangent spanned, a loss of shape could result when using larger tangent lengths.

Robust methods offer the opportunity for increased quality and optimised operational performance which is an important consideration in the development of healthcare devices (Clarkson et al. 2018). As a result, instead of maintaining these assumptions, this study employed a more systematic approach to optimisation. This guided the exploration of the 'optimal' mathematical fit and tangent length for spinal and trunk shape measurements respectively.

The 6th order polynomial was identified as the best mathematical fit for spinal shape measurements. This is different to the 5th order polynomial used in previous research (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). Although the inferences seen in this literature remain acceptable and were demonstrated as appropriate for this previous research, this study has been able to quantify that a 6th order polynomial may be more optimal for the range of spinal shapes likely to be experienced in the clinic. It was concluded that Lowess was the best surface fitting function for trunk shape measurements.

Concerning tangent length, the conclusion of this study was that even though the tangent length should not significantly affect the spinal and trunk shape angles calculated, they should not be used interchangeably. Differences were evident across the tangent lengths investigated but the pattern of change depended on the variable analysed. This demonstrates that tangent length may be influenced by the anatomical region being investigated. Cases of a surface device 'house' with larger dimensions results in essentially a 'large' tangent length for orientation calculation. For example, a wireless EM device for surface mounting (Liberty Latus) uses sensors 8.9cm x 4.2cm x 2.5cm (Polhemus 2020b). This can result in a tangent length of 9cm (depending on mounting) compared to another sensor (XSens DOT) with dimensions of 3.6cm x 3.0cm x 1.1cm (Xsens 2020). The findings of this thesis therefore suggest that these devices should not be used interchangeably.

In summary, compared to previous research of surface curvature measurements with an EM system, this study demonstrated a systematic approach to the methodology enabling informed recommendations to be made concerning 'optimal' data processing techniques. The 6th order polynomial and Lowess functions were found to be the optimal mathematical functions for spinal shape and trunk shape respectively. Conclusions as to the optimal tangent length were unable to be reached for spinal and trunk shape. However, recommendations concerning the use of tangent lengths consistently, not interchangeably, were made. Furthermore, specific applications of this method should be considered when deciding the optimal tangent length.

5.7 Clinical applications

The measurement of spinal and trunk posture are common features evident in the clinical assessment of many conditions (Fortin et al. 2011). As a result, measurement techniques for spinal and trunk shape are important tools required for clinicians. Yet, commonly these examinations involve qualitative assessment such as observation or clinical outcome measures (Fedorak et al. 2003; Pazos et al. 2007), where a flaw of such approaches lies within its subjectivity (Fortin et al. 2011). Alternatively, for quantitative measures, reliance remains on radiographic evaluation as the gold-standard (Pazos et al. 2007). However, surface curvature methods have been developed to overcome the limitation of repeated radiation exposure associated with this gold-standard. In spite of this, none of these surface methods have been extended

beyond spinal shape measurements. For these reasons, it is well recognised that there is clinical need for a new measurement method for evaluating both spinal and trunk posture.

This thesis has presented a new and innovative method of measuring spinal and trunk shape with an EM system. Through research on phantom models, this study has demonstrated that it is possible to measure and detect changes to both spinal and trunk shape objectively without involving repeated exposure to radiation or invasive methods. Since the data in this study was found to be reliable and valid, this method has the potential for routine clinical practice.

This holds exciting prospects for healthcare. For instance, spinal shape forms a fundamental part of the clinical assessment of conditions like osteoporosis. Therefore, the method could become a standard measure for such a disease, especially in cases involving curvature change but in the absence of trauma or suspected fracture. Similarly, going beyond the limits of spinal shape means the method, as developed in this thesis, can also provide information pertaining to trunk shape. This could be applied clinically in the study of pregnancy related postural changes whereby, use of this method could permit opportunities to investigate these over-time. Therefore, complimenting existing research in this area.

Combined evaluation of spine and trunk shape also holds significant potential in scoliosis where postural change is a consequence (Heitz et al. 2018) and closely correlated with spinal curve progression (Dalleau et al. 2012). Thus, this method could serve as a mass screening tool for scoliosis in adolescents, resulting in early, non-invasive identification. Equally, the method could be appropriate in neurological populations where pathologies are often linked with postural impairments (Genthon et al. 2007; Fortin et al. 2011; Trompetto et al. 2014). This highlights that the application of this method is not restricted to musculoskeletal disorders. To this end, the method has potential to assist in the identification of conditions, improve monitoring of such conditions and evaluate the effect of treatment interventions. This could lead to increased understanding of the implications to posture, kinematics and functioning in the presence of disease.

5.8 Method Evaluation

This EM method is clinically attractive because it is non-invasive, non-ionising and of a relatively low cost (£2000-£6000). Nevertheless, disadvantages of an EM system are documented in the literature. This includes optimum operational zones which may affect the accuracy of measurements depending on the distance between the source and sensor, constraining data collection (Milne et al. 1996; Bull and Amis 1997; Schuler et al. 2005; Wong and Wong 2008; Franz et al. 2014). Measurements could also be adversely affected by the presence of metallic objects (Milne et al. 1996; Burnett et al. 2004; Ng et al. 2009; Franz et al. 2014). Consequently, this method may not be suitable for some patients such as those with metallic implants or prostheses (Wong and Wong 2008). To this extent, case by case considerations are important but rarely is there a clinical measurement that suits all individuals.

Besides this, the method does not allow for dynamic motion capture and although the tracing of spinal shape is quick and simple, trunk shape tracing might be slightly more time consuming. In addition, the method does involve the use of mathematical equations that may not be understood by clinicians and currently, there is no software available with computations 'built in'. Even so, understanding of these calculations is not required in order to perform spinal and trunk shape measurements. Furthermore, specific algorithms or programs could be written to perform all the data processing tasks, producing results almost immediately and making the method easier for use in clinical contexts.

In summary, this thesis has addressed a clinical need by developing a method using an EM system to measure trunk shape in addition to spinal shape. This objective quantification of spinal and trunk shape in clinical applications could benefit the examination and understanding of many diseases. Meanwhile the non-invasive, non-ionising, simple and low-cost features of this method make it clinically attractive. However, limitations to this novel method are evident but solutions to minimise these exist, meaning the method has potential for future routine clinical use.

Chapter 6. Limitations and conclusions

6.1 Study limitations

A limitation of the present study is that it did not investigate human participants. Therefore, it is unknown whether humans can remain stationary or tolerate specific positions for the time is takes to complete data capture with this method. Thus, future work investigating these considerations is warranted

Furthermore, variability in humans is a common phenomenon, and source of error, that can directly impact reliability of measurement tools. Therefore, extrapolation of the findings is limited. Reliability of this method in human populations for spinal shape measurements have been shown (Singh et al. 2010; Gonzalez-Sanchez et al. 2014). However, similar research applying this novel method in-vivo to measure spinal and trunk shape is further required. In addition, inferences regarding the method's use in intended populations for its clinical application cannot be made without research into those with injury or disease and over-time.

Another limitation is that this study failed to investigate elements of data capture that may impact results. For example, the sampling time could alter the accuracy of angles calculated with the EM system. Consequently, additional research may be needed to investigate such influences. This could provide further recommendations for optimisation of this method, increasing its robustness.

6.2 Conclusion

This thesis delivers a valid and reliable novel method for measuring spinal and trunk shape utilising an EM system. This work has built upon earlier research, overcoming the insufficiencies of previous literature and surface measurement methods to address a clinical need.

The results illustrate proof of concept of this method in determining spinal and trunk shape which are integral components in the clinical assessment of many disorders. For spinal shape measurements, reliability results are excellent and compare favourably with previous studies. Trunk shape reliability with the method is also excellent and all measurements with the EM system have high absolute reliability, showing its capacity to detect alteration in spinal and trunk shape. Furthermore, computed angles are similar to corresponding angles obtained with the flexicurve, demonstrating this method's validity.

The results show the 6^{th} order polynomial equation is optimal for spinal shape measurements with this EM method. Comparatively, the Lowess function is recommended for trunk shape analysis. Conclusions as to the optimal tangent length were unable to be reached but are likely influenced by the variable and region being investigated. Thus, specific applications of the method should be considered when deciding which tangent length to use. Additionally, the data capture technique used for the three measurement repeats should not affect reliability results.

6.3 Recommendations for future work

- Research trunk shape measurements in-vivo with this method.
- Determine the effect of human error and variability on the reliability and validity of trunk shape measurements with this method.
- Explore between operator reliability of spinal and trunk shape measurements with this method.
- Explore within-day and between-day reliability of spinal and trunk shape measurements with this method.
- Investigate different data capture parameters, including time, that may alter results and make further recommendations to this method based on these findings.
- Measure spinal and trunk shape with this method in individuals with known deformity or disease.
- Investigate this method's usefulness in monitoring changes to spinal and trunk shape over time due to natural disease progression or treatment intervention.

References

Adams, M. A., Dolan, P., Marx, C. and Hutton, W. C., 1986. An electronic inclinometer technique for measuring lumbar curvature. *Clinical biomechanics* [online], 1 (3), 130-134.

Ahmed, A. S., Ramakrishnan, R., Ramachandran, V., Ramachandran, S. S., Phan, K. and Antonsen, E. L., 2018. Ultrasound diagnosis and therapeutic intervention in the spine. *Journal of spine surgery* [online], 4 (2), 423-432.

Arksey, H., and O'Malley, L., 2005. Scoping studies: towards a methodological framework. *International journal of social research methodology* [online], 8 (1), 19-32.

Aulisa, A. G., Guzzanti, V., Falciglia, F., Galli, M., Pizzetti, P. and Aulisa, L., 2017. Curve progression after long-term brace treatment in adolescent idiopathic scoliosis: comparative results between over and under 30 cobb degrees – SOSORT 2017 award winner. *Scoliosis and spinal disorders* [online], 12, 36.

Barrett, E., Lenehan, B., O'sullivan, K., Lewis, J. and McCreesh, K., 2018. Validation of the manual inclinometer and flexicurve for the measurement of thoracic kyphosis. *Physiotherapy theory and practice* [online], 34 (4), 301-308.

Breen, A. C., Muggleton, J. M. and Mellor, F. E., 2006. An objective spinal motion imaging assessment (OSMIA): reliability, accuracy and exposure data. *BMC musculoskeletal disorders* [online]. Available from: https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-7-1 [Accessed 14 October 2019].

Breen, A. C., Teyhen, D. S., Mellor, F. E., Breen, A. C., Wong, K. W. N. and Deitz, A., 2012. Measurement of intervertebral motion using quantitative fluoroscopy: Report of an international forum and proposal for use in the assessment of degenerative disc disease in the lumbar spine. *Advances in Orthopaedics* [online]. Available from: https://www.hindawi.com/journals/aorth/2012/802350/#B22 [Accessed 03 February 2020].

Breen, A., Hemming, R., Mellor, F. and Breen, A., 2018. Intrasubject repeatability of in vivo intervertebral motion parameters using quantitative fluoroscopy. *European spine journal* [online], 28 (2), 450-460.

Briggs, A. M., Wrigley, T. V., Tully, E. A., Adams, P.E, Greig, A. M. and Bennell, K. L., 2007. Radiographic measures of thoracic kyphosis in osteoporosis: cobb and vertebral centroid angles. *Skeletal radiology* [online], 36 (8), 761-767.

Buchbinder, R., Blyth, F. M., March, L. M., Brooks, P., Woolf, A. D. and Hoy, D. G., 2013. Placing the global burden of low back pain in context. *Best practice and research clinical rheumatology* [online], 27 (5), 575-589.

Bull, A. M. J. and Amis, A. A., 1997. Accuracy of an electromagnetic tracking device. *Journal of biomechanics* [online], 30 (8), 857-858.

Burnett, A. F., Cornelius, M. W., Dankaerts, W. and O'Sullivan, P. B., 2004. Spinal kinematics and trunk muscle activity in cyclists: a comparison between healthy controls and non-specific chronic low back pain subjects- a pilot investigation. *Manual therapy* [online], 9 (4), 211- 19.

Burton, K., 1986. Regional lumbar sagittal mobility; measurement by flexicurves. *Clinical biomechanics* [online], 1 (1), 20-26.

Cakir, B., Richter, M., Käfer, W., Wieser, M., Puhl, W. and Schmidt, R., 2006a. Evaluation of lumbar spine motion with dynamic x-ray. *Spine* [online], 31 (11), 1258-1264.

Cakir, B., Richter, M., Puhl, W. and Schmidt, R., 2006b. Reliability of motion measurements after total disc replacement: the spike and the fin method. *European spine journal* [online], 15 (2), 165-173.

Castro, A. P. G., Pacheco, J. D., Lourenco, C., Queiros, S., Moreira, A. H. J., Rodrigues, N. F. and Vilaca, J. L., 2017. Evaluation of spinal posture using Microsoft Kinect™: a preliminary case-study with 98 volunteers. *Porto biomedical journal* [online], 2 (1), 18-22.

Chandel, V. and Ghose, A., 2018. Demo abstract: EMeasure: using a smart device with consumer-grade accelerometer as an accurate measuring scale [online]. In: 17th ACM/IEEE international conference on information processing in sensor networks (IPSN), Porto 11-13 April 2018. Porto: Institute of Electrical and Electronics Engineers. Available from: https://ieeexplore.ieee.org/abstract/document/8480043 [Accessed 25 November 2020].

Chleboun, G. S., Amway, M. J., Hill, J. G., Root, K. J., Murray, H. C. and Sergeev, A. V., 2012. Measurement of segmental lumbar spine flexion and extension using ultrasound imaging. *Journal of orthopaedic and sports physical therapy* [online], 42 (10), 880-885.

Christe, G., Kade, F., Jolles, B. M. and Favre, J., 2017. Chronic low back pain patients walk with locally altered spinal kinematics. *Journal of biomechanics* [online], 60, 211-218.

Chun, S-W., Lim, C-Y., Kim, K., Hwang, J. and Chung, S. G., 2017. The relationships between low back pain and lumbar lordosis: a systematic review and meta-analysis. *The spine journal* [online], 17 (8), 1180-1191.

Clarkson, J., Dean, J., Ward, J., Komashie, A. and Bashford, T., 2018. A systems approach to healthcare: from thinking to practice. *Future healthcare journal* [online], 5 (3), 151-155.

Cloud, B. A., Zhao, K. D., Breighner, R., Giambini, H. and An, K-N., 2014. Agreement between fiber optic and optoelectronic systems for quantifying sagittal plane spinal curvature in sitting. *Gait posture* [online], 40 (3), 369-374.

Cuesta-Vargas, A. I., 2015. Development of a new ultrasound-based system for tracking motion of the human lumbar spine: reliability, stability and repeatability during forward bending movement trials. *Ultrasound in medicine and biology* [online], In Press. Available from: https://www.umbjournal.org/article/S0301-5629(15)00209-4/fulltext [Accessed 12 October 2019].

Cushman, D., Flis, A., Jensen, B., McCormick, Z., 2016. The effect of body mass index on fluoroscopic time and radiation dose during sacroiliac joint injections. *Physical medicine and rehabilitation* [online], 8 (8), 767-772.

Czaprowski, D., Pawlowska, P., Gebicka, A., Sitarski, D. and Kotwicki, T., 2012. Inta- and interobserver repeatability of the assessment of anterioposterior curvatures of the spine using saunders digital inclinometer. *Ortopedia traumatologia rehabilitacja* [online], 2 (6), 145-153.

Dalleau, G., Leroyer, P., Beaulieu, M., Verkindt, C., Rivard, C-H. and Allard, P., 2012. Pelvis morphology, trunk posture and standing imbalance and their relations to the cobb angle in moderate and severe untreated AIS. *PLoS one* [online], 7 (7).

Davey, E. and England, A., 2014. AP versus PA positioning in lumbar spine computed radiography: image quality and individual organ doses. *Radiography* [online], 21 (2), 188-196.

Davis, R. J., Lee, D. C., Wade, C. and Cheng, B., 2015. Measurement performance of a computer assisted vertebral motion analysis system. *International journal of spine surgery* [online], 9 (36).

de Oliveira, T. S, Candotti, C. T., La Torre, M., Pelinson, P. P. T., Furlanetto, T. S., Kutchak, F. M. and Loss, J. F., 2012. Validity and reproductibility of the measurements obtained using the flexicurve instrument to evaluate the angles of thoracic and lumbar curvatures of the spine in the sagittal plane. *Rehabilitation research and practice* [online]. Available from: https://www.hindawi.com/journals/rerp/2012/186156/ [Accessed 7 November 2020].

Dombrowski, M. E., Rynearson, B., LeVasseur, C., Adgate, Z., Donaldson, W. F., Lee, J. Y., Aiyangar, A. and Anderst, W. J., 2018. ISSLS prize in bioengineering science 2018: dynamic imaging of degenerative spondylolisthesis reveals mid-range dynamic lumbar instability not evident on static clinical radiographs. *European spine journal* [online], 27 (4), 752-762.

du Rose, A. and Breen, A., 2016. Relationships between lumbar inter-vertebral motion and lordosis in healthy adult males: a cross sectional cohort study. *BMC musculoskeletal disorders* [online], 17, 121.

Dunleavy, K., Mariano, H., Wiater, T. and Goldberg, A., 2010. Reliability and minimal detectable change of spinal length and width measurements using the flexicurve for usual standing posture in healthy young adults. *Journal of back and musculoskeletal rehabilitation* [online], 23 (4), 209-214.

Dunn, J., Henrikson, N. B. and Morrison, C. C., 2018. Screening for adolescent idiopathic scoliosis evidence report and systematic review for the US preventative services task force. *JAMA* [online], 319 (2), 173-187.

Durbridge, G., 2011. Magnetic resonance imaging: fundamental safety issues. *Journal of orthopaedic and sports physical therapy* [online], 41 (11), 820-828.

Elliot, B., Hardcastle, P., Burnett, A. and Foster, D., 1992. The influence of fast bowling and physical factors on radiologic features in high performance young fast bowlers. *Sports medicine, training and rehabilitation* [online], 3 (2), 113-130.

Fedorak, C., Ashworth, N., Marshall, J. and Paull, H., 2003. Reliability of the visual assessment of cervical and lumbar lordosis: how good are we? *Spine* [online], 28 (16), 1857-1859.

Fenster, A., Downey, D. B. and Cardinal, H. N., 2001. Three-dimensional ultrasound imaging. *Physics in medicine and biology* [online], 46 (5), R67-R99.

Field, E. and Hutchinson, J. O., ed. 2013. Field's anatomy, palpation & surface markings. 5th edition. Edinburgh: Elsevier.

Fortin, C., Feldman, D. E., Cheriet, F. and Labell, H., 2011. Clinical methods for quantifying body segment posture: a literature review. *Disability and rehabilitation* [online], 33 (5), 367-383.

Franz, A. M., Haidegger, T., Birkfellner, W., Cleary, K., Peters, T. M. and Maier-Hein, L., 2014. Electromagnetic tracking in medicine – A review of technology, validation and applications. *IEEE transactions on medical imaging* [online], 33 (8), 1702-1725.

Fritz, J. M., Erhard, R. E. and Hagen, B. F., 1998. Segmental instability of the lumbar spine. *Physical therapy*, 78 (8), 889 – 896.

Galbusera, F., Niemeyer, F., Tao, Y., Cina, A., Sconfienza, L. M., Kienle, A., Wilke, H-J., 2021. ISSLS prize in bioengineering science 2021: in vivo sagittal motion of the lumbar spine in low back pain patients-a radiological big data study. *European spine journal* [online], In Press. Available from: https://link.springer.com/article/10.1007/s00586-021-06729-z [Accessed 25 January 2021].

Genthon, N., Vuillerme, N., Monnet, J-P., Petit, C. and Rougier, P., 2007. Biomechanical assessment of the sitting posture maintenance in patients with stroke. *Clinical biomechanics* [online], 22 (9), 1024-1029.

Gercek, E., Hartmann, F., Kuhn, S., Degreif, J., Rommens, P. M. and Rudig, L., 2008. Dynamic angular three-dimensional measurement of multisegmental thoracolumbar motion in vivo. *Spine* [online], 33 (21), 2326-2333.

Ghandhari, H., Nabizadeh, N., Nikouei F., Mahabadi, M. A., Mahdavi, S. M., Kamaly, T. and Aghdam, A. A., 2020. Comparison of cobb angles on radiographs with magnetic resonance imaging in idiopathic scoliosis. *Journal of research in orthopedic science* [online], 7 (1), 29-34.

Global burden of disease and injury incidence and prevalence collaborators, 2017. Global, regional and national incidence, prevalence and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the global burden of disease study 2016. *The lancet* [online]. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)32154-2/fulltext#%20 [Accessed 11 September 2019].

Golightly, Y. M., Goode, A. P., Cleveland, R. J., Nelson, A. E., Hannan, M. T., Hillstrom, H. J., Kraus, V. B., Schwartz, T. A., Renner, J. B. and Jordan, J. M., 2016.

Relationship of joint hypermobility with low back pain and lumbar spine osteoarthritis: a cohort study. *Annals of the rheumatic diseases* [online], 75, 659.

Gonzalez-Sanchez, M., Luo, J., Lee, R. and Cuesta-Vargas, A. I., 2014. Spine curvature analysis between participants with obesity and normal weight participants: a biplanar electromagnetic device measurement. *BioMed research international* [online]. Available from: https://www.hindawi.com/journals/bmri/2014/935151/ [Accessed 6 November 2020].

Grindle, D. M., Mousavi, S. J., Allaire, B. T., White, A. P. and Anderson, D. E., 2020. Validity of flexicurve and motion capture for measurements of thoracic kyphosis vs standing radiographic measurements. *JOR spine* [online], Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/jsp2.1120 [Accessed 27 April 2020].

Haas, M., Nyiendo, J., Peterson, C., Thiel, H., Sellers, T., Cassidy, D. and Yong-Hing, K., 1990. Interrater reliability of roentgenological evaluation of the lumbar spine in lateral bending. *Journal of manipulative physiological therapeutics*, 13 (4).

Haldeman, S., Kopansky-Giles, D., Hurwitz, E. L., Hoy, D., Erwin, W. M., Dagenais, S., Kawchuk, G., Stromqvist, B. and Walsh, N., 2012. Advancements in the management of spine disorders. *Best practice & clinical rheumatology* [online], 26 (2), 263-280.

Hart, D. L. and Rose, S. J., 1986. Reliability of a noninvasive method for measuring the lumbar curve*. *The journal of orthopaedic and sports physical therapy* [online], 8 (4), 180-184.

Hartwig, V., Giovannetti, G., Vanello, N., Lombardi, M., Landini, L. and Simi, S., 2009. Biological effects and safety in magnetic resonance imaging: a review. *International journal of environment research and public health* [online], 6 (6), 1778-1798.

Harvey, S., Hukins, D., Smith, F., Wardlaw, D. and Kader, D., 2016. Measurement of lumbar spine intervertebral motion in the sagittal plane using videofluoroscopy. *Journal of back and musculoskeletal rehabilitation* [online], 29 (3), 445-457.

Hawes, M. C. and O'Brien, J. P., 2006. The transformation of spinal curvature into spinal deformity: pathological processes and implications for treatment. *Scoliosis* [online], 1.

Heidari, P., Farahbakhsh, F., Rostami, M., Noormohammadpour, P. and Kordi, R., 2015. The role of ultrasound in diagnosis of the causes of low back pain: a review of the literature. *Asian journal of sports medicine* [online], 6 (1).

Heitz, P-H., Aubin-Fournier, J-F., Parent, E. and Fortin, C., 2018. Test-retest reliability of posture measurements in adolescents with idiopathic scoliosis. *The spine journal* [online], 18 (12), 2247-2258.

Hellström, M., Jacobsson, B., Swärd, L. and Peterson, L., 1990. Radiologic abnormalities of the thoraco-lumbar spine in athletes. *Acta radiologica* [online], 31 (2), 127-132.

Hides, J. A., Richardson, C. A. and Jull, G. A., 1998. Use of real-time ultrasound imaging for feedback in rehabilitation. *Manual therapy* [online], 3 (3), 125-131.

Hinman, M. R., 2004a. Comparison of thoracic kyphosis and postural stiffness in younger and older women. *Spine* [online], 4 (4), 413-417.

Hinman, M. R., 2004b. Interrater reliability of flexicurve postural measures among novice users. *Journal of back and musculoskeletal rehabilitation* [online], 17 (1), 33-36.

Hodges, P., Van den hoorn, W., Dawson, A. and Cholewicki, J., 2009. Changes in the mechanical properties of the trunk in low back pain may be associated with recurrence. *Journal of biomechanics* [online], 42 (1), 61-66.

Hoy, D., March, L., Brooks, P., Woolf, A., Blyth, F., Vos, T. and Buchbinder, R., 2010. Measuring the global burden of low back pain. *Best practice & research: clinical rheumatology* [online], 24 (2), 155-165.

Huang, K-Y., Line, R-M., Lee, Y-L. and Li, J-D., 2009. Factors affecting disability and physical function in degenerative lumbar spondylolisthesis of L4-5: evaluation with axially loaded MRI. *European spine journal* [online], 18 (12), 1851-1857.

Israel, M., 1959. A quantitative method of estimating flexion and extension of the spine; a preliminary report. *Military medicine* [online], 124 (3), 181-186.

Janicki, J. A. and Alman, B., 2007. Scoliosis: review of diagnosis and treatment. *Paediatrics child health* [online], 12 (19), 771-776.

Janssen, M., Nabih, A., Moussa, W., Kawchuk, G. N. and Carey, J. P., 2011. Evaluation of diagnosis techniques used for spinal injury related back pain. *Pain research and treatment* [online]. Available from: https://www.hindawi.com/journals/prt/2011/478798/ [Accessed 27 January 2021].

Jensen, R. K., Jensen, T. S., Koes, B. and Hartvigsen, J., 2020. Prevalance of lumbar spine stenosis in general and clinical populations: a systematic review and meta-analysis. *European spine journal* [online], 29, 2143-2163.

Kalichman, L., Kim, D. H. and Hunter, D. J., 2009. Spondylolysis and spondylolisthesis: prevalence and association with low back pain in the adult community-based population. *Spine* [online], 34 (2), 199-205.

Karayannis, N. V., Smeets, R. J. E. M., Van den hoorn, W. and Hodges, P. W., 2013. Fear of movement is related to trunk stiffness in low back pain. *Public library of science (PLOS) one* [online]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0067779 [Accessed 26 November 2019].

Kim, H. M., Choi, K. H. and Kim, T. W., 2013. Patients' radiation dose during videofluoroscopic swallowing studies according to underlying characteristics. *Dysphagia* [online], 28 (2), 153-158.

Kluszczynski, M., Wasik, J., Ortenburger, D., Zarzycki, D. and Siwik, P., 2017. Prognostic value of measuring the angles of lumbar lordosis and thoracic kyphosis

with the saunders inclinometer in patients with low back pain. *Polish annals of medicine* [online], 24 (1), 31-35.

Konieczy, M. R., Senyurt, H. and Krauspe, R., 2013. Epidemiology of adolescent idiopathic scoliosis. *Journal of children's orthopaedics* [online], 7 (1).

Koo, T. K. and Li, M. Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* [online], 15 (2), 155-163.

Krüger, M., Straube, A., and Eggert, T. 2017. The propagation of movement variability in time: a methodological approach for discrete movements with multiple degrees of freedom. *Frontiers in computational neuroscience* [online], *11* (93).

Kulig, K., Powers, C. M., Landel, R. F., Chen, H., Fredericson, M., Guillet, M. and Butts, K., 2007. Segmental lumbar mobility in individuals with low back pain: in vivo assessment during manual and self-imposed motion using dynamic MRI. *BMC musculoskeletal disorders* [online]. Available from: https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-8-8 [Accessed 26 November 2019].

Kuwahara, W., Deie, M., Fujita, N., Tanaka, N., Nakanishi, K., Sunagawa, T., Asaeda, M., Nakamura, H., Kono, Y. and Ochi, M., 2016. Characteristics of thoracic and lumbar movements during gait in lumbar spinal stenosis patients before and after decompression surgery. *Clinical biomechanics* [online], 40, 45-51.

Lai, Z. H., Sá dos Reis, C. and Sun, Z., 2020. Effective dose and image optimisation of lateral lumbar spine radiography: a phantom study. *European Radiology Experimental* [online], 4. Available from: https://eurradiolexp.springeropen.com/articles/10.1186/s41747-019-0132-3 [Accessed 25 January 2021].

Lam, S. C. B., McCane, B. and Allen, R., 2009. Automated tracking in digitized videofluoroscopy sequences for spine kinematic analysis. *Image and vision computing* [online], 27 (10), 1555-1571.

Landel, R., Kulig, K., Fredericson, M., Li, B. and Powers, C. M., 2008. Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. *Physical therapy* [online], 88 (1), 43-49.

Latalski, M., Danielewicz-Bromberek, A., Fatyga, M., Latalska, M., Krober, M. and Zwolak, P., 2017. Current insights into the aetiology of adolescent idiopathic scoliosis. *Archives if orthopaedic and trauma surgery* [online], 137, 1327-1333.

Leblanc, R., Labelle, H., Forest, F. and Poitras, B., 1998. Morphologic discrimination among health subjects and patients with progressive and nonprogressive adolescent idiopathic scoliosis. *Spine* [online], 23 (10), 1109-1115.

Leone, A., Guglielmi, G., Cassar-Pullicino, V. N., Bonomo, L., 2007. Lumbar intervertebral instability: A review. *Radiology* [online], 245 (1), 62-77.

Lewis, J. S. and Valentine, R. E., 2010. Clinical measurement of the thoracic kyphosis. A study of the intra-rater reliability in subjects with and without shoulder pain. *BMC musculoskeletal disorders* [online], 11. Available from: https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-11-39 [Accessed 7 November 2020].

Liang, C-Z., Li, F-C., Li, H., Tao, Y., Zhou, X. and Chen, Q-X., 2012. Surgery is an effective and reasonable treatment for degenerative scoliosis: a systematic review. *The journal of international medical research* [online], 40, 339-405.

Livanelioglu, A., Kaya, F., Nabiyev, V., Demirkiran, G. and Flrat, T., 2016. The validity and reliability of “spinal mouse” assessment of spinal curvatures in the frontal plane in pediatric adolescent idiopathic thoraco-lumbar curves. *European spine journal* [online], 25, 476-482.

Logan, G. S., Pike, A., Copsey, B., Parfrey, P., Etchegary, H. and Hall. A., 2019. What do we really know about the appropriateness of radiation emitting imaging for low back pain in primary emergency care? A systematic review and meta-analysis of medical record reviews. *Plos one* [online]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0225414 [Accessed 25 January 2021].

Lovell, F. W., Rothstein, J. M. and Personius, W. J., 1989. Reliability of clinical measurements of lumbar lordosis taken with a flexible rule. *Physical therapy* [online], 69 (2).

Luczak, S., Grepl, R. and Bodnicki, M., 2017. Selection of MEMS accelerometers for tilt measurements. *Journal of sensors* [online]. Available from: https://www.hindawi.com/journals/js/2017/9796146/ [Accessed 25 November 2020].

Luinge, H. J. and Veltink, P. H., 2005. Measuring orientation of human body segments using miniature gyroscopes and accelerometers. *Medical and biological engineering and computing* [online], 43, 273-282.

MacIntyre, N. J., Lorbergs, A. L. and Adachi, J. D., 2014. Inclinometer-based measures of standing posture in older adults with low bone mass are reliable and associated with self-reported, but not performance-based, physical function. *Osteoporosis international* [online], 25, 721-728.

Maetzel, A. and Li, L., 2002. The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best practice and research: clinical rheumatology* [online], 16 (1), 23-30.

Mahato, N. K., Montuelle, S. and Clark, B. C., 2019. Assessment of in vivo lumbar inter-veterbral motion: reliability of a novel dynamic weight-bearing magnetic resonance imaging technique using a side bending task. *Asian spine journal* [online], 13 (3), 377-385.

Maigne, J-Y., Lapeyre, E., Morvan, G. and Chatellier, G., 2003. Pain immediately upon sitting down and relieved by standing up is often associated with radiological lumbar instability or marked anterior loss of disc space. *Spine* [online], 28 (12), 1327-1334.

Maniadakis, N. and Gray, A., 2000. The economic burden of back pain in the UK. *Pain* [online], 84(1), 95-103.

Manninen, H., Kiekara, O., Soimakallio, S. and Vainio, J., 1988. Reduction in radiation dose an imaging costs in scoliosis radiography. Application of large-screen intensifier photofluorography. *Spine* [online], 13 (4), 409-412.

Mannion, A. F., Knecht, K., Balaban, G., Dvorak, J. and Grob, D., 2004. A new skin-surface device for measuring the curvature and global and segmental ranges of motion of the spine: reliability of measurement and comparison with data reviewed from the literature. *European spine journal* [online], 13, 122-136.

Margarido, C. B., Arzola, C., Balki, M. and Carvalho, J. C., 2010. Anesthesiologists' learning curves for ultrasound assessment of the lumbar spine. *Canadian journal of anaesthesia* [online], 57 (2), 120-126.

Marquez, J. M. S., Perez-Grueso, F. J. S., Fernadez-Baillo, N. and Garay, E. G., 2012. Gradual scoliosis correction over time with shape-memory metal: a preliminary report of an experimental study. *Scoliosis* [online], 7, 20.

Marshburn, T. H., Hadfield, C. A., Sargsyan, A. E., Garcia, K., Ebert, D. and Dulchavsky, S. A., 2014. New heights in ultrasound: first report of spinal ultrasound from the international space station. *The journal of emergency medicine* [online], 46 (1), 61-70.

Martin, B. I., Deyo, R. A., Mirza, S. K., Turner, J. A., Com-stock, B. A., Hollingworth, W. and Sullivan, S. D., 2008. Expenditures and health status among adults with back and neck problems. *JAMA* [online], 299 (60), 656-664.

Martin, B., Mirza, S. K., Spina, N., Spiker, W. R., Lawrence, B. and Brodke, D. S., 2019. Trends in lumbar fusion procedure rates and associated hospital costs for degenerative spinal diseases in the united states, 2004 to 2015. *Spine* [online], 44 (5), 369-376.

Mathworks, 2019. *MATLAB. R*2019b [computer program]. Natick: Mathworks.

McGregor, A. H., Anderton, L., Gedroyc, W. M. W., Johnson, J. and Hughes, S. P. F., 2002. The use of interventional open MRI to assess the kinematics of the lumbar spine in patients with spondylolisthesis. *Spine* [online], 27 (14), 1582-1586.

Mellor, F. E., Thomas, P. W., Thompson, P. and Breen, A. C., 2014a. Proportional lumbar spine inter-vertebral motion patterns: a comparison of patients with chronic, non-specific low back pain and healthy controls. *European spine journal* [online], 23 (10), 2059-2067.

Mellor, F. E., Thomas, P. and Breen, A., 2014b. Moving back: radiation dose received from lumbar spine quantitative fluoroscopy compared to lumbar spine radiographs with suggestions for dose reduction. *Radiography* [online], 20 (3), 251-257. Available from: https://www.radiographyonline.com/article/S1078-8174(14)00041-8/fulltext#secsectitle0095 [Accessed 25 January 2021].

Mentiplay, B, F., Clark, R. A., Mullins, A., Bryant, A. L., Bartold, S. and Paterson, K., 2013. Reliability and validity of the Microsoft Kinect for evaluating static foot posture. *Journal of foot and ankle research* [online], 6. Available from: https://jfootankleres.biomedcentral.com/articles/10.1186/1757-1146-6-14 [Accessed 12 November 2020].

Metaxas, V., Messaris, G. A., Lekatou, A. N., Petsas, T. G. and Panayiotakis, G. S., 2019. Patient dose in digital radiography utilising BMI classificiation. *Radiation protection dosimetry* [online], 184 (2), 155-167.

Michelini, G., Corridore, A., Torlone, S., Bruno, F., Marsecano, C., Capasso, R., Caranci, F., Barile, A., Masciocchi, C. and Splendiani, A., 2018. Dynamic MRI in the evaluation of the spine: state of the art. *Acta biomdeica* [online], 89, 89-101.

Milne, A. D., Chess, D. G., Johnson, J. A. and King, G. J. W., 1996. Accuracty of an electromagnetic tracking device: A study of the optimal operating range and metal interference. *Journal of biomechanics* [online], 29 (6), 791-793.

Mirbagheri, S. S., Rasa, A., Farmani, F., Amini, P. and Nikoo, M-R., 2015. Evaluating kyphosis and lordosis in students by using a flexible ruler and their relationship with severity and frequency of thoracic and lumbar pain. *Asian spine journal* [online], 9 (3), 416-422.

Modica, M. J., Kanal, K. M. and Gunn, M. L., 2011. The obese emergency patient: imaging challenges and solutions. *Radiographics* [online], 31 (3), 811-823.

Morrissy, R. T., Goldsmith, G. S., Hall, E. C., Kehl, D. and Cowie, G. H., 1990. Measurement of the cobb angle on radiographs of patients who have scoliosis. Evaluation of intrinsic error. *The journal of bone and joint surgery* [online], 72 (3), 320-327.

Muggleton, J. M. and Allen, R., 1997. Automatic location of vertebrae in digitilized videofluoroscopic images of the lumbar spine. *Medical engineering and physics* [online], 19 (1), 77-89.

Murray, C. J. L., Vos, T., Lozano R., Naghavi, M., Flaxman, A. D., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J. A., Abdalla, S., Aboyans, V., Abraham, J., Ackerman, I., Aggarwal, R., Ahn, S. Y., Ali, M. K., ALMazroa, M. A., Alvarado, M., Anderson, H. R., Anderson, L. M., Andrews K. G., Atkinson, C., Baddour, L. M., Bahalim, A. N., Barker-Collo, S., Barrero, L. H., Bartels, D. H., Basanez, M-G., Baxter, A., Bell, M. L., Benjamin, E. J., Bennett, D., Bernabe, E., Bhalla, K., Bhandari, B., Bikbov, B., Abdulhak, A. B., Birbeck, G., Black, J. A., Blencow, H., Blore, J. D., Blyth, F., Bolliger, I., Bonaventure, A., Boufous, S., Bourne, R., Boussinesq, M., Braithwaite, T., Brayne, C., Bridgett, L., Brooker, S., Brooks, P., Brugha, T. S., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Buckle, G., Budke, C. M., Burch, M., Burney, P., Burstein, R., Calabria, B., Campbell, B., Canter, C. E., Carabin, H., Carapetis, J., Carmona, L., Cella, C., Charlson, F., Chen, H., Cheng, A., Chou, D., Chugh, S., Coffeng, L., Colan, S., Colquhoun, S., Colson, K, Condon, J., Connor, M. D., Cooper, L. T., Corriere, M., Cortinovis, M., Courville de Vaccaro, K., Couser, W., Cowie, B. C., Criqui, M. H., Cross, M., Dabhadkar, K. C., Dahiya, M., Dahodwala, N., Damsere-Derry, J., Danaei, G., Davis, A., De Leo, D., Degenhardt, L., Dellavalle, R., Delossantos, A., Denenberg, J., Derrett, S., Jarlais, D. C. D., Dharmaratne, S. D., Dherani, M., Diaz-Torne, C., Dolk, H., Dorsey, E. R., Driscoll, T., Duber, H., Ebel, B., Edmond, K., Elbaz, A., Ali, S. E., Erskine, H., Erwin, P. J., Espindola, P., Ewoigbokhan, S. E., Farzadfar, F., Feigin, V., Felson, D. T., Ferrari, A., Ferri, C. P., Fevre, E. M., Finucane, M. M., Flaxman, S., Flood, L., Foreman, K., Forouzanfar, M. H., Fowkes, F. G. R., Fransen, M., Freeman, M. K., Gabbe, B. J., Gabriel, S. E., Gakidou, E., Ganatra, H. A., Garcia, B., Gaspari, F., Gillum, R. F., Gmel, G., Gonzalez-Medina, D., Gosselin, R., Grianger, R., Grant, B., Groeger, J., Guillemin, F., Gunnell, D., Gupta, R., Haagsma, J., Hagan, H., Halasa, Y. A., Hall,

W., Haring, D., Haro, J. M., Harrison, J. E., Havmoeller, R., Hay, R. J., Higashi, H., Hill, C., Hoen, B., Hoffman, H., Hotez, P. J., Hoy, D., Huang, J. J., Ibeanusi, S. E., Jacobsen, K. H., James, S. L., Jarvis, D., Jasrasaria, R., Jayaraman, S., Johns, N., Jonas, J. B., Karthikeyan, G., Kassebaum, N., Kawakami, N., Keren, A., Khoo, J-P., King, C. H., Knowlton, L. M., Kobusingye, O., Koranteng, A., Krishnamurthi, R., Laden, F., Lalloo, R., Laslett, L. L., Lathleam T., Leasher, J. L., Lee, Y. Y., Leigh, J., Levinson, D., Lim, S. S., Limb, E., Lin, J. K., Lipnick, M., Lipshultz, S. E., Liu, W., Loane, M., Ohno, S. L., Lyons, R., Mabweijano, J., MacIntyre, M. F., Malekzadeh, R., Mallinger, L., Manivannan, S., Marcenes, W., March, L., Marjolis, D. J., Marks, G. B., Marks, R., Matsumori, A., Matzopoulos, R., Mayosi, B. M., McAnulty, J. H., McDermott, M. M., McGill, N., McGrath, J., Medina-Mora, M. E., Meltzer, M., Memish, Z. A., Mensah, G/ A/. Merriman, T. R., Meyer, A. C., Miglioli, V., Miller, M., Miller, T. R., Mitchell, P. B., Mock, C., Mocumbi, A. O., Moffitt, T. E., Mokdad, A. A., Monasta, L., Montico, M., Moradi-Lakeh, M., Moran, A., Morawska, L., Mori, R., Murdoch, M. E., Mwaniki, M. K., Naidoo, K., Nair, N., Naldi, L., Narayan, V., Nelson, P. K., Nelson, R. G., Nevitt, M. C., Newton, C. R., Nolte, S., Normal, P., Normal., R., O'Donnell, M., O'Hanlon, S., Olives, C., Omer, S. B., Ortblad, K., Osborne, R., Ozgediz, D., Page, A., Pahari, B., Pandian, J. D., Rivero, A. P., Patten, S. B., Pearce, N., Padilla, R. P., Perez-Ruiz, F., Perico, N., Pesudovs, K., Phillips, D., Phillips, M. R., Pierce, K., Pion, S., Polanczyk. G. V., Polinder, S., Pope III, C. A., Popova, S., Porrini, E., Piurmalek, F., Prince, M., Pullan, R. L., Ramaiah, K. D., Ranganathan, D., Razavi, H., Regan, M., Rehm, J. ., Rein, D. B., Remuzzi, G., Richardson, K., Rivara, F. P., Roberts, T., Robinson, C., Leon, F. R. D., Ronfani, L., Room, R., Rosenfeld, L. C., Rushton, L., Sacco, R. L., Saha, S., Sampson, U., Sanchez-Riera, L., Sanman, E., Schwebel, D. C., Scott, J. G., Segui-Gomez, M., Shahraz, S., Shepard, D. S., Shin, H., Shivakoti, R., Silbeberg, D., Singh, D., Singh, G. M., Singh, J. A., Singleton, J., Sleet, D. A., Sliwa, K., Smith, E., Smith, J. L., Stapelberg, N. J. C., Steer, A., Steiner, T., Stolk, W. A., Stovner, L. J., Sudfield, C., Syed, S., Tamburlini, G., Tavakkoli, M., Taylor, H. R., Taylor, J. A., Taylor, W. J., Thomas, B., Thomson, W. M., Thurston, G. D., Tleyjeh, I. M., Tonelli, M., Towbin, J. A., Truelsen, T., Tsilimbaris, M. K., Ubeda, C., Undurraga, E. A., van der Werf, M. J., Os, J. V., Vavilala, M. S., Venketasubramanian, N., Wang, M., Wang, W., Watt, K., Weatherall, D. J., Weinstock, M.A., Weintraub, R., Weisskopf, M. G., Weissman, M. M., White, R. A., Whiteford, H., Wiebe, N., Wiersma, S. T., Wilkonson, J. D.,

Williams, H. C., Williams, S. R. M., Witt, E., Wolfe, F., Woolf, A. D., Wulf, S., Yeh, P-H., Zaidi, A. K. M., Zheng, Z-J., Zonies, D. and Lopez, A. D., 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet* [online], 380 (9859), 2197-2223.

National heart, lung and blood institute (NHLBI), 2019. *Study quality assessment tools* [online]. Bethesda: NHLBI. Available from: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools [Accessed 25 October 2019].

National heart, lung and blood institute (NHLBI), ca.2020. *Laboratory of cardiac energetics: What are the technological advantages and limitations (disadvantages) of MRI* [online], Bethesda: NHLBI. Available from: https://dir.nhlbi.nih.gov/labs/lce/cmri/mri-advantages-limitation.asp [Accessed 30 March 2020].

Ng, L., Burnett, A., Campbell, A. and O'Sullivan, P., 2009. Caution: The use of an electromagnetic device to measure trunk kinematic on rowing ergometers. *Sports biomechanics* [online], 8 (3), 255-259.

Oakley, P. A. and Harrison, D. E., 2018. Radiophobia: 7 reasons why radiography used in spine posture rehabilitation should not be feared or avoided. *Dose-response* [online], 16 (2).

Okawa, A., Shinomiya, K., Komori, H., Muneta, T., Arai, Y. and Nakai, O., 1998. Dynamic motion study of the whole lumbar spine by videofluoroscopy. *Spine* [online], 23 (16), 1743-1749.

Parent, R., 2012. Computer animation [online]. 3rd edition. Amsterdam: Elsevier.

Park, S-A., Fayyazi, A. H., Yonemura, K. S. and Fredrickson, B. E., 2012. An in vivo kinematic comparison of dynamic lumbar stabilization to lumbar discectomy and posterior lumbar fusion using radiostereometric analysis. *International journal of spine surgery* [online], 1 (6), 87-92.

Parkinson, S., Campbell, A., Dankaerts, W., Burnett, A. and O'Sullivan, P., 2013. Upper and lower lumbar segments move differently during sit to stand. *Manual therapy* [online], 18 (5), 390-394.

Pazos, V., Cheriet, F., Danserau, J., Ronsky, J., Zernicke, R. F. and Labelle, H., 2007. Reliability of trunk shape measurements based on 3-D surface reconstructions. *European spine journal* [online], 16 (11), 1882-1891.

Pazos, V., Cheriet, F., Song, L., Labelle, H. and Dansereau, J., 2005. Accuracy assessment of human trunk surface 3D reconstructions from an optical digitising system. *Medical & biological engineering & computing* [online], 43 (1), 11-15.

Pearson, A. M., Spratt, K. F., Genuario, J., McGough, W., Kosman, K., Lurie, J., and Sengupta, D. K., 2011. Precision of lumbar intervertebral measurements. *Spine* [online], 36 (7), 572-580.

Pellios, S., Kenanidis, E., Potoupnis, M., Tsiridis, E., Sayegh, F. E., Kirkos, J. and Kapetanos, G. A., 2016. Curve progression after bracing for adolescent idiopathic scoliosis: long term comparative results between two matched groups of 18 versus 23 hours daily bracing. *Scoliosis and spinal disorders* [online], 11, 3.

Plaszewski, M., Nowobilski, R., Kowalski, P. and Cieslinski, M., 2012. Screening for scoliosis: different countries' perspectives and evidence-based health care. *International journal of rehabilitation research* [online], 35 (1).

Plocharski, M., Lindstroem, R., Lindstroem, C. F., Østergard, R., 2018. Motion analysis of the cervical spine during extension and flexion: Reliability of the verterbral marking procedure. *Medical engineering and physics* [online], In Press. Available from: https://www.sciencedirect.com/science/article/abs/pii/S135045331830119X [Accessed 03 February 2020].

Polhemus, 2020a. *Patriot* [brochure] [online]. Vermont: Polhemus.

Polhemus, 2020b. *Liberty Latus* [brochure] [online]. Vermont: Polhemus.

Post, R. B. and Leferink, V. J. M., 2004. Spinal mobility: sagittal range of motion measure with the spinal mouse, a new non-invasive device. *Archives of orthopaedic and trauma surgery* [online], 124 (3), 187-192.

Pourahmadi, M. R., Takamjani, I. E., Jaberzadeh, S., Sarrafzadeh, J., Sanjari, M. A., Bagheri, R. and Taghipour, M., 2018. Kinematics of the spine during sit-to-stand movement using motion analysis systems: A systematic review of literature. *Journal of sport rehabilitation* [online], 13, 1-17.

Preim, B. and Botha, C., 2014. Visual computing for medicine [online]. 2nd edition. Amsterdam: Elsevier.

Quek, J., Brauer, S. G., Treleaven, J. and Clark, R. A., 2017. The concurrent validity and intrarater reliability of the Microsoft Kinect to measure thoracic kyphosis. *International journal of rehabilitation research* [online], 40 (3), 279-284.

Raciborski, F., Gasik, R. and Klak, A., 2016. Disorders of the spine. A major health and social problem. *Reumatologia* [online], 54 (4), 196-200.

Ramdas, J. and Jella, V., 2018. Prevalance and risk factors of low back pain. *International journal of advances in medicine* [online], 5 (5), 1120-1123.

Reeves, P., Choelwicki, J., Milner, T. and Lee, A. S., 2008. Trunk antagonist co-activation is associated with impaired neuromuscular performance. *Experimental brain research* [online], 188 (3), 457-463.

Roghani, T., Zavieh, M. K., Rahimi, A., Talebian, S., Manshadi, F. D., Baghban, A. A., King, N. and Katzman, W., 2017. The reliability of standing sagittal measurements of spinal curvature and range of motion in older women with and without hyperkyphosis using a skin-surface device. *Journal of manipulative and physiological therapeutics* [online], 40 (9), 685-691.

Rozumalski, A., Schwartz, M. H., Wervey, R., Swanson, A., Dykes, D, Novacheck, T., 2008. The in vivo three-dimensional motion of the human lumbar spine during gait. *Gait posture* [online], 28 (3), 378-384.

Schmid, S., Studer, D., Haslet, C-C., Romkes, J., Taylor, W. R., Brunner, R. and Lorenzetti, S., 2015. Using skin markers for spinal curvature quantification in main thoracic and adolescent idiopathic scoliosis: An explorative radiographic study. *PLOS one* [online]. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0135689 [Accessed 24 November 2020].

Schuler, N. B., Bey, M. J., Shearn, J. T. and Butler, D. L., 2005. Evaluation of an electromagnetic position tracking device for measuring in vivo, dynamic joint kinematics. *Journal of biomechanics* [online], 38 (10), 2113-2117.

Schwab, F., Dubey, A., Gamez, L., El Fegoun, A. B., Hwang, K., Pagala, M. and Farcy, J-P., 2005. Adult scoliosis: prevalence, SF-36, and nutritional parameters in an elderly volunteer population. *Spine* [online], 30 (9), 1082-1085.

Sedrez, J. A., Candotti, C. T., Rosa, M. I. Z., Medeiros, F. S., Marques, M. T. and Loss, J. F., 2016. Test-retest, inter- and intra-rater reliability of the flexicurve for evaluation of the spine in children. *Brazilian journal of physical therapy* [online], 20 (2), 142-147.

Sett, P. and Crockard, A., 1991. The value of magnetic resonance imaging (MRI) in the follow-up management of spinal injury. *Paraplegia* [online], 29 (6), 396-410.

Singh, D. K., Bailet, M. and Lee, R., 2010. Biplanar measurement of thoracolumbar curvature in older adults using an electromagnetic tracking device. *Archives of physical medicine and rehabilitation* [online], 91, 137-142.

Stergiou, N. and Decker, L M., 2011. Human movement variability, nonlinear dynamics and pathology: Is there a connection?. *Human movement science* [online], 30 (5), 869-888.

Sui, F., Zhang, D., Lam, S. C. B., Zhao, L., Wang, D., Bi, Z. and Hu, Y., 2011. Auto-tracking system for human lumbar motion analysis. *Journal of x-ray science and technology* [online], 19 (2), 205-218.

Takayanagi, K., Takahashi, K., Yamagata, M., Moriya, H., Kitahara, H. and Tamaki, T., 2001. Using cineradiography for continuous dynamic-motion analysis of the lumbar spine. *Spine* [online], 26 (17), 1858-1865.

Tarantino, U., Fanucci, E., Lundusi, R., Celi, M., Altobelli, S., Gasbarra, E., Simonetti, G. and Manenti, G., 2013. Lumbar spine MRI in upright position for diagnosing acute and chronic low back pain: statistical analysis of morphological changes. *Journal of orthopaedics and traumatology* [online], 14 (1), 15-22.

Teixeira, F. A. and Carvalho, G. A., 2007. Reliability and validity of thoracic kyphosis measurements using flexicurve method. *Brazilian journal of physical therapy* [online], 11 (3), 199-204.

The world federation for ultrasound in medicine and biology (WFUMB), 2013. WFUMB policy and statements on safety of ultrasound. *Ultrasound in medicine and biology* [online], 39 (5), 926-929.

Topiladou, A., Tzagarakis, G., Souvatzis, X., Kontakis, G. and Katonis, P., 2014. Evaluation of the reliability of a new non-invasive method for assessing the functionality and mobility of the spine. *Acta of bioengineering and biomechanics* [online], 16 (1).

Tozawa, R., Katoh, M., Aramaki, H., Kawasaki, T., Nishikawa, Y., Kumamoto, T. and Fujinawa, O., 2016. Absolute and relative reliability of interspinous process ultrasound imaging measurements. *The journal of physical therapy science* [online], 28 (8), 2210-2213.

Tozawa, R., Katoh, M., Aramaki, H., Kumamoto, T. and Fujinawa, O., 2015. Reliability and validity of an ultrasound-based imaging method for measuring interspinous process distance in the lumbar spine using two different index points. *The journal of physical therapy science* [online], 27 (7), 2333-2336.

Trompetto, C., Marinelli, L., Mori, L., Pelosin, E., Curra, A., Molfetta, L. and Abbruzzese, G., 2014. Pathophysiology of spasticity: implications for neurorehabilitation. *Biomed research international* [online], Available from: https://www.hindawi.com/journals/bmri/2014/354906/ [Accessed 27 April 2020].

Trudelle-Jackson, E., Fleisher, L. A., Borman, N., Morrow, J. R. and Frierson, G. M., 2010. Lumbar spine flexion and extension extremes of motion in women of different age and racial groups: the WIN study. *Spine* [online], 35 (16), 1539-1544.

Vrtovec, T., Pernus, F. and Likar, B., 2009. A review of methods for quantitative evaluation of spinal curvature. *European spine journal* [online], 18, 593-607.

Was, J., Sitarski, D., Ewertowska, P., Bloda, J. and Czaprowski, D., 2016. Using smartphones in the evaluation of spinal curvatures in a sagittal plane. *Rehabilitacja* [online], 4, 29-38.

Wassenaar, M., Van Rijn, R. M., Van Tulder, M. W., Verhagen, A. P., Van Der Windt, D. A. W. M., Koes, B. W., De Boer, M. R., Ginai, A. Z. and Ostelo, R. W. J. G., 2012. Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patient with low back pain or sciatica: a diagnostic systematic review. *European spine journal* [online], 21 (2), 220-227.

Williams, J. M., Haq, I. and Lee, R. Y., 2010. Dynamic measurement of lumbar curvature using fibre-optic sensors. *Medical engineering and physics* [online], 32, 1043-1049.

Williams, J. M., Haq, I. and Lee, R. Y., 2012. Dynamic lumbar curvature measurement in acute and chronic low back pain sufferers. *Archives of physical medicine and rehabilitation* [online], 93, 2094-2099.`

Williams, J. M., Haq, I. and Lee, R. Y., 2013. The effect of pain relief on dynamic changes in lumbar curvature. *Manual therapy* [online], 18 (2), 149-154.

Wong, W. Y. and Wong, M. S., 2008. Trunk posture monitoring with inertial sensors. *European spine journal* [online], 17 (5), 743-753.

Wright, N. C., Looker, A. C., Saag K. G., Curtis, J. R., Delzell, E. S., Randall, S. and Dawson-Hughes, B., 2014. The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. *Journal of bone and mineral research* [online], 29 (11), 2520-2526.

Xsens, 2020. *Xsens DOT user manual* [online]. Enschede: Xsens.

Yeager, M. S., Cook, D. J. and Cheng, B. C., 2014. Reliability of computer-assisted lumbar intervertebral measurements using a novel vertebral motion analysis system. *The spine journal* [online], 14 (2), 274-281.

Youdas, J. W., Suman, V. J. and Garrett, T. R., 1995. Reliability of measurements of lumbar spine sagittal mobility obtained with the flexible curve. *Journal of orthopaedic and sports physical therapy* [online], 21 (1).

Yukawa, Y., Matsumoto, T., Kollor, H., Minamide, A., Hashizume, H., Yamada, H. and Kato, F., 2019. Local sagittal alignment of the lumbar spine and range of motion in 627 asymptomatic subjects: age-related changes and sex-based differences. *Asian spine journal* [online], 13 (4), 663-671.

Zabjek, K. F., Leroux, M. A., Coillard, C., Rivard, C-H. and Prince, F., 2005. Evaluation of segmental postural characteristics during quiet standing in control and idiopathic scoliosis patients. *Clinical biomechanics* [online], 20 (15), 483-490.

Zhang, R., Duan, Y., Zhao, Y. and He, X., 2018. Temperature compensation of elasto-magneto-electric (EME) sensors in cable force monitoring using BP neural network. *Sensors* [online], 18 (7), 2176.