The reliability and reproducibility of sagittal spinal curvature measurement using the Microsoft Kinect V2

Erin Hannink[a,b,c]
Email: erin.hannink-2017@brookes.ac.uk | Tel: 01865 483272
Gipsy Lane, Headington, Oxford, OX3 0BP, United Kingdom

Thomas Shannon[d]

Karen L. Barker[b,c*]

Helen Dawes[a*]

a: Centre for Movement, Occupational and Rehabilitation Sciences, Oxford Brookes University, Gipsy Lane, Headington, Oxford, OX3 0BP, United Kingdom

b: Physiotherapy Research Unit, Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Foundation Trust, Windmill Road, Oxford, OX3 7LD, United Kingdom

c: Nuffield Department of Orthopaedic, Rheumatoid and Musculoskeletal Sciences, University of Oxford, Windmill Road, Oxford, OX3 7LD, United Kingdom

d: Faculty of Health Sciences, Staffordshire University, College Road, Stoke-on-Trent, ST4 2DE, United Kingdom

*: Joint senior authorship

ABSTRACT

BACKGROUND: Abnormal sagittal spinal curvature is associated with pain, decreased mobility, respiratory problems and increased mortality. Time-of-flight technology of the Microsoft Kinect sensor can reconstruct a three-dimensional image of the back quickly and inexpensively.

OBJECTIVE: To estimate the extent of the reproducibility of sagittal spine curvature measurement using the Microsoft Kinect sensor.

METHODS: Simultaneous measurement of thoracic and lumbar spine using the Microsoft Kinect sensor in 37 participants. Two investigators gave standardised instructions and each captured 3 images. Thoracic kyphosis and lumbar lordosis angle indexes were calculated using maximum height divided by the length.

RESULTS: Adult participants (mean age in years (SD)= 51.7 (20.6); 57% female; BMI in kg/m^2 (SD)= 24.9 (3.3)) kyphosis and lordosis indexes showed high intra-rater and inter-rater ICC values (0.960-0.973). The means of the first images from both raters had significantly larger kyphosis indexes compared to the second and third images, yet no difference between means in lordosis data.

CONCLUSIONS: The results indicate that the Microsoft Kinect sensor has a reproducible method with high intra-rater and inter-rater reliability. The difference between the means over repeated measures suggest the second image capture is more consistent. It is a reproducible and quick method in clinical and research settings.

1. INTRODUCTION

A surface topography method using the Microsoft Kinect sensor version 2 (Microsoft Corporation, Seattle, Washington, U.S.A) has the potential to improve the ease of measurement, increase accuracy, and provide a robust analysis of spinal curvature for clinicians and researchers. The time-of-flight (ToF) technology of the Kinect sensor V2 reconstructs a three-dimensional image of the back quickly and inexpensively [1]. Kinect sensors are depth cameras that have been used in various commercial and research fields, both in and out of the healthcare scope. Within the area of human movement and posture, it has been used in several different capacities to measure aspects of spinal posture, from the movement analysis of postural control and ergonomic positions to the cosmetic defect of the surface of the back [2–7]. A number of these studies [2–5] employed the first version of the Kinect sensor that used structured light to compare the regular geometry of the projected infrared light to the distortion of the light on the object. Several studies have since researched similar areas of spine and posture measurement using the second version of the Kinect sensor which uses a ToF [7,8], a fundamentally different technology for surface measurement [9].

While other studies have used the Kinect sensors to measure topography of the back to estimate the lateral and rotational curvature associated with adolescent idiopathic scoliosis (AIS) [1,6], the focus of this study was to examine specifically the spinal curvature in the sagittal plane which is in itself an important outcome to measure and understand due to its numerous health implications. Abnormal sagittal spinal curvature, in particular hyperkyphosis, occurs not only in the kyphoscoliotic population, but also commonly in people with degenerative spinal diseases such as osteoporosis, vertebral fractures and degenerative disc disease, as well as congenital diseases, such as Scheurmann's Disease. Hyperkyphosis is associated with poor health-related outcomes with the ageing process, such as pain, decreased mobility, respiratory problems, and increased mortality [10–12]. Because the normative values of sagittal spine alignment are widely defined and morph with the natural course of ageing, the ability to measure change over time is crucial. Older women have been shown to increase kyphosis

Cobb angle by a mean of 7 degrees over 15 years [13]. Monitoring sagittal alignment is pertinent as change is more rapid with the progression of degenerative spinal conditions, e.g. osteoporosis and degenerative disc disease. The current gold standard for measuring sagittal spine curvature is the Cobb method using a lateral view spinal radiograph; however its drawbacks include its cost, exposure to ionising radiation, and reliance on the morphology of two vertebral endplates versus the entire spinal shape [14,15].

The Kinect sensor offers a non-invasive, non-ionising radiation solution that can measure multiple regions of the spine concurrently and quickly. While the Kinect sensor has been researched for various healthcare applications, the reliability of the Kinect sensor V2 specifically applied to the sagittal plane of the spine in the thoracic and lumbar regions simultaneously has not been researched. The primary aim of this study is to estimate the extent of the reliability and reproducibility of sagittal spinal curvature measurement method using the Kinect sensor.

2. METHODS

2.1 Design

This cross-sectional study was evaluated according to the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) and on the COSMIN pathway for reliability [16].

2.2 Participants

Adults were eligible to participate if they could stand independently and if they had no neurological disorders affecting their posture. There were no exclusion criteria based on existence of abnormal spinal curvature or alignment. All participants provided an informed consent and the study was approved the Oxford Brookes University Faculty Research Ethics Committee in accordance with the Declaration of Helsinki. A sample size calculation for intraclass correlation coefficients with

4

significance level at 0.05, power at 0.80, acceptable reliability at 0.75 and expected reliability at 0.90 required at least 19 participants [17].

2.3 Procedures

Investigator 1 was an experienced musculoskeletal physiotherapist who palpated the participants' spine and placed adhesive markers to identify the specific anatomical landmarks: C7, L1, and right and left posterior superior iliac spine (PSIS) (Figure 1). The spinous process of C7 was identified by palpating the prominent spinous process in the cervicothoracic junction, confirmed by active cervical extension and flexion. The spinous process of L1 was identified by its relationship to L4 which is located horizontal to the level of the iliac crests and two vertebrae superior to the sacral base; after L4 was confirmed, the investigator counted the spinous processes up to L1. Bilateral PSIS locations were identified by surface dimples at the level of S2. Participants were instructed to stand facing away from the sensor with their heels on a line marked one metre from the sensor. The sensor was aligned parallel to the ground and adjusted in height to be level with the participants' mid-scapular region (Figure 2). Both investigators gave instructions for 'best posture' with arm positioning forward with their shoulders and elbows at 90 degrees. Participants stood in their 'best posture' for less than 3 seconds as the image was captured. Between each image capture, the participant was instructed to walk away then return to the same standing position and posture. Investigator 1 captured three images using this procedure and was blinded to the results of each capture. On the same research visit, the second investigator, a non-clinician, who was blinded to previous investigator's results, captured three subsequent images using the same anatomical landmarks previously identified.

2.4 Kinect data analysis

The Kinect sensor was used with Kinect Fusion software (Microsoft Corporation, Seattle, Washington, U.S.A). All images were uploaded to MeshLab, a 3D mesh processing open source software system [18], then processed through a bespoke software programme, Parser (Oxford Metrics

plc/Staffordshire University Stanford Polygon PLY, version 2.6) [19], which required manual identification of the landmarks in order to then obtain coordinates along the spine and subsequently the angle indexes were calculated. The angle indexes were calculated using the length of the target spinal region and the maximum height which is an accepted method widely used with Flexicurve measurement and validated against the Cobb angle [20], e.g. kyphosis index = (height / (length from C7 to L1)) x 100; the same method was applied to the lumbar region with the length extending from L1 to bisection of right and left PSIS (Figure 1).

2.5 Statistical analysis

Data was analysed using SPSS version 24 software (SPSS, Inc., Chicago, Illinois, U.S.A.). Descriptive statistics and reliability coefficients were calculated. Reliability was based on Intraclass Correlation Coefficient (ICC) with 95% confidence intervals; intra-rater reliability was based on a single rater, absolute agreement, two-way mixed effects model (ICC 3,1), and inter-rater reliability was based on a single rater, consistency, two-way random-effects model (ICC 2,1) [21]. Levels of reliability used were: poor reliability is <0.40, moderate reliability is 0.40-0.75, good reliability is 0.75-0.90, and excellent reliability is >0.90 [22]. Standard error of measurement (SEM) and minimal detectable difference (MDD) were determined by the equations: $SEM = SD\sqrt{1\text{-reliability}^2}$ and $MDD = SEM*1.96*\sqrt{2}$. Comparison of multiple means analysed using repeated measures ANOVA (p<0.05).

3. RESULTS

Thirty-seven participants aged between 18 and 79 years old had a mean (SD) age of 51.7 (20.6) years old. Their mean BMI (SD) was 24.9 (3.3) kg/m^2 and 57% were female. Descriptive statistics of kyphosis and lordosis indexes are reported in Tables 1 and 2. The kyphosis and lordosis indexes have very high intra-rater and inter-rater ICC scores (ICC= 0.960-0.973) (Tables 1-2).

The means of the six images taken were compared. There were no significant differences between the means in the lumbar region, $F(3.91,133.01) = 1.566$, $p = 0.188$; however there were significant differences between the means in the thoracic region, $F(5,170) = 5.317$, $p < 0.001$ (Table 1). To look at the differences in the thoracic spine as they related to age and the severity of kyphosis, means within these subgroups showed that in participants who were less than 65 years old or who had normal kyphosis, image captures 1 and 4 were significantly higher than image capture 2 (Figure 3). In participants who were 65 years or older or who were hyperkyphotic, there were no significant differences between the means of each capture (Figure 3). The data were further reformatted to pool the first, second and third images from both investigators (Tables 1-2) in order to determine if posture changed based on the capture order in reference to the introduction of a new investigator. For pooled thoracic data, there was a significant effect between the means, $F(1.76,122.86) = 11.257$, $p < 0.001$; the mean of the first images pooled had a significantly larger kyphosis index ($p<0.001$) compared to the second images and third images (Table 1). For pooled lumbar data, there was no statistically significant differences, $F(2,140) = 0.004$, $p = 0.996$ (Table 2).

The participant-facing time it took for instruction and positioning was less than 30 seconds and actual image capture was three seconds. For each participant, the image capture and analysis all images were less than five minutes.

4. DISCUSSION

The ICC results show that the Kinect sensor has high intra-rater and inter-rater reliability in the both thoracic and lumbar regions. It was shown to be a quick method, as it took less than five minutes to complete both image capture and analysis of both regions. In any setting, this would be a quick, reliable, and inherently safe non-radiographic method to measure thoracic and lumbar curvature concurrently to use as an outcome measure with a patient-friendly visual representation of the back.

While there have been numerous non-radiological methods developed to measure sagittal plane spinal curvature, most have focused on measuring the thoracic region, some measure the lumbar region, and few measure both regions simultaneously. Commonly used non-invasive measurement methods of thoracic kyphosis, such as the flexicurve and DeBrunner kyphometer, have shown high intra-rater and inter-rater reliability and moderate validity [20,23]. In the lumbar region, the flexicurve has been shown to have moderate to high reliability and moderate validity [24,25]. In one study that measured thoracic and lumbar regions concurrently with a flexicurve, the method involved taking a digital picture of the flexicurve against graph paper and manually plotting coordinates from the curve which would be time consuming for a clinician to perform [25]. While modern non-radiographic methods have been developed, such as the SpinalMouse and rastersstereography, cost has been a barrier to the practical constraints of a clinical setting [23].

While the use of Kinect sensor in AIS populations to measure lateral and rotational curvature has been growing quickly, to date, only one published study has used the Kinect sensor V2 to specifically measure thoracic kyphosis. Quek et al. used the Kinect sensor to measure kyphosis in sitting and standing positions and found good reliability and concurrent validity against the flexicurve (ICC=0.81-0.98) [7]. Their participant population had a mean (SD) age of 31 (11) years and standing kyphosis index mean (SD) of 9.78 (2.4) [7]. Our study found similar reliability values in a population which more closely represents the adult life span, including older adults with degenerative spinal changes. In this study participants averaged 51.7 years old with a kyphosis index mean (SD) of 10.36 (3.2), indicating a larger range in spinal shape and a notably wider diversity in age. A fifth of the participants in our study surpassed the threshold for hyperkyphosis, which is a kyphosis index ≥ 13.1 [10,26]. With this diverse sample we can expand the generalisability of our results since it is known that adults undergo a natural course of change in sagittal spine curvature in the later decades of life with degenerative spinal diseases furthering the change, as hyperkyphosis is prevalent in 20-40% of older adults [10,27].

Understanding the difference between the means of repeated measurements led to a potentially important difference due to degenerative spinal changes. The significant differences between means appeared to be driven by younger adults within normal spinal curvature ranges. This suggests that degenerative changes may limit the variability in 'best posture' possibly due to reduction in spinal mobility, yet the small sample size limits extrapolation of these results. Additionally, analysing the difference between repeated measurements based on the order by pooling the data of the two investigators was important to help identify the best protocol for future use. The data in the thoracic and lumbar regions indicate only one significant mean difference, which was in the thoracic spine during the first image capture, suggesting that the second image capture is more reliable to use. Since taking a second image adds only 15 seconds to the testing time, the method remains a quick execute.

Since the first investigator palpated and identified the anatomical landmarks, these study results focus on the error of the sensor hardware, bespoke software, procedural instructions, and participants' interpretation of 'best posture.' It is known that spinal landmark palpation is variable, even between experienced physiotherapists [28,29], therefore removing additional anatomical palpation eliminated error stemming from two investigators' palpation skills. . Innovative research has been conducted to develop automatic estimation of anatomical landmarks which would offer a good solution to error generated from palpation, however further validation for this system is still required [30].

While the sample size is adequate for ICC analysis of reliability for the full sample, it is not powered to measure differences in subgroups based on age or the degree of curvature, which is a limitation to the study. This study has a more clinically applicable sample population to examine sagittal spine curvature but lacks the large cohort that some studies [6] have used to examine lateral and rotational spinal curvature. With these limitations in mind, the results of this study lay the groundwork to test other aspects of Kinect sensor measurement reliability on the COSMIN pathway, such as test-retest reliability.

Overall, using the Kinect sensor to simultaneously measure thoracic and lumbar curvature in the sagittal plane is a reliable, quick, reproducible method. Since many of the participants had normal spinal alignment, future research should measure a population diagnosed with specific spinal deformity or malalignment and validate this method against the gold standard.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1]   Shannon T, Chockalingam N. Investigating a Low Cost Method to quantify Cosmetic Defect. Stud Health Technol Inform. 2012;176:282–5.
[2]   Dutta T. Evaluation of the Kinect^{TM} sensor for 3-D kinematic measurement in the workplace. Appl Ergon. 2012 Jul 1;43(4):645–9.
[3]   Diego-Mas JA, Alcaide-Marzal J. Using Kinect^{TM} sensor in observational methods for assessing postures at work. Appl Ergon. 2014;45:976–85.
[4]   Clark RA, Pua Y-H, Fortin K, Ritchie C, Webster KE, Denehy L, et al. Validity of the Microsoft Kinect for assessment of postural control. Gait Posture. 2012;36:372–7.
[5]   Galna B, Barry G, Jackson D, Mhiripiri D, Olivier P, Rochester L. Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. Gait Posture. 2014 Apr 1;39(4):1062–8.
[6]   Castro APG, Pacheco JD, Lourenç O C, Queirós S, Moreira AHJ, Rodrigues NF, et al. Evaluation of spinal posture using Microsoft Kinect TM : A preliminary case-study with 98 volunteers. Porto Biomed J. 2017;2(1):18–22.
[7]   Quek J, Brauer SG, Treleaven J, Clark RA. The concurrent validity and intrarater reliability of the Microsoft Kinect to measure thoracic kyphosis. Int J Rehabil Res. 2017;40(3):279–84.
[8]   Plantard P, Muller A, Pontonnier C, Dumont G, Shum HPH, Multon F. Inverse dynamics based on occlusion-resistant Kinect data: Is it usable for ergonomics? Int J Ind Ergon. 2017 Sep 1;61:71–80.
[9]   Sarbolandi H, Lefloch D, Kolb A. Kinect range sensing: Structured-light versus Time-of-Flight Kinect. Comput Vis Image Underst. 2015;139:1–20.
[10]   Katzman WB, Wanek L, Shepherd JA, Sellmeyer DE. Age-Related Hyperkyphosis: Its Causes, Consequences, and Management. J Orthop Sport Phys Ther. 2010;40(6):352–60.
[11]   Kado DM, Prenovost K, Crandall C. Narrative Review : Hyperkyphosis in Older Persons. Ann Intern Med. 2007;(147):330–8.
[12]   Good CR, Auerbach JD, O'Leary PT, Schuler TC. Adult spine deformity. Curr Rev Musculoskelet Med. 2011;4(4):159–67.
[13]   Kado DM, Huang M, Karlamangla A, Cawthon P, Katzman WB, Hillier T, et al. Factors

Associated With Kyphosis Progression in Older Women: 15 years experience in the Study of Osteoporotic Fractures. J Bone Min Res. 2013;28(1):179–87.

[14]   Vrtovec T, Pernuš F, Likar B. A review of methods for quantitative evaluation of spinal curvature. Eur Spine J. 2009;18(5):593–607.

[15]   Briggs AM, Wrigley T V., Tully EA, Adams PE, Greig AM, Bennell KL. Radiographic measures of thoracic kyphosis in osteoporosis: Cobb and vertebral centroid angles. Skeletal Radiol. 2007;36(8):761–7.

[16]   Kottner J, Audig L, Brorson S, Donner A, Gajewski BJ. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. 2011;64:96–106.

[17]   Walter SD, Eliasziw M, Donner A. SAMPLE SIZE AND OPTIMAL DESIGNS FOR RELIABILITY STUDIES. 1998;17(April 1997):101–10.

[18]   Cignoni P, Cignoni P, Callieri M, Callieri M, Corsini M, Corsini M, et al. MeshLab: an Open-Source Mesh Processing Tool. Sixth Eurographics Ital Chapter Conf. 2008;129–36.

[19]   Shannon T, Chockalingam N, Jevtić N. The Radiological and Clinical Assessment of a Cohort of AIS Patients in Serbia and Bulgaria. Spine J. 2017 Nov 1;17(11):S329.

[20]   Greendale GA, Nili NS, Huang MH, Seeger L, Karlamangla AS. The reliability and validity of three non-radiological measures of thoracic kyphosis and their relations to the standing radiological Cobb angle. In: Osteoporosis International. 2011. p. 1897–905.

[21]   Rankin G, Stokes M. Reliability of assessment tools in rehabiliation: an illustration of appropriate statistical analyses. Clin Rehabil. 1998;12(98):187–99.

[22]   Fleiss JL. The Design and Analysis of Clinical Experiments. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 1986. 1–32 p.

[23]   Barrett E, McCreesh K, Lewis J. Reliability and validity of non-radiographic methods of thoracic kyphosis measurement: A systematic review. Manual Therapy. 2014. p. 10–7.

[24]   Hart DL, Rose SJ. Reliability of a noninvasive method for measuring the lumbar curve*. J Orthop Sports Phys Ther. 1986;8(4):180–4.

[25]   de Oliveira TS, Candotti CT, La Torre M, Pelinson PPT, Furlanetto TS, Kutchak FM, et al. Validity and Reproducibility of the Measurements Obtained Using the Flexicurve Instrument to Evaluate the Angles of Thoracic and Lumbar Curvatures of the Spine in the Sagittal Plane. Rehabil Res Pract. 2012;2012(August):1–9.

[26]   Eum R LS. Is kyphosis related to mobility, balance, and disability? Am J Phys Med Rehabil. 2013;92(11):980–9.

[27]   Ensrud KE, Black DM, Harris F, Ettinger B, Cummings SR. Correlates of kyphosis in older women. The Fracture Intervention Trial Research Group. J Am Geriatr Soc. 1997;45(6):682–7.

[28]   Harlick JC, Milosavljevic S, Milburn PD. Palpation identification of spinous processes in the lumbar spine. Man Ther. 2007;12(1):56–62.

[29]   McKenzie AM, Taylor NF. Can physiotherapists locate lumbar spinal levels by palpation? Physiotherapy. 1997;83(5):235–9.

[30]   Bonnet V, Yamaguchi T, Dupeyron A, Andary S, Seilles A, Fraisse P, et al. Automatic estimate of back anatomical landmarks and 3D spine curve from a Kinect sensor. Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatronics. 2016;2016-July:924–9.

Table 1. Descriptive statistics and reliability of the thoracic spine

	THORACIC REGION								
	Investigator 1			Investigator 2			Both investigators		
	Image 1	*Image 2*	*Image 3*	*Image 4*	*Image 5*	*Image 6*	*1st images pooled*	*2nd images pooled*	*3rd images pooled*
n	37	36	37	37	36	36	74	72	73
Mean	10.6	10.26ɫ	10.12	10.62	10.45	10.33	10.61‡	10.35	10.22
SD	3.19	3.09	3.19	3.1	3.18	3.23	3.17	3.16	3.24
Range	4.76-17.75	4.47-16.98	4.36-17.46	5.06-17.52	4.45-18.32	4.55-18.09	4.76-17.75	4.45-18.32	4.36-18.09
SEM	0.54	0.53	0.54	0.53	0.54	0.55	0.54	0.54	0.55
MDD	1.50	1.47	1.50	1.47	1.50	1.52	1.50	1.50	1.52
ICC coefficient (95% CI)	Intra-rater 0.960 (0.926, 0.979)			Intra-rater 0.975 (0.956, 0.987)			Inter-rater 0.971 (0.954, 0.984)		

ɫ significant difference from mean of Image 1 (p=0.009) and mean of Image 4 (p=0.001)
‡ significant difference from 2nd pooled images (p<0.001) and 3rd pooled images (p<0.001)

Table 2. Descriptive statistics and reliability of the lumbar spine

	LUMBAR REGION								
	Investigator 1			Investigator 2			Both investigators		
	Image 1	*Image 2*	*Image 3*	*Image 4*	*Image 5*	*Image 6*	*1st images pooled*	*2nd images pooled*	*3rd images pooled*
n	37	36	37	37	36	36	74	72	73
Mean	9.74	9.55	9.51	9.29	9.37	9.51	9.52	9.46	9.51
SD	4.23	4.07	4.08	4.19	4.34	4.34	4.19	4.18	4.18
Range	1.77-21.68	1.74-20.47	0.95-19.59	1.90-20.85	1.43-21.22	0.93-19.34	1.77-21.68	1.43-21.22	0.93-19.59
SEM	0.72	0.69	0.69	0.71	0.74	0.74	0.71	0.71	0.71
MDD	2.00	1.91	1.91	1.97	2.05	2.05	1.97	1.97	1.97
ICC coefficient (95% CI)	Intra-rater 0.973 (0.954, 0.985)			Intra-rater 0.977 (0.960, 0.988)			Inter-rater 0.971 (0.954, 0.984)		

8. FIGURE CAPTIONS

Figure 1. (a) Reconstructed 3D image using MeshLab software from original viewpoint of the sensor which was placed behind the participant perpendicular to the coronal plane. (b) Same image rotated to partially view the sagittal plane of the spine. (c) Schematic in pure sagittal view of anatomical markers used to calculate kyphosis index and lordosis index of the spine; e.g. kyphosis index = (height/length) x 100.

Figure 2. This schematic shows the Kinect sensor set up for image capture. The participants stood parallel to the face of the Kinect sensor, with their back exposed. The Kinect sensor was mounted onto a tripod and connected to a laptop.

Figure 3. After stratifying repeated measures by age and by the degree of thoracic kyphosis (hyperkyphosis ≥ 13.1), the hyperkyphotic group and older adults showed no significant within group differences. With statistical differences denoted by the asterisk signs: adults <65 y.o. showed significant differences between captures 1 and 2 ($p=0.045$) and captures 2 and 4 ($p=0.007$); the normal kyphosis group showed difference between the 1st and 2nd captures ($p=0.003$).
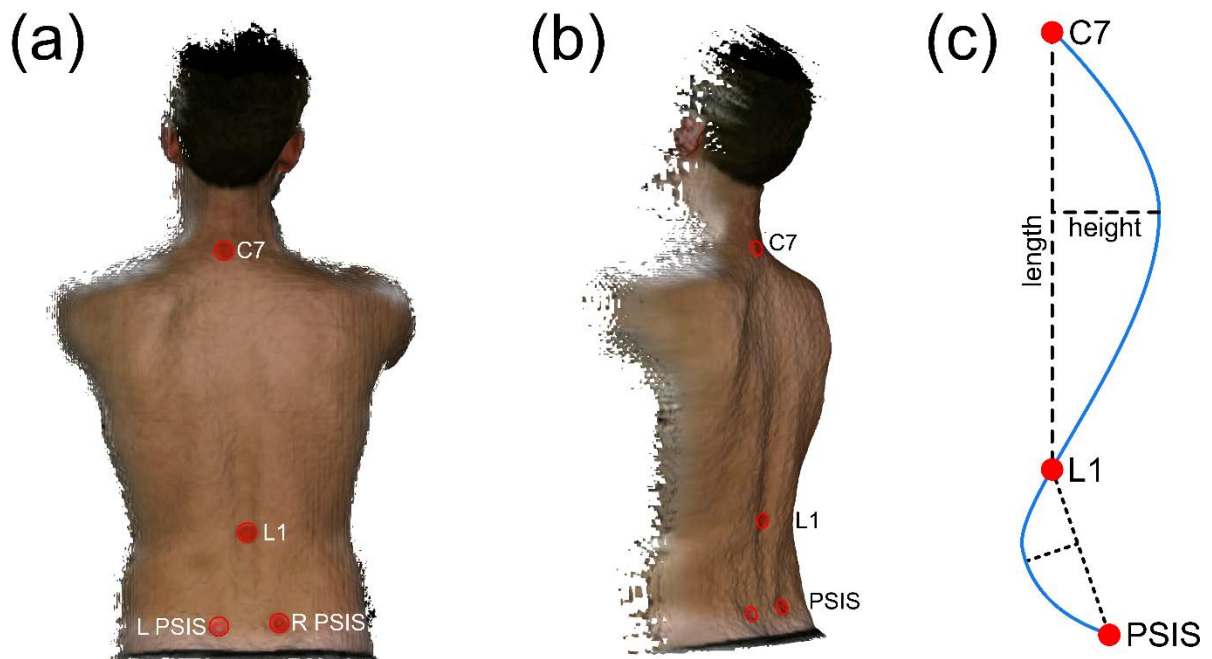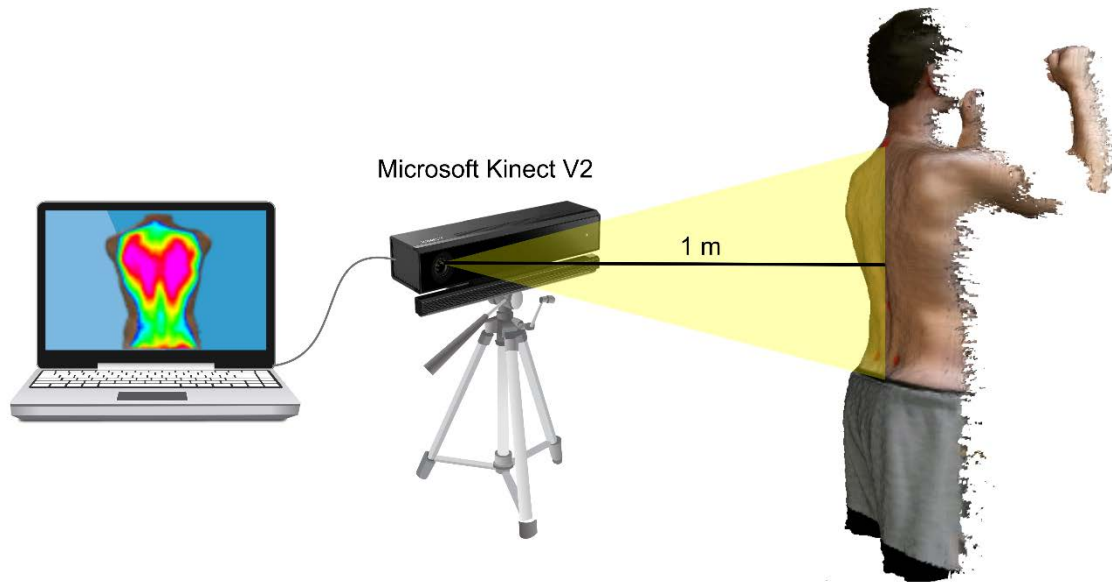
Figure 1.

Figure 2.

Microsoft Kinect V2

1 m

Figure 3.

Repeated measures by sub groups