January 2019

The Reliability Of Indirect Assessments Of Standing Forward Head Posture In Asymptomatic Adults

Bailey Strand

THE RELIABILITY OF INDIRECT ASSESSMENTS OF STANDING FORWARD HEAD POSTURE IN ASYMPTOMATIC ADULTS

by

Bailey Renae Strand
Bachelor of Science, University of North Dakota, 2018

A Thesis
Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

May
2019

PERMISSION

Title	The Reliability of Indirect Assessments of Standing Forward Head Posture in Asymptomatic Adults
Department	Kinesiology and Public Health Education
Degree	Master of Science

In presenting this document in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her/his absence, by the Chairperson of the department or the Dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Name Bailey Strand
Date 5/2/2019

TABLE OF CONTENTS

LIST OF TABLES

Table Page

ACKNOWLEDGMENTS

I would like to express my thanks to my committee members for their guidance and patience; without it I would not have been able to accomplish this. I couldn't be more grateful for their support during my time in the master's program at the University of North Dakota.

ABSTRACT

Background: Measurement reliability has important decision-making implications for physical therapists and researchers when assessing individuals. Given that physical therapists often visually assess forward head posture (FHP) to guide treatment, the aims of this study were to quantify: (a) the reliability of indirect (visual) assessments of standing FHP in asymptomatic adults, and (b) the magnitude of the biological (postural) and technical errors involved.

Methods: A reliability analysis of two indirect assessments (visual assessments of real and 3D body scanned people) of FHP was undertaken. The sample comprised 10 physical therapists and 50 asymptomatic participants. Participants were scanned using the *Vitus Smart* 3D whole body scanner and visually assessed for FHP by the physical therapists. One week later, the physical therapists visually assessed the scanned images of the participants plus 15 duplicates, and two weeks later, the physical therapists and participants again presented with physical therapists repeating their visual assessments. Reliability (both intra- and inter- rater) of indirect assessments was determined by Cohen's Kappa (κ). Total error was estimated as the unexplained error in repeated indirect assessments of real people (i.e., 1 minus the intra-rater κ for real people); technical error as the unexplained error in repeated visual assessments of scanned people (i.e., 1 minus the intra-rater κ for scanned people); and biological error as the difference between the total and technical errors.

Results: The intra-rater reliability of indirect assessments of real and scanned people was moderate (κ [95%CI]: real, 0.45 [0.34, 0.56]); scan, 0.46 [0.39, 0.53]), while the inter-rater reliability was slight (κ [95%CI]: real, −0.02 [−0.09, 0.05]; scan, 0.09 [0.06, 0.12]). Nearly all of the FHP error was due to technical error.

Conclusions: The intra- and inter- rater reliability of indirect assessments of standing FHP was moderate and slight, respectively. It appears that nearly all of the error in indirect assessments of FHP is due to technical error, highlighting that efforts to improve reliability should focus on minimizing technical errors.

Keywords: visual assessments; repeatability; biological error; technical error; 3D scanner

CHAPTER I

INTRODUCTION

Posture is usually measured and assessed by practitioners such as physical therapists, occupational therapists, orthopedic surgeons, chiropractors, and other exercise professionals. Typically, posture has been assessed using indirect (visual) observations; meaning assessments are made by visually observing patients statically (i.e., when motionless) or dynamically (i.e., in motion). The posture of the patients is often described using qualitative thresholds of magnitude (e.g., mild, moderate, or severe), although these thresholds are poorly defined and operationalized (1,2) . Because posture is meaningful linked with pain/dysfunction longitudinally (2,3), physical therapists often use these indirect assessments to help guide treatment.

The cost of back pain to the U.S. economy has been estimated at US $100–200 billion, due to medical expenses, productivity loss and absenteeism (4). In fact, living with back pain is one of the leading causes of living with a disability in the United States (5). Neck problems are also becoming more prevalent, with neck and/or back pain now the second leading cause of disability in the United States (6). Shin et al. (7) reported that people with more prominent forward head postures (FHP) tend to experience more neck pain. Falla et al. (3) found that people with neck pain had more difficulty maintaining their head on neck posture (with the magnitude of FHP increasing over time) after as little as 10 minutes of playing video games relative to their asymptomatic peers. FHP has also been meaningful linked with a more marked cervical lordosis and cervicogenic headaches (8).

Reliability has important decision-making implications for physical therapists and

researchers for: (a) assessing an individual with a single measurement or repeated measurements, (b) estimating the extent of individual responses to treatment, (c) comparing the reliability of different tests/devices or measurers, and (d) sample size estimation in experimental or longitudinal studies (11). Physical therapists currently use spinal posture examinations including patient history and visual assessments in their clinical practice before treatment is prescribed. Unfortunately, no data are available on the reliability of indirect (visual) assessments of FHP. Furthermore, because the error in postural measurement is due to both technical error (i.e., the error in the measurement process) and biological (postural) error (i.e., the within-subject error), knowledge of the magnitude of each error is critical to improving measurement precision. Potential technical errors include landmarking, equipment, calibration or technique, and potential biological errors include growth, physical activity, or diurnal variability. Unfortunately, no information is available regarding the magnitude of technical and biological errors associated with visual assessments of standing FHP. The aims of this study were to quantify: (a) the reliability of indirect (visual) assessments of standing FHP, and (b) the magnitude of the associated biological and technical errors.

CHAPTER II

METHODS

Participants and Sampling

Fifty participants (18 females, 32 males; mean±SD: age, 27±12 years; height, 174±11 cm; mass, 72±14 kg) and 10 registered physical therapists (mean±SD: age, 38±11 years; clinical experience, 16±12 years) were recruited by convenience. To be eligible, participants had to be able to stand unsupported for 15 seconds on a raised platform, and present asymptomatic for back and neck pain. The Human Research Ethics Committee of the University of South Australia and the Institutional Review Board of the University of North Dakota approved this study.

Procedures

Upon arrival, participants completed a demographic questionnaire and were then measured for height (cm) and mass (kg) using a stadiometer and digital weighing scale, respectively. They then changed into the appropriate undergarments, which included form fitting briefs for men and form fitting briefs and sports bras for women, and were scanned using a *Vitus Smart* 3D whole-body scanner (Human Solutions GmbH, Kaiserslautern, Germany). Participants were scanned in their "normal" standing posture using the procedures of Schranz et al (14), where they took a few steps in place and then moved their head and shoulders around to find a comfortable standing position. Once the scan was complete, the physical therapists entered the room one at a time to visually assess the FHP of each participant. The physical therapists were allowed to move around

the participants and palpate as necessary, but they could not ask the participants to move from their normal standing posture. FHP was assessed as the degree to which the head is anteriorly/posteriorly positioned relative to 'normal', using a spectrum of postural deviations ranging from normal, mild, moderate, to severe. While FHP was defined to the physical therapists, the grading criteria were not. Two weeks later, and at the same time of day, the physical therapists re-assessed the participants using the same procedure, with participants randomly presented.

Movie (.avi) files of the 3D scan of each participant were generated and visually assessed by the physical therapists one week later. The order of the movie files was randomized with 15 randomly selected duplicate scans included to assess the intra-rater reliability of indirect visual assessments of scanned people. The physical therapists were aware that duplicate movie files were added, but were asked to assess the FHP of each scan independently using the same grading criteria as described above.

Statistical Analyses

Both intra-rater (the same physical therapists rating the same participants on separate occasions) and inter-rater (different physical therapists rating the same participants on a single occasion) reliability of indirect (ordinal) assessments of real and scanned people were quantified using Cohen's Kappa (κ). Kappa coefficients were qualitatively interpreted using Landis and Koch (15) scale of magnitudes, where: <0.00 indicated poor agreement; 0.00–0.20 indicated slight agreement; 0.21–0.40 indicated fair agreement; 0.41–0.60 indicated moderate agreement; 0.61–0.80 indicated substantial agreement; and 0.81–1.00 indicated almost perfect agreement. Intra- and inter-rater

reliability were also examined visually by generating frequency distributions of absolute intra- and inter-rater differences.

Assuming that the biological and technical errors were independent, and that Kappa coefficients behave similarly to Pearson's correlation coefficients, then the total error was estimated as the unexplained error in repeated indirect assessments of real people (i.e., 1 minus the intra-rater kappa κ coefficient for real people) and technical error as the unexplained error in repeated indirect assessments of scanned people (i.e., 1 minus the intra-rater kappa coefficient κ for scanned people). Biological error (i.e., the within-subject error free from technical error) was therefore estimated as the difference between the total and technical errors.

CHAPTER III

RESULTS

Intra-and Inter-rater Reliability of Two Indirect (Visual) Assessments

The intra- and inter- rater reliability of indirect assessments of FHP in real people was moderate (κ [95% CI]: 0.45 [0.34, 0.56]) and slight (κ [95% CI]: −0.02 [−0.09, 0.05]), respectively. Similarly, the intra- and inter- rater reliability of indirect assessments of FHP in scanned people was moderate (κ [95% CI]: 0.46 [0.39, 0.53]) and slight (κ [95% CI]: 0.09 [0.06, 0.12]), respectively.

The most common intra-rater difference of indirect assessments (real and scanned) of FHP was zero (i.e., the same visual assessment was given by the same physical therapists for both the test and retest), with 98–99% within one point (Table 1). The most common inter-rater difference of indirect assessments of FHP was one (i.e., a 1-point rating difference between the test measures of two physical therapists) in real people and zero in scanned people, with 97–98% within one point (Table 1).

Table 1. Frequency Distribution of Intra- and Inter-rater Differences in Indirect Assessments of FHP in Asymptomatic Adults.

	|point difference|	**real vs. real**	**scan vs. scan**
Intra-rater	2	2	1
	1	25	25
	0	73	74
Inter-rater	2	3	2
	1	52	41
	0	45	57

Note: Intra- and inter-rater differences are expressed as absolute rating-point differences, with frequencies represented as percentages.

Measurement Error Due to Biological and Technical Errors

Most of the error in standing FHP was due to technical error. The total error was estimated as the unexplained error in repeated visual assessments of real people ($\therefore 1 - 0.45^2 = 0.80$ or 80%); technical error as the unexplained error in repeated visual assessments of scanned people ($\therefore 1 - 0.46^2 = 0.79$ or 79%); and biological error as the difference between the total and technical errors ($\therefore 0.80 - 0.79 = 0.01$ or 1%).

CHAPTER IV

DISCUSSION

This study examined the reliability of indirect assessments of standing FHP made by physical therapists. The key findings were: (a) intra-rater reliability of indirect assessments was moderate and better than inter-rater reliability; (b) inter-rater reliability of indirect assessments was slight and no better than chance alone in the case of real people; and (c) nearly all of the errors in indirect assessments was due to technical error.

Implications

These findings have several important implications for physical therapists and researchers as indirect assessments of real people are regularly used to guide treatment and diagnosis. Given that nearly all of the error in indirect FHP assessments was due to technical error, efforts to improve measurement precision should therefore aim to reduce technical errors in the visual assessment process. This may include (but is not restricted to) the operationalization of measurement definitions and grading criteria, strict adherence to measurement protocols, and extensive postural training for physical therapists.

Comparisons with Other Studies

Reliability data on visual assessments of standing posture are scant. In a sample of 28 physical therapists, chiropractors, physiatrists, rheumatologists, and orthopedic surgeons who assessed the cervical and lumbar lordosis of photographed participants (with and without back pain), Fedorak et al. (12) reported the collective intra-rater

reliability (qualitatively interpreted using Landis and Koch's [15] thresholds) as moderate (κ [(95% CI]: 0.50 [0.02, 0.98]) and the inter-rater reliability as slight (κ [(95% CI]: 0.16 [0.00, 0.48]). The intra- and inter-rater reliability of physical therapists (n=7) was moderate (κ [(95% CI]: 0.49 [0.09, 0.89]) and fair (κ [(95% CI]: 0.29 [0.00, 0.46]), respectively (12). Similarly, using the same physical therapist and asymptomatic participants sample as the present study, Larson (13) reported that the intra- and inter-rater reliability of indirect assessments of lumbar lordosis by 10 physical therapists was fair-to-moderate and slight, respectively. Larson (13) also estimated that most (~80%) of the total intra-rater error in visual assessments of lumbar lordosis was due to technical error, which is somewhat less than that observed in this study.

Strengths and Limitations

This is the first study to examine the reliability of visual assessments of standing FHP made by physical therapists. The use of two different visual assessment types (visual assessments of real and 3D body scanned people) allowed for both the biological and technical errors associated with FHP to be estimated. This study also used a 2-week test-retest measurement interval which reflects clinical practice reasonably well. However, the study design (where people were tested and retested at the same time of day and on the same day of the week) may not reflect clinical practice well and likely minimized diurnal variation (i.e., biological error); meaning that the reliability estimates reported in this study likely reflect a best-case scenario.

It is also important to remember that the results of this study reflect only reliability estimates for FHP in asymptomatic adults and are not necessarily generalizable to other postural aspects or symptomatic adults. The convenience sample strategy almost

certainly resulted in a sample unrepresentative of that typically observed by physical therapists. The homogenous group of physical therapists (who were trained at a single institution) may not have been representative of all physical therapists, and a more heterogeneous group of physical therapists may have increased the variability in reliability estimates.

Conclusion

This study reported that the intra- and inter-rater reliability of indirect assessments of standing FHP was moderate and slight, respectively. These findings have important decision-making implications when assessing single and change measurements in individuals. The other key finding that nearly all of the error in indirect assessments was due to technical error, highlights that efforts to improve measurement precision should focus on minimizing technical error of measurement.

REFERENCES

1. Griegel-Morris P, Larson K, Mueller-Klaus K, Oatis CA. Incidence of common postural abnormalities in the cervical, shoulder, and thoracic regions and their association with pain in two age groups of healthy subjects. Phys Ther. 1992;72(6):425–31.

2. Silva AG, Punt TD, Sharples P, Vilas-Boas J, Johnson MI. Head Posture and Neck Pain of Chronic Nontraumatic Origin: A Comparison Between Patients and Pain-Free Persons. Archives of Physical Medicine & Rehabilitation. 2009;90(4):669–74.

3. Falla D, Jull G, Russell T, Vicenzino B, Hodges P. Effect of neck exercise on sitting posture in patients with chronic neck pain. Phys Ther. 2007;87(4):408–17.

4. Ma VY, Chan L, Carruthers KJ. Incidence, Prevalence, Costs, and Impact on Disability of Common Conditions Requiring Rehabilitation in the United States: Stroke, Spinal Cord Injury, Traumatic Brain Injury, Multiple Sclerosis, Osteoarthritis, Rheumatoid Arthritis, Limb Loss, and Back Pain. Archives of Physical Medicine & Rehabilitation. 2014;95(5):995.e1.

5. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. 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. 2017;390(10100):1211–59.

6. Strine TW, Hootman JM. US national prevalence and correlates of low back and neck pain among adults. Arthritis Care & Research. 2007;57(4):656–65.

7. Shin YJ, Kim WH, Kim SG. Correlations among visual analogue scale, neck disability index, shoulder joint range of motion, and muscle strength in young women with forward head posture. Journal of exercise rehabilitation. 2017;13(4):413.

8. Farmer PK, Snodgrass SJ, Buxton AJ, Rivett DA. An Investigation of Cervical Spinal Posture in Cervicogenic Headache. Phys Ther. 2015;95(2):212–22.

9. Hopkins WG. Measures of reliability in sports medicine and science. Sports medicine. 2000;30(1):1–15.

10. Schranz N, Tomkinson G, Olds T, Daniell N. Three-dimensional anthropometric analysis: Differences between elite Australian rowers and the general population. J Sports Sci. 2010;28(5):459–69.

11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977:159–74.

12. Fedorak C, Ashworth N, Marshall J, Paull H. Reliability of the visual assessment of cervical and lumbar lordosis: how good are we? Spine. 2003;28(16):1857–9.

13. Larson NA. A Comparison of Direct and Indirect Assessments of Standing Lumbar Lordosis in Asymptomatic Adults [dissertation]. ProQuest Dissertations Publishing; 2018.