J. Phys. Ther. Sci. 28: 3398–3402, 2016

The Journal of Physical Therapy Science

Original Article

Inter- and intra-rater agreement of static posture analysis using a mobile application

David M. Boland, DPT[1)], Eric V. Neufeld, BS[1)], Jack Ruddell, BS[1)], Brett A. Dolezal, PhD[1)*], Christopher B. Cooper, MD[1)]

[1)] *Exercise Physiology Research Laboratory, Departments of Medicine and Physiology, David Geffen School of Medicine at the University of California, Los Angeles: 10833 Le Conte Ave, Los Angeles, California, USA*

Abstract. [Purpose] To determine the intra- and inter-rater agreement of a mobile application, PostureScreen Mobile® (PSM), that assesses static standing posture. [Subjects and Methods] Three examiners with different levels of experience of assessing posture, one licensed physical therapist and two untrained undergraduate students, performed repeated postural assessments of 10 subjects, fully clothed or minimally clothed, using PSM on two nonconsecutive days. Anterior and right lateral images were captured and seventeen landmarks were identified on them. Intraclass correlation coefficients (ICCs) were calculated for each of 13 postural measures to evaluate inter-rater agreement on the first visit (fully or minimally clothed), as well as intra-rater agreement between the first and second visits (minimally clothed). [Results] Eleven postural measures were ultimately analyzed for inter- and intra-rater agreement. Inter-rater agreement was almost perfect (ICC≥0.81) for four measures and substantial (0.60<ICC≤0.80) for three measures during the fully clothed exam. During the minimally clothed exam, inter-rater agreement was almost perfect for four measures and substantial for four measures. Intra-rater agreement between two minimally clothed exams was almost perfect for two measures and substantial for five measures. [Conclusion] PSM is a widely available, inexpensive postural screening tool that requires little formal training. To maximize inter- and intra-rater agreement, postural screening using this mobile application should be conducted with subjects wearing minimal clothing. Assessing static standing posture via PSM gives repeatable measures for anatomical landmarks that were found to have substantial or almost perfect agreement. Our data also suggest that this technology may also be useful for diagnosing forward head posture.

Key words: Reliability, Posture, Mobile application

(This article was submitted Jul. 18, 2016, and was accepted Aug. 20, 2016)

INTRODUCTION

Postural anomalies are widespread and are exhibited by people of all demographics. While commonly attributed to lack of physical activity, poor posture is especially prominent in individuals who spend long periods of time sitting at a desk or computer, as is necessary with many jobs in developed countries[1–3)]. Poor posture is a major causative factor of many musculoskeletal problems, including joint stress, pain, and various diseases[4–11)]. Detecting irregularities in posture is an important component in understanding sources of pain and limited movement, and for subsequently developing mitigation strategies such as corrective exercises. Due to its dynamic nature, human posture is difficult to quantify. Objective analysis often utilizes elaborate technology, such as x-ray imaging and 3D motion capture, which provide a valid and reliable assessment; however, these capabilities are found only in specialized clinics[12)]. In an effort to address this limitation, portable device technology has been developed hand-in-hand with medical technology to provide non-invasive, user-friendly, and

*Corresponding author. Brett A. Dolezal (E-mail: BDolezal@mednet.ucla.edu)

affordable posture-assessing software designed for use by physical therapists, chiropractors, and other health and fitness professionals. PostureScreen Mobile® (PostureCo Inc., Trinity, FL, USA) is a novel, commercially sold photographic mobile application that enables the identification of deviations from the ideal standing posture. The PostureScreen Mobile® (PSM) application facilitates the assessment of posture in a variety of settings. While the convenience of this tool is evident, the information provided would only be useful if the measurements were repeatable between multiple visits and reproducible between multiple examiners. Therefore, the purpose of this study was to investigate the inter- and intra-rater agreement when using PSM, as well as understanding whether standard clothing can obscure the reliability of such measurements.

SUBJECTS AND METHODS

One experienced licensed physical therapist (a Doctor of Physical Therapy) and two inexperienced examiners (undergraduate students) performed repeated postural assessments of 10 subjects wearing both full and minimal clothing on two nonconsecutive days using PSM installed on a tablet (iPad; Apple Inc., Cupertino, CA, USA). An initial set of posture images was captured with subjects dressed in standard street clothing. A second set of images was captured on the same visit with subjects wearing minimal clothing (shorts for men, and shorts and a sports bra for women). A third set of images was captured 48 hours later in the minimally clothed condition. Each image was then analyzed separately by all three examiners. Subjects were recruited randomly from the UCLA community, based on availability, not on postural qualifications. This study was approved by the UCLA Institutional Review Board and all participants gave their written informed consent. Participants completed standard health, medical, and exercise history forms following enrollment and prior to data acquisition. Before each assessment, the room was set up in a standardized way to ensure uniform conditions throughout the test. Tape was placed on the floor to indicate where the subject should stand for the anterior picture. A second piece of tape was laid perpendicular to the first piece to indicate the positioning for the lateral picture. The tablet was placed on a stand exactly 10 feet away from the subject markers at a height of 4.5 feet to standardize the image angle. The height and weight of each subject was entered into PSM. Height was measured to the nearest 0.5 cm using a wall-mounted stadiometer, and weight was measured to the nearest 0.1 kg with an octopolar bioelectrical impedance scale (Biospace InBody R20; InBody Co. Ltd., Seoul, South Korea). Subjects were instructed to stand at the tape marks in a comfortable position with weight evenly distributed on both legs and to look at a marker placed on the wall in front of them at eye level. To help eliminate potential deviations from their routine posture, subjects were asked to stand on their right foot for a few seconds and then shift their balance to the left foot for a few seconds. Subjects then placed both feet on the ground and an image was captured using PSM and the camera function of the iPad. This process was repeated to obtain both an anterior and right lateral view of each subject. Three sets of postural images were taken of each subject over two visits. An initial set of images was captured with subjects dressed in standard street clothing with their shoes removed. A second set of images was captured on the same visit with subjects wearing minimal clothing (shorts for men, and shorts and a sports bra for women) also with their shoes removed. An identical third examination was conducted 48 hours later in the minimally clothed condition. Subjects were asked to refrain from strenuous physical activity between visits in order to reduce external factors influencing postural variation. Three examiners analyzed each set of data by following the PSM prompts to select 17 specific anatomical landmarks. Twelve landmarks were placed on the anterior view image at the following locations: the right pupil, left pupil, midpoint between the nose and upper lip, top of the right acromioclavicular (AC) joint, episternal notch, top of the left AC joint, right lateral rib T8 level, left lateral rib T8 level, right anterior superior iliac spine (ASIS), left ASIS, center of the right talus, and center of the left talus. Five landmarks were placed on the lateral view image at the following locations: the external auditory meatus, center of the shoulder at the cervicothoracic junction, greater trochanter, center of the tibiofemoral joint, and center of the malleolus. PSM then calculated the following 13 quantitative data points using proprietary algorithms: *head shift (lateral), head shift (longitudinal), head tilt, shoulder shift (lateral), shoulder shift (longitudinal), shoulder tilt, ribcage shift, hip shift (lateral), hip shift (longitudinal), hip tilt, head weight, effective head weight,* defined as how heavy the head feels to its supporting structures based on any postural deviations, and *knee shift*. Intraclass correlation coefficients (ICCs) were derived by calculating between-subject variance as a fraction of the total variance of each of the 13 postural measurements made by all three examiners to determine inter-rater agreement. Inter-rater agreement was calculated separately for the fully clothed and minimally clothed conditions. Additionally, ICCs were calculated for the measurements of the two minimally clothed examinations conducted by the experienced examiner on separate occasions to determine the intra-rater agreement of the measurements. ICCs greater than 0.60 but less than or equal to 0.80 ($0.60<ICC\leq0.80$) were considered to indicate "substantial" agreement, while ICCs equal to or greater than 0.81 ($ICC\geq0.81$) were considered to indicate "almost perfect" agreement[13]. The postural measures of *head weight* and *effective head weight* were omitted from the analysis because these were considered to be affected more by body weight than by the placement of landmarks by the raters.

RESULTS

Ten subjects (five females) completed the study (age 21 ± 1 years; height 172 ± 11 cm; weight 70 ± 14 kg; mean ± SD). Inter-rater agreement for both the fully and minimally clothed conditions and the intra-rater agreement for the minimally clothed condition are reported in Table 1. Inter-rater agreement of the fully clothed exam was at least substantial ($ICC>0.60$)

Table 1. Levels of agreement of repeated measurements

Postural measurements	Intraclass correlation coefficients†		
	Fully clothed (3 Examiners)	Minimally clothed (3 Examiners)	Minimally clothed (2 Visits)
Head shift lateral	0.87*	0.83*	0.46
Head shift AP	0.55	0.72	0.73
Head tilt	0.91*	0.86*	0.78
Shoulder shift lateral	0.67	0.64	0.33
Shoulder shift AP	0.75	0.72	0.68
Shoulder tilt	0.58	0.88*	0.78
Hip shift lateral	0.95*	0.93*	0.76
Hip shift AP	0.64	0.55	0.84*
Hip tilt	0.10	0.26	0.59
Head weight	1.00	1.00	0.90
Effective head weight	0.98	0.96	0.84
Rib cage shift	0.19	0.58	0.59
Knee shift	0.88*	0.75	0.86*

†Data are intraclass correlation coefficients (ICCs) where 0.40<ICC≤0.60 indicate moderate agreement, 0.60<ICC≤0.80 indicate substantial agreement and ICC≥0.81 indicate almost perfect agreement. *Head weight* and *effective head weight* were omitted from the analysis because they were considered to be affected more by body weight than by the placement of landmarks by the raters.
*ICC≥0.81

for seven measures and specifically almost perfect (ICC≥0.81) for: *head shift lateral*, *head tilt, hip shift lateral*, and *knee shift.* Inter-rater agreement of the minimally clothed exam was almost perfect for the previously identified measures minus *knee shift* but plus *shoulder tilt.* In the minimally clothed condition, only *hip tilt* failed to reach at least moderate agreement (0.40<ICC≤0.60). Intra-rater agreement (between visits) for the minimally clothed exams was at least substantial (ICC>0.60) for seven measures and specifically almost perfect for: *hip shift anterior/posterior (hip shift AP)* and *knee shift*. There were large differences between inter- and intra-rater agreement for *head shift lateral* (0.826 vs. 0.457) and *shoulder shift lateral* (0.642 vs. 0.332), with both failing to have substantial intra-rater agreement despite having met or exceeded this level of agreement in the inter-rater analysis. In contrast, *hip tilt* possessed a low inter-rater agreement (ICC=0.256) but registered a much higher intra-rater agreement (ICC=0.588).

DISCUSSION

To the best of our knowledge, this is the first published study that has examined both the inter- and intra-rater agreement of the mobile application, PostureScreen Mobile®, for screening static posture. Acceptable levels of agreement were found among the measurements of three different examiners of varying experience. Overall inter-rater agreement was substantial whether the examination was conducted on a fully clothed or minimally clothed subject, with 7 of the 11 measures exceeding an ICC of 0.60 in the fully clothed condition. In the fully clothed condition, the examiners' evaluation of head position in the frontal plane (*head shift lateral, head tilt*), weight shift in the frontal plane (*hip shift lateral*), and one of three measures relating to weight shift in the sagittal plane (*knee shift*) using PSM were especially reproducible (ICC≥0.81). Unsurprisingly, removing the shirt (i.e., creating a minimally clothed condition) improved the agreement of measures relying on upper body landmarks on the acromioclavicular joints and rib cage. For example, the *rib cage shift* measure improved from slight agreement (0.01<ICC≤0.20) to moderate agreement. Importantly, the *head shift anterior/posterior (head shift AP)* measure showed substantial agreement (ICC=0.724) in the minimally clothed condition. *Head shift AP* may be the most clinically relevant measure provided by PSM as it enables a quantitative diagnosis of forward head posture. Forward head posture is a common postural deviation in developed countries and it is believed to be a causative factor of many musculoskeletal diseases, including migraine and chronic tension headaches[10, 11]. Traditional means of diagnosing this postural condition involve manually determining the angle between the horizontal and a line running from the C7 vertebra to the tragus of the ear from a sagittal photograph[14]. A quick, yet reliable, means of determining this angle could greatly simplify management of cervical-spine-related disorders. Posture measures that failed to reach substantial agreement in the minimally clothed condition tended to rely on anatomical landmarks that are either difficult to observe (e.g., anterior superior iliac spine, greater trochanter) or non-specific (e.g., rib cage). *Hip tilt* stood out as having notably poor inter-rater agreement. Even among highly experienced examiners, identification of pelvic deviation via hands-on palpation can prove challenging and inexact[15–17]. Therefore, it is unsurprising that attempts to identify this phenomenon via photographic observation were less

reliable. However, the strong intra-rater agreement of *hip tilt* could indicate that this measure, unlike the other 10, requires an experienced examiner, meaning the undergraduate examiners were not able to reliably estimate the location of the landmarks. However, since we did not assess intra-rater agreement for the inexperienced examiners, this cannot be stated conclusively. Alternatively, it is possible that each examiner has their own consistent, even if inaccurate, idea of where the ASIS is located when analyzing the photographs. In addition to evaluating the performance of the examiner, intra-rater agreement could be considered a partial measure of postural consistency within participants, at least among measures with strong inter-rater agreement. Between-visit variations suggest certain aspects of posture may be highly inconsistent within individuals on a day-to-day basis. All lower body weight shift measures (*hip shift AP, hip shift lateral,* and *knee shift*) appear to have remained fairly consistent across visits, possibly because we asked subjects to unweight each foot prior to each exam, thus standardizing the lower extremity weight shift preference for each visit. In contrast, none of the weight shift measures above the waist reached almost perfect agreement, indicating the variable nature of weight shift preference from moment-to-moment even with a stable lower body. A previous study by Hopkins[18], tested the validity and intra-rater agreement of PSM by comparing it with the Vicon 3D motion analysis system (Vicon Motion Systems, Oxford, United Kingdom) as the gold-standard measure[19]. This investigator found that the shift and tilt measurements using PSM were repeatedly higher than those measured by the Vicon system in the frontal view. In the lateral view, the results reversed—PSM tended to underestimate shift and tilt measurements. These significant differences in variation persisted with and without the use of anatomical markers. Unlike this study, Hopkins found that the trial-to-trial variance with PSM exhibited low intra-rater agreement. Other software for the quantification of posture exists. One such posture assessment tool with 29 measurements was found to be highly reliable in a study by Ferreira et al.[20], with just 14.0% of measurements for inter-rater agreement and 19.3% for intra-rater agreement failing to reach an ICC of at least 0.70. While these findings were better than those of our present study of PSM (with 30.7% of inter-rater and 38.5% of intra-rater measures with ICCs lower than 0.700), there is a trade-off between ease-of-use and time spent performing analysis. The study of Ferreira et al.[20] involved five fully trained physical therapists and the software required identification of 32 anatomic points. It is noteworthy that PSM performed reliably while requiring identification of only 17 points and when employed by inexperienced examiners. One important finding of this study is the inconsistency of posture between subject visits. There are several factors that may influence day-to-day posture including energy level, recent exertion, and involuntary compensation to alleviate temporary discomfort. In order to minimize these confounding factors, subjects were instructed not to perform any heavy lifting between visits. Nevertheless, any of these factors may have led to lower ICCs that underestimate the true inter-rater agreement. Future studies should build on the work of Raine et al.[14] to identify acceptable ranges for each postural measure to allow categorization (e.g., normal, at risk, abnormal), which would likely improve the clinical utility of this device. PSM appears to be a reliable and user-friendly tool for characterizing static standing posture. To maximize agreement, assessments should be conducted with the subject wearing minimal clothing so as to better identify anatomical landmarks. Our results indicate that usage of PSM requires minimal formal training since two inexperienced examiners reliably matched the trained physical therapist in the determination of most measures. Moreover, PSM may provide a cheaper, more feasible alternative to more advanced methods for initial patient screenings or for situations that do not require high precision. For assessments where a very high degree of accuracy is required, a more robust software package or motion analysis system may be preferred.

Conflicts of interest

This study did not receive any funding, and authors have no conflicts of interest to declare.

ACKNOWLEDGEMENTS

The authors wish to thank Jonathan Lee and Andrew Chang for volunteering to serve as our two inexperienced examiners. Data analysis was completed by Sitaram Vangala, M.S., from the UCLA Department of Medicine Statistics Core.

REFERENCES

1) Burdorf A, Naaktgeboren B, de Groot HC: Occupational risk factors for low back pain among sedentary workers. J Occup Med, 1993, 35: 1213–1220. [Medline]
2) Jin K, Sorock GS, Courtney T, et al.: Risk factors for work-related low back pain in the People's Republic of China. Int J Occup Environ Health, 2000, 6: 26–33. [Medline] [CrossRef]
3) Hamilton MT, Healy GN, Dunstan DW, et al.: Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior. Curr Cardiovasc Risk Rep, 2008, 2: 292–298. [Medline] [CrossRef]
4) Poca MA, Sahuquillo J, Topczewski T, et al.: Posture-induced changes in intracranial pressure: a comparative study in patients with and without a cerebrospinal fluid block at the craniovertebral junction. Neurosurgery, 2006, 58: 899–906, discussion 899–906. [Medline] [CrossRef]
5) Adams MA, Hutton WC: The effect of posture on the lumbar spine. J Bone Joint Surg Br, 1985, 67: 625–629. [Medline]
6) Tattersall R, Walshaw MJ: Posture and cystic fibrosis. J R Soc Med, 2003, 96: 18–22. [Medline]
7) Kado DM, Huang MH, Karlamangla AS, et al.: Hyperkyphotic posture predicts mortality in older community-dwelling men and women: a prospective study.

J Am Geriatr Soc, 2004, 52: 1662–1667. [Medline] [CrossRef]

8) Burdorf A: Sources of variance in exposure to postural load on the back in occupational groups. Scand J Work Environ Health, 1992, 18: 361–367. [Medline] [CrossRef]
9) Sjaastad O, Saunte C, Hovdahl H, et al.: “Cervicogenic” headache. An hypothesis. Cephalalgia, 1983, 3: 249–256. [Medline] [CrossRef]
10) Fernández-de-las-Peñas C, Alonso-Blanco C, Cuadrado ML, et al.: Trigger points in the suboccipital muscles and forward head posture in tension-type headache. Headache, 2006, 46: 454–460. [Medline] [CrossRef]
11) Fernández-de-Las-Peñas C, Cuadrado ML, Pareja JA: Myofascial trigger points, neck mobility and forward head posture in unilateral migraine. Cephalalgia, 2006, 26: 1061–1070. [Medline] [CrossRef]
12) Fortin C, Feldman DE, Cheriet F, et al.: Clinical methods for quantifying body segment posture: a literature review. Disabil Rehabil, 2011, 33: 367–383. [Medline] [CrossRef]
13) Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics, 1977, 33: 159–174. [Medline] [CrossRef]
14) Raine S, Twomey LT: Head and shoulder posture variations in 160 asymptomatic women and men. Arch Phys Med Rehabil, 1997, 78: 1215–1223. [Medline] [CrossRef]
15) Gajdosik R, Simpson R, Smith R, et al.: Pelvic tilt. Intratester reliability of measuring the standing position and range of motion. Phys Ther, 1985, 65: 169–174. [Medline]
16) Alviso DJ, Dong GT, Lentell GL: Intertester reliability for measuring pelvic tilt in standing. Phys Ther, 1988, 68: 1347–1351. [Medline]
17) Burdett RG, Brown KE, Fall MP: Reliability and validity of four instruments for measuring lumbar spine and pelvic positions. Phys Ther, 1986, 66: 677–684. [Medline]
18) Hopkins BC: Validity of PostureScreen Mobile® in the measurement of standing posture [Master’s thesis]. Provo: Brigham Young University, 2014.
19) Ehara Y, Fujimoto H, Miyazaki S, et al.: Comparison of the performance of 3D camera systems. Gait Posture, 1995, 3: 166–169. [CrossRef]
20) Ferreira EA, Duarte M, Maldonado EP, et al.: Postural assessment software (PAS/SAPO): validation and reliabiliy. Clinics (Sao Paulo), 2010, 65: 675–681. [Medline] [CrossRef]