Journal of Back and Musculoskeletal Rehabilitation 22 (2009) 173–178
DOI 10.3233/BMR-2009-0231
IOS Press

Computerized back postural assessment in physiotherapy practice: Intra-rater and inter-rater reliability of the MIDAS system

R.T. McAlpine[a], J.A. Bettany-Saltikov[b,*] and J.G. Warren[c]
[a]*Bishop Auckland General Hospital, Cockton Hill Road, Bishop Auckland, UK*
[b]*University of Teesside, School of Health and Social Care, Middlesbrough, UK*
[c]*School of Computing, University of Teesside, Middlesbrough, UK*

Abstract. *Background and purpose:* Assessment of spinal posture during physiotherapy practice is routine, yet few objective measures exist to this end. The Middlesbrough Integrated Digital Assessment System (MIDAS) is a low cost portable system able to record 3D information on posture. The purpose of this study was to assess both the intra-rater and inter-rater reliability of the MIDAS system.

Methods: Twenty-five healthy subjects were recruited. A repeated measures design was used to record fifteen pre-palpated landmarks on the back of each subject. To limit the sources of variability, the principal researcher palpated the landmarks for each subject. Each of three raters took two measurements on each subject in a standardized upright posture. X (medio-lateral), Y (antero-posterior) and Z (height) landmark positions were recorded via a computer interface.

Results: Both intra-rater agreement (mean ICCs – rater 1 $r = 0.970$, rater 2 $r = 0.965$ and rater 3 $r = 0.965$, $p < 0.001$) and inter-rater agreement (mean ICCs $r = 0.967$, $p < 0.001$) was very high between repeated measures and between markers. Error values for the z-axis (height) were the lowest.

Conclusions: The MIDAS demonstrated both high inter-rater and intra-rater reliability and provides an objective method for the assessment of posture in physiotherapy practice.

Keywords: Posture assessment, evidence-based practice, low-cost system, portable

1. Introduction

Retraining of posture is a traditionally integral physiotherapeutic intervention in the treatment of back pain [18], with the benefit of postural correction exercises for the relief of back pain being well documented [16,17,23,25]. Evaluation of posture by physiotherapists is based on subjective observations and patients' self report of what activities lead to increased pain. Improvements from treatment may be detected subjectively, but it is currently difficult to compare patients or to quantify improvements [14,15]. Further, difficulties arise when trying to compare measures taken by different physiotherapists as standardized objective assessment tools may not have been used. This lack of objective measures for assessing posture does not agree with standards set out by governing bodies in the United Kingdom, which stipulate that treatments should be based on objective markers and evidence-based practice [5].

Previous studies with physiotherapists concluded that there was a need for an assessment system to provide objective, accurate results, displayed quantitatively and visually for evidence based practice [27]. Several systems have been proposed, all of which have benefits and drawbacks. Posture charts are available that

*Address for correspondence: Dr. Josette Bettany-Saltikov, Senior Lecturer in Research Methods, University of Teesside, School of Health and Social Care, Victoria Road, Middlesbrough, TS1 3BA, UK. Tel.: +44 1642 384985; Fax: +44 1642 384905; E-mail: j.b.saltikov@tees.ac.uk.

ISSN 1053-8127/09/$17.00

rate positions of body segments relative to others in the sagittal or coronal plane and assign a score (good, fair or poor) [24], the sum of all the measures giving an overall posture score on a particular date. Drawbacks with this method include the presence of more than one postural deviation at each level, or abnormalities that arise in more than one plane, e.g. a scoliosis combined with a rib hump deformity. As the chart only allows one deviation per section, only a generalized description can be documented.

The same drawbacks exist with the chart as comparing subjective interpretations of one person taken by different physiotherapists – what one clinician identifies as a significant deviation may only be minor to another. An alternative system proposed is the Postural Analysis Digitizing System (PADS) [14] where photographs are taken of the patients' head, neck and shoulders laterally. For a tool to be clinically useful, it is essential that it fits in with the range of patients physiotherapists routinely encounter [27]. As a tool for assessing posture therefore the PADS inability to assess total trunk posture would be too limiting for routine physiotherapy practice.

Many spinal measures for orthopedic monitoring have been proposed but most are concerned with finding relationships between Cobb angles [24], measured by radiograph and the surface changes influenced by structures beneath. Potentially, orthopedic systems could be adapted for assessing posture by physiotherapists but to the authors' knowledge, all studies using these systems are concerned with finding a non-invasive, non-radiographic method of correlating surface deformation with Cobb angles. The systems have not been used primarily for postural assessment and have been criticized for their complexity of operation [6], lack of portability and significant costs [27].

This study uses a relatively low-cost, portable system, known as the MIDAS (Middlesbrough Integrated Digital Assessment System), with software specifically designed for the assessment of back posture. The equipment consists of a mechanical arm designed for creating 3D computer models of physical objects [11]. The MIDAS is able to measure the distance in three dimensions from the base of the unit to the tip of the arm, these measurements are then automatically stored by the program on the computer for analysis. The patient is positioned in front of the MIDAS, allowing measurement of landmarks of their back for posture assessment. Modifications of this system are already in place in some hospitals for guidance during CT scans [12]. It is hoped that the MIDAS system could be implemented as a means of quantifying posture in physiotherapy departments. Before the MIDAS system can be implemented for this purpose however, evidence of the reliability and validity must be demonstrated [20]. Previous work using the MIDAS system have demonstrated very high intra-rater reliability (ICC > 0.999, $p < 0.0001$) on an anatomical mannequin [30] and on 50 subjects ($r = 0.92$ to 0.99, $p < 0.001$) in an upright position [28]. No studies have reported inter-rater reliability in standing back posture to the authors' knowledge. The aims of this study were to evaluate the inter-rater reliability and intra-rater reliability of the MIDAS system on standing back posture in healthy young adult subjects.

2. Materials and methods

2.1. Subjects

A repeated-measures design [29] using a highly standardized protocol [9] was implemented for this study, with both intra-rater and inter-rater reliability being reported for the three testers. A non-probabilistic convenience sample [9] of University of Teesside (UoT) students were recruited for this study ($n = 20$ females, 5 males). Ages ranged 23–28 (mean $= 24.68$), height ranged 1.55–1.94 m (mean $= 1.69$ m), weight ranged 45–100 kg (mean $= 67.6$ kg) and body mass index ranged 16.9–33.1 (mean $= 23.3$). Subjects were excluded if they had any lower limb or back injury that prevented the subject standing for the duration of data collection, any vestibular problems that prevented the subject maintaining normal balance for the duration of data collection or a known allergy to self-adhesive stickers when in contact with the skin. Ethical approval was granted by the UoT School of Health and Social Care ethics committee. Consent and data protection issues were observed throughout. Written consent was given by all participants prior to commencement, after which, all participants were given a subject number for identification during the study. Subject details were held securely in the laboratory and were only held for the duration of data collection, allowing data to be removed in the event of a withdrawal. Information was supplied to subjects on dissemination of the research at their request.

2.2. Instrumentation

The MIDAS (Fig. 1) is a tool for acquiring a static 3-D computer recording of a physical object. A coun-

Table 1
Key to standing back anatomical landmarks measured

Label	Anatomical point
AL	Left acromion processes
AR	Right acromion process
SL	Left inferior angle of scapulae
SR	Right inferior angle of scapulae
ICL	Left iliac crest
ICR	Right iliac crest
PSL	Left posterior superior iliac spine
PSR	Right posterior superior iliac spine
C2	2nd cervical vertebra
CA	Anterior cervical vertebra
VP	Vertebra prominens
TA	Anterior thoracic vertebra
T12	12th thoracic vertebra
LA	Anterior lumbar vertebra
SA	Sacral point

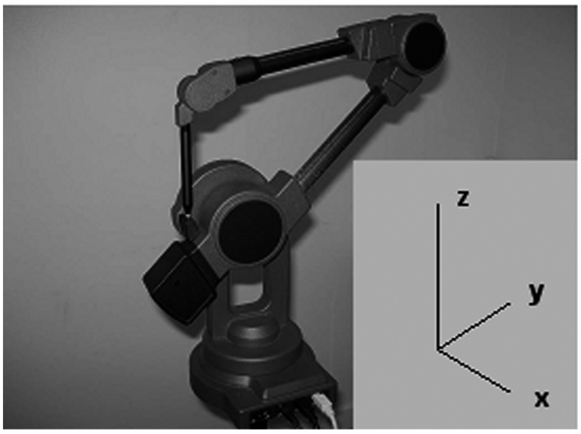

Fig. 1. The Microscribe in resting position.

terbalanced mechanical arm has optical sensors in each joint for X, Y, Z coordinate awareness with a mean accuracy of 0.23 mm [11]. A footplate was created with marks to standardize foot position and a chart was placed on the wall in front of the subject with markers to focus on. Previous work has found improvements in repeat measurements with foot and vision standardization [1,14]. The MIDAS was placed on an adjustable tripod for positioning and connected via a serial port to a laptop PC for data storage. A set of scales and a stadiometer were used to obtain weight and height measurements of subjects.

2.3. Procedure

A pilot study using one subject and the three testers was carried out to ensure the procedure was correct and the results were suitable for analysis. Subjects read

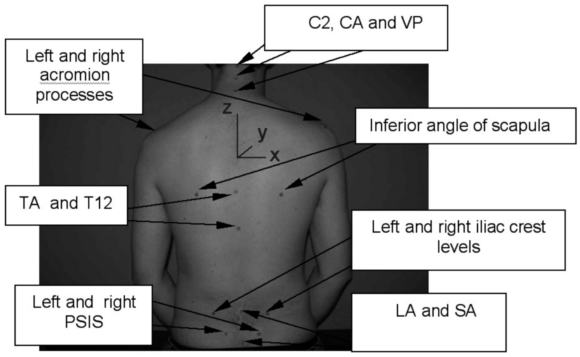

Fig. 2. Back Landmarks identified with self-adhesive markers.

the subject information sheet and after consenting to participate were attired so that their back was visible for landmark identification. To limit the intra-rater and inter-rater sources of error, palpation and marking of anatomical landmarks with 8mm self-adhesive stickers was carried out by the principal researcher with the subject standing and left in place for the other testers to measure. Training of the testers was carried out to ensure a standardized procedure was used throughout data recording. Subjects stood and fixed their vision to a point on a wall chart, in agreement with other studies of postural assessment tools [1,2,14]. The landmarks used were identical to those used in previous MIDAS studies. Landmarks were identified as shown in Table 1 and Fig. 2. The order of testers was randomized using a random sampling numbers table [13], this step being taken to minimize the effect of postural changes over time that may have introduced a bias. Data collection involved one tester touching the MIDAS stylus tip to each of the marked points in a standardized order dictated by the software and pressing the foot pedal of the MIDAS to store the position on the computer. The same tester repeated the process once more, totaling two sets of data per subject, per tester (six per subject in total). The anatomical markers remained *in situ* for the second and third tester to repeat the process.

2.4. Statistical analysis

Data was analyzed using the Statistical Package for Social Sciences (SPSS) version 13. Data was checked for normal distribution before Intra-class Correlation Coefficient's (ICC 2,1) were calculated based on methods described by Batavia [7]. Sim and Wright [26] suggest that in order to calculate intrarater and inter-rater reliability, agreement and not association should be sought.

Table 2
Intra-rater mean values for all points (all values $p < 0.05$)

	ICC	95% CI lower bound	95% CI upper bound
Rater 1	0.970	0.934	0.987
Rater 2	0.965	0.915	0.985
Rater 3	0.965	0.924	0.984

Calculating the proportion of agreement is a simple evaluation, but has been criticised for not accounting for chance agreement [10]. Kappa is a measure of agreement that accounts for chance but along with proportion of agreement, is suited to categorical or ordinal data [4]. The acceptance of the ICC (for similar study designs) over other less appropriate methods e.g. Kappa) has permeated the research community [8].

The intraclass correlation coefficient was selected as the most appropriate statistical test as it allows for numerous sets of data from numerous raters to be compared. The ICC test however isn't without limitations, as the potential exists to select the wrong calculation equation and identifying a bias in results is not possible [22]. A potential solution would be to use the ICC in conjunction with Bland and Altman graphs [3], which allow a visual representation of the data and allows identification of a bias easily. The drawbacks with Bland and Altman tests however are that the calculations become very complex if more than two raters are being used and a sample size of greater than 50 subjects is required. The current study had a sample size of 25 subjects due to resource limitations, therefore the ICC test alone was selected as the test with the least limitations for this study. An ICC value of 1 would represent prefect reliability and a value of 0.75–0.90 represents good to excellent reliability [21].

3. Results

The intra-rater mean scores (CI's) were all very high ($r = 0.970$ (0.934–0.987), $p < 0.001$ rater 1), $r = 0.965$ (0.915–0.985, $p < 0.001$ rater 2) and $r = 0.965$ (0.924–0.984, $p < 0.001$ rater 3) (Table 2). The inter-rater ICC values (CI's) ranged from 0.903–0.978, ($p < 0.05$) with a mean of 0.967 (Table 3). Consistently across all raters and all fifteen points measured, the values for the z-axis were higher than the x or y-axis scores (0.997 vs. 0.885 and 0.918) possibly indicating greater reliability in this aspect of the MIDAS (Table 4).

4. Discussion

The results from this study indicate that the intra-rater and inter-rater agreement, as measured by the ICC was very high. Mean *intra-rater* ICC values (CIs) for all points ranged from $r = 0.915$ to $r = 0.987$ and the mean *inter-rater* ICC values (CIs) for all points was $r = 0.967$ (0.903 to 0.978). These values correspond to Portney and Watkins [21] good to excellent reliability category (0.75–0.90). Therefore, this demonstrates excellent agreement between repeated measures made by individual raters and repeated measures between raters with the MIDAS on healthy young adults. This study furthers that of two other reliability studies using the MIDAS [30,28] that found very high intra-rater agreement on an anatomical mannequin and fifty healthy subjects ($r = 0.999$, $p < 0.0001$ 2005 $r = 0.92$ to 0.99, $p < 0.00001$).

However, some variability existed between the results of the three axes measured. The z-axis corresponds to height measurements and therefore should be unaffected by factors such as postural sway, whereas the x and y axes are susceptible to medio-lateral (ML) and antero-posterior (AP) sway. The length of path of the centre of pressure (LCOP) has been found to vary by up to 450 mm, when subjects are measured in quiet standing on a force platform for 30 seconds [19]. The order of testers was randomized in this study in an attempt to minimize changes in posture over time. Despite this step, postural sway was unlikely to be controlled. Further studies with the MIDAS are currently addressing the small variations demonstrated in the x and y-axes which are postulated to be due to variations in postural sway.

Previous research has found the acceptance of the MIDAS by physiotherapists to be good, with conclusions being drawn for advancements in the system [27]. Van Schaik et al. [27] found that ease of use and perceived usefulness improved the prospects of the MIDAS being used in clinical practice [27]. Long-term studies on both asymptomatic and symptomatic back pain populations with the MIDAS may provide information on factors predisposing people to developing back pain, which, clinically would be valid as there is currently a lack of comparable data [19]. The present study and recent research [27,28] have provided areas for future development of the MIDAS – functionality and ease of use. Within functionality, it was recognized that clinically relevant measures are required and Van Schaik et al. [27] comment that different outputs from the MIDAS may be required. Currently the MIDAS

Table 3
Inter-rater mean values for all testers and subjects (all values $p < 0.05$)

	ICC	95% CI lower bound	95% CI upper bound
Inter-rater mean	0.967	0.903	0.978
Mean of all x points	0.885	0.815	0.915
Mean of all y points	0.918	0.846	0.960
Mean of all z points	0.997	0.993	0.999

Table 4
Comparisons of the reliability for the x, y and z measurements for each of the three raters (all values $p < 0.001$). Two points, the left and right acromion processes have been used to illustrate the mean results of the X, Y and Z axis for each of the three different raters. The last three rows show the results of the means of the 3 axes for all the points

	Intrarater ICC rater 1	Intrarater ICC rater 2	Intrarater ICC rater 3
Left acr process x	0.954	0.977	0.961
Left acr process y	0.872	0.894	0.874
Left acr process z	0.998	0.999	0.998
Right acr proc x	0.977	0.972	0.933
Right acr proc y	0.941	0.922	0.948
Right acr proc z	0.998	0.998	0.999
Mean of all x points	0.912	0.912	0.894
Mean of all y points	0.940	0.924	0.844
Mean of all z points	0.997	0.998	0.998

software allows calculation of landmark heights relative to points on the spine [28] however, it is unable to display this visually which may impact on the systems' ease of use. Future development could provide such an output, thus allowing clinicians and patients to observe postural changes over time visually.

The limitations of this study have been discussed in part throughout the discussion, namely the variation in the x and y axes which may be due to postural sway and the inability to display the results in a user friendly format. Other major limitations of the study include the use of a convenience sample, the use of one researcher alone to apply the landmark stickers and the use of asymptomatic subjects. The convenience sample was due to necessity at the time of subject recruitment. The rational for the use of one researcher to apply the landmark stickers related to the fact that the purpose of the current study was to evaluate the reliability of the assessment tool itself rather than the assessment of the reliability of palpating the spinous processes by different raters. To address these shortcomings further studies are required, as in everyday use, different users would be selecting and recording the landmarks. Future studies would benefit from a larger sample size and several testers each palpating and selecting the landmarks.

5. Conclusions

The aims of this study were to assess the intra-rater (with three raters) and inter-rater (with three raters) reliability of the MIDAS 3D posture assessment tool when repeated measures were made on a convenience sample of healthy young adults. The results confirm that agreement between repeated measures for intra-rater values and inter-rater values were very high, as analyzed by ICC (2, 1). It is concluded that subject postural sway may have caused slightly lower values for the x and y values. Future studies need to focus on controlling body movement, testing the reliability of spinal process palpation with different users, developing a visual representation of the MIDAS' output and using a patient population, with the aim of leading to clinical uptake of this low-cost portable posture assessment tool.

References

[1] M. Batavia, *Clinical Research for Health Professionals: A User-Friendly Guide*, Butterworth-Heineman, Boston, 2001, 45–93.

[2] J.A. Bettany and D.J. Harrison, The Stanmore Footplate: An Improved Method in ISIS Optical Scanning, in: *Surface Topography and Spinal Deformity, Proceedings of the 6th International Symposium*, A. Alberti, B. Drerup and E. Hierholzer, eds, Gustav Fischer Verlag/VCH Publishing, 1992, pp. 42–46.

[3] J.M. Bland and D.G. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, *Lancet* **1** (1986), 1–32.

[4] S. Burt and L. Punnett, Evaluation of Interrater Reliability for Posture Observations in a Field Study, *Applied Ergonomics* **30** (1999), 121–135.

[5] Chartered Society of Physiotherapy, *Core Standards of Physiotherapy Practice*, London, Standard, 2000, 10.1c.

[6] J.R. Cobb, An Outline for the Study of Scoliosis (Instructional Course Lectures 5) (1948), American Academy of Orthopedic Surgeons, Park Ridge, Illinois, 261, in: N.D. Scutt, P.H. Dangerfield and J.C. Dorgan, The Relationship Between Surface and Radiological Deformity in Adolescent Idiopathic Scoliosis: Effect of Change in Body Position, *European Spine Journal*, **5** (1996), 85–90.
[7] E. Domholdt, *Physical Therapy Research*, (2nd ed.), W.B. Saunders Company, USA, 2000, 239–247.
[8] M. Eliasziw, S.L. Young, M.G. Woodbury and K. Fryday-Field, Statistical Methodology for the Concurrent Assessment of Interrater and Intrarater Reliability: Using Goniometric Measurements as an Example, *Physical Therapy* **74** (1994), 777–788.
[9] M. Eliasziw, S.L. Young, M.G. Woodbury and K. Fryday-Field, Statistical Methodology for the Concurrent Assessment of Interrater and Intrarater Reliability: Using Goniometric Measurements as an Example, *Physical Therapy* **74** (1994), 777–788.
[10] J. Fleiss, (1973) Statistical Methods for Rates and Proportions, Wiley, New York, in: *Evaluation of Interrater Reliability for Posture Observations in a Field Study*, S. Burt and L. Punnett, eds, Applied Ergonomics, 1999, p. 30, p. 122.
[11] C.J. Goldberg, M. Kaliszer, D.P. Moore, E.E. Fogarty and F.E. Dowling, Surface Topography, Cobb angles, and Cosmetic Change in Scoliosis, *Spine* **26**(4) (2001), E55–E63.
[12] Immersion Corporation (online), http://www.immersion.com/digitizer/docs/MSG2_0704_V1.pdf, 2005 [accessed 24/11/08].
[13] J.L. Jaremko, P. Poncet, J. Ronsky, J. Harder, J. Dansereau, H. Labelle and R.F. Zernicke, Estimation of Spinal Deformity in Scoliosis From Torso Surface Cross Sections, *Spine* **26**(14) (2001), 1583–1591.
[14] B. Lafferty-Braun and L.R. Amundson, Quantitative Assessment of Head and Shoulder Posture, *Archives of Physical and Medical Rehabilitation* **70** (1989), 322–329.
[15] D.J. Magee, *Orthopedic Physical Assessment*, (4th ed.), Saunders, 2002, 876–903.
[16] A.F. Mannion, M. Müntener, S. Taimela and J. Dvorak, 1999 Volvo Award Winner in Clinical Studies: A Randomized Clinical Trial of Three Active Therapies for Chronic Low Back Pain, *Spine* **24**(23) (1999), 2435–2448.
[17] M. Moore, Upper Crossed Syndrome and its Relationship to Cervicogenic Headache, *Journal of Manipulative Physiological Therapy* **27** (2004), 414–420.
[18] P.B. O'Sullivan, K.M. Grahamslaw, M. Kendell, S.C. Lapenskie, N.E. Möller and K.V. Richards, The Effect of Different Standing and Sitting Postures on Trunk Muscle Activity in a Pain-Free Population, *Spine* **27**(11) (2002), 1238–1244.
[19] G.L. Pellecchia, Postural Sway Increases with Attentional Demands of Concurrent Cognitive Task, *Gait and Posture* **18** (2003), 29–34.
[20] Philips Medical Systems (online), http://www.medical.philips.com (2008), [accessed 24/11/08].
[21] L.G. Portney and M.P. Watkins, *Foundations of Clinical Research: Applications to Practice*, Appleton and Lange, Norwalk, 1993.
[22] G. Rankin and M. Stokes, Reliability of assessment tools in rehabilitation: an illustration of appropriate statistical analyses, *Clinical Rehabilitation* **12** (1998), 187–199.
[23] E. Rasmussen-Barr, L. Nilsson-Wikmar and I. Arvidsson, Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain, *Manual Therapy* **8**(4) (2003), 233–241.
[24] Reedco Research, *Reedco Posture Score Sheet, Orthopedic Assessment and Treatment of the Geriatric Patient*, Mosby, St. Louis, 1974, 72.
[25] C.A. Richardson and G.A. Jull, Muscle Control–Pain Control. What Exercises Would You Prescribe? *Manual Therapy* **1** (1995), 2–10.
[26] J. Sim and C. Wright, *Research in Health Care: Concepts, Designs and Methods*, Nelson Thornes, 2000, 18–21, 123–135 and 294–360.
[27] P. Van Schaik, J.A. Bettany-Saltikov and J.G. Warren, Clinical Acceptance of a Low-cost Portable System for Postural Assessment, *Behaviour and Information Technology* **21**(1) (2002), 47–57.
[28] J.G. Warren, J. Bettany-Saltikov, P. Van Schaik and S.L. Papastefanou, 3-D measurement of posture and back shape using a low cost, portable system – a reliability study, in: *Research in Spinal Deformities 3*, A. Tanguy and B. Peuchot, eds, IOS Press, Amsterdam 2002, pp. 100–104.
[29] J.G. Warren, J. Bettany-Saltikov, P. Van Schaik and S. Papastefanou, Evidence-Based Postural Assessment for use in Therapy and Rehabilitation, *International Journal of Therapy and Rehabilitation* **1212** (2005) 527–532.
[30] S. Wood-Dauphinee and J.I. Williams, Much Ado About Reliability, *Physiotherapy Canada* **41**(5) (1989), 234–235.