Physiotherapy Theory and Practice, 26(5):327–333, 2010

ISSN: 0959-3985 print/1532-5040 online
DOI: 10.3109/09593980903094230

informa
healthcare

DESCRIPTIVE REPORT

Reliability of shoulder range of motion comparing a goniometer to a digital level

Michael J Mullaney, DPT,[1] Malachy P McHugh, PhD,[2] Christopher P Johnson, MPT, MCMP,[3] and Timothy F Tyler, PT, ATC[4]

[1]*Research Consultant, Nicholas Institute of Sports Medicine & Athletic Trauma, Lenox Hill Hospital, New York, NY, and Mullaney & Associates Physical Therapy, Matawan, NJ, USA*

[2]*Director of Research, Nicholas Institute of Sports Medicine & Athletic Trauma, Lenox Hill Hospital, New York, NY, USA*

[3]*Research Assistant/Physical Therapist, Nicholas Institute of Sports Medicine & Athletic Trauma, Lenox Hill Hospital, New York, NY, USA*

[4]*Research Consultant, Nicholas Institute of Sports Medicine & Athletic Trauma Lenox Hill Hospital, New York, NY, USA, and PRO Sports Therapy of Westchester, Scarsdale, NY, USA*

ABSTRACT

The clinical use of digital levels, for joint measurement, may be a viable alternative to standard goniometry. The purpose of this study was to determine the intra- and intertester reliability of a construction grade digital level compared to the standard universal goniometer for measurements for active assisted shoulder range of motion (ROM). Two experienced physical therapists measured shoulder flexion, external rotation (ER), and internal rotation (IR) ROM bilaterally, on two different occasions, in 20 patients (9 males, 11 females, 18–79 years old) with unilateral shoulder pathology, using a goniometer and a digital level. Relative reliability was assessed by using intraclass correlation coefficients (ICC), and absolute reliability was assessed by using 95% limits of agreement (LOA). Intratester ICCs ranged from 0.91 to 0.99, and LOA ranged from 3° to 9° for measurements made with the goniometer and digital level. Intertester ICCs ranged from 0.31 to 0.95, and LOA ranged from 6° to 25°. For the comparison of goniometric vs. digital level ROM, ICCs ranged from 0.71 to 0.98. ER and IR ROM were 3–5° greater for the digital level than the goniometer ($p < 0.01$). Goniometric vs. digital level LOA ranged from 6° to 11° for shoulder flexion. Both measurement techniques had excellent intratester reliability, but for intertester reliability ICCs were 20% lower and LOA were 2.3 times higher than intratester values. Reliability estimates were similar between the digital level and the goniometer. However, because glenohumeral rotation was 3–5° greater for the digital level than the goniometer (systematic error), the two methods cannot be used interchangeably. On the basis of the average intratester LOA for the goniometer and the digital level, a change of 6–11° is needed to be certain that true change has occurred. For comparison of measures made by two different therapists, a change is of 15° is required to be certain a true change has occurred. A digital level can be used to reliably measure shoulder ROM but should not be used interchangeably with a standard goniometer.

INTRODUCTION

Physical therapists use goniometry to objectively measure passive and active range of motion (ROM). The measurements are generally used to assess limitations in ROM, determine appropriate interventions, and document treatment progression (Ekstrand, Wiktorsson, Oberg, and Gillquist, 1982; Gajdosik and Bohannon, 1987; Riddle, Rothstein, and Lamb, 1987; Rothstien, Miller, and Roettger, 1983). Clinically, the universal goniometer is most commonly used (Hayes, Walton, Szomor, and Murrell, 2001; Rome and Cowieson, 1996; Rothstien, Miller, and Roettger, 1983; Somers et al, 1997; Watkins, Riddle, Lamb, and Personius, 1991; Youdas, Bogard, and Suman, 1993). Numerous

Accepted for publication 4 June 2009.

Address correspondence to Michael J Mullaney, DPT, Research Consultant, Nicholas Institute of Sports Medicine & Athletic Trauma, Lenox Hill Hospital, 177 East 77th Street, New York, NY 10021 and Mullaney & Associates Physical Therapy, 127 Main Street, Suite E, Matawan, NJ 07747 USA.
E-mail: mjmullaney@verizon.net

studies have investigated the reliability of goniometry for both active and passive ROM. Boone et al (1978) looked at the intertester and intratester reliability of active ROM measurements taken on both upper and lower extremities using a goniometer in normal healthy subjects. They found that intertester reliability was greater for upper extremity motion ($r = 0.86$) than lower extremity ($r = 0.58$) and intratester reliability was high for both upper ($r = 0.89$) and lower extremities ($r = 0.80$). Riddle et al (1987) compared intertester and intratester reliability of passive shoulder ROM in a clinical setting using two different sized goniometers. They demonstrated that the ICC values for intertester reliability were high for flexion and abduction when using a large and small goniometer (0.84–0.89 respectively), but poor for extension (0.27–0.26). Intratester reliability for shoulder flexion, extension and abduction were demonstrated to be high (0.94–0.98). Intertester reliability of horizontal abduction/adduction and internal rotation measurements were taken with large and small goniometers and shown to be poor (0.28–0.55), whereas external rotation measurements were high (0.88–0.90). Intratester reliability was high (0.87–0.99) for both horizontal abduction/adduction and internal/external rotation.

An alternative approach is to use a digital level to document joint angle relative to a predetermined zero position. This may be a viable alternative to standard goniometry, but the reliability of this technique has not been established. An advantage to the digital level is that zero degrees can be fixed to the plane of interest while for testing with a goniometer the reference arm of the goniometer must be manually held in the reference position while the other arm is rotated with the joint being tested. This may be advantageous in situations such as glenohumeral internal and external rotation at 90° of abduction where the reference arm has no bony landmark that it can reference to. The digital level may also be a valuable instrument in passive range of motion measurement situations where the examiner may only have one free hand to take a measurement, such as passive straight leg raise.

The reliability of the standard universal goniometer is well established, and this method of measuring ROM is widely accepted. However, when a tester is measuring active assisted ROM, it may be difficult to keep the reference arm of the goniometer stationary while rotating the joint. It also may be difficult to read the goniometer at end ROM, and removing the goniometer from the joint to read the value can result in unintended movement of the goniometer. Because a digital level can be referenced to a fixed angle, and the digital display is more easily read and can be locked at end ROM, this instrument may have some clinical value. However, the reliability of ROM measurements made with a digital level has not been reported. Therefore, the purpose of this study was to determine the intratester and intertester reliability of a construction grade digital level compared to the standard universal goniometer for measurements for active assisted shoulder range of motion (ROM).

METHODS

Subjects

Twenty subjects with unilateral shoulder pathology (9 males, 11 femailes) between the ages of 18 and 79 (51 ± 18) volunteered to participate in this study. Patients included in this study were required to attain 90° abduction (ABD) during active assisted testing. Nine patients had rotator cuff pathology; three patients had shoulder instability; and two patients had coracoclavicular ligament reconstructions. Other diagnoses included proximal humeral fracture, clavicle fracture, frozen shoulder, greater tuberosity fracture, biceps tenodesis, and rheumatoid arthritis. Lenox Hill Hospital Institutional Review Board approved the study, and all subjects read and signed an informed consent form prior to participation.

Procedures/protocol

Patients were selected from a sample of convenience. Prior to measurements, to blind the examiner from the ROM measurements, the universal goniometer scale was covered, and all measurements taken with the digital level were done so with the digital screen facing away from the measuring examiner. The digital level was calibrated to zero prior to each patient being measured. Two separate examiners with various years of experience (5 years and 3 years) measured the patients' active assisted ROM. Two trials per examiner, per shoulder were taken, with measurements taken on the uninvolved side first. The order in which measurements were taken (Flexion, ER, IR), which tool was used (goniometer vs. digital level), and which examiner measured was randomized. All measurements were recorded with the patient in the supine position and the nontesting examiner recorded all measurements.

Universal goniometer measurements

A 12-inch, 360° goniometer, marked in 1° increments, with two adjustable overlapping arms was used. Shoulder flexion ROM was taken by asking the patient to raise their arm straight over-head as far as possible. Standard measurement positioning was used by placing the stationary arm parallel to the midline of the thorax, and the moving arm aligned with the shaft of the humerus and lateral epicondyle (Norkin and White, 1988).

Shoulder ER was taken by passively placing the patient's arm in 90° abduction with the elbow flexed 90° and asking the patient to rotate their arm backward as far as possible so that their palm was facing the ceiling. Standard goniometric positioning was used by placing the stationary arm perpendicular to the floor, and the moving arm was aligned with shaft of the ulna and styloid process (Norkin and White, 1988).

Shoulder IR was taken by passively placing the patients arm in 90° abduction with the elbow flexed 90° and asking the patient to rotate their arm forward as far as possible so that their palm was facing the floor. Positioning of the goniometer for measurement was also used in a standardized fashion (Norkin and White, 1988).

Digital level measurements

An M-D construction grade Smarttool digital level (M-D Building Products, Oklahoma City, OK, USA) was used in this study (Figure 1). This digital level displays degrees, percent slope, and pitch to 1/10° accuracy and can be set to zero in either the horizontal or vertical plane. All instructions for patient positioning were the same as for the goniometer. All landmarks described for placement of the digital level were consistent with anatomical position. Shoulder flexion measurements were taken by placing the digital level along the posterior aspect of the arm, parallel with the midline of the humerus, resting on the soft tissue of the triceps. Shoulder ER measurements were taken by placing the digital level on the anterior aspect of the forearm, parallel with the midline of the ulna, resting on the soft tissue of the flexor mass. Shoulder IR measurements were taken by placing the digital level on the posterior aspect of the forearm, parallel with the midline of the ulna, resting on the soft tissue of the extensor mass.

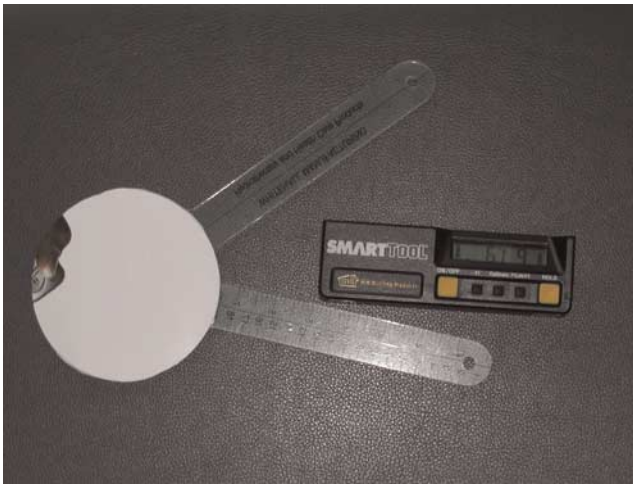

FIGURE 1 Smarttool digital level and standard goniometer.

Data analysis

Methods of assessing reliability based on correlation coefficients such as the Pearson correlation coefficient or intraclass correlation coefficient (ICC) have been referred to as relative reliability (Atkinson and Nevill, 1998). The ICC is the most commonly used method of measuring relative reliability. The ICC value represents the relationship between the within-subjects variability and the between-subjects variability. The within-subjects variability represents measurement error while the between-subjects variability represents the heterogeneity of the sample. Intratester, intertester, and inter-instrument relative reliability were assessed by using the ICC (3,1) (Portney and Watkins, 1993). SPSS 10.0 was used for computing ICC values, and the two-way mixed model single-measure ICC is reported.

There are three limitations to using ICCs for assessing the reliability of clinical tests. First, the derived result does not provide a measure of measurement error in the actual units of measurement for the particular clinical measure. Second, the ICC value does not indicate if there is a systematic error (bias) between tests. Third, the ICC value is influenced by the range of measured values (sample heterogeneity). Higher ranges (greater heterogeneity) are associated with higher ICCs, independent of actual measurement error (Atkinson and Nevill, 1998; Weir, 2005). Therefore, it is desirable to also provide a measure of absolute reliability such as the standard error of the measurement (SEM); the coefficient of variation; or the 95% limits of agreement (LOA) (Atkinson and Nevill, 1998). The 95% LOA have been used to assess the reliability of cervical and lumbar spine range of motion (Hoving et al, 2005; Jones, Stratton, Reilly, and Unnithan, 2002) but have not been widely used in the physical therapy literature. This method provides an estimate of the range of error within which 95% of repeated measures can be expected to fall (assuming no systematic error). Clinically, the 95% LOA represent the threshold a clinician can use for distinguishing a real change from measurement error (McHugh, Lexell, and Bangsbo, 2003). LOA are analgous to the often reported minimal detectable change (MDC) with one important advantage; LOA are unaffected by sample heterogeneity. The 95% LOA are calculated as:

$$\mathrm{LOA} = \mathrm{SDd} * 1.96 \quad (1)$$

where *SDd* is the standard deviation of the difference between two repeated measures on a sample. If the mean difference between two tests was significantly different from zero, the LOA were not reported, and the test was classified as having bias (systematic error).

Average intratester and intertester ICCs and LOA were compared by using paired *t*-tests; 12 paired values: 2 sides (involved, noninvolved); 3 motions (Flex, IR, ER); 2 instruments (goniometer, digital level). Besides ICC values and 95% limits of agreement for inratester, intertester and interinstrument comparisons, intersubject mean ± SD is reported for range of motion measurements.

RESULTS

Intratester and intertester reliability

Mean values for ROM for the involved and noninvolved arms, between tests, between testers, and between instruments are displayed in Table 1. Intratester ICCs ranged from 0.91 to 0.99 for both the goniometer and the digital level (Table 2). Intratester LOA ranged from 3° to 9° for the goniometer and digital level. Intratester ICCs were not different between tester 1 and tester 2, between the involved and noninvolved arms, and between the goniometer and the digital level (Table 2). Intratester LOA were 2° lower on the noninvolved side ($p < 0.001$) but were not different between tester 1 and tester 2 or between the goniometer or the digital level.

Intertester ICCs ranged from 0.62 to 0.95 for the goniometer and 0.31 to 0.93 for the digital level (Table 3). Intertester LOA ranged from 7° to 23° for the goniometer and 6° to 25° for the digital level (Table 3). Intertester ICCs were higher on the involved vs. the noninvolved side (0.89 vs. 0.71, $p < 0.001$). Intertester ICCs were not different between test 1 and test 2, or between the goniometer and the digital level. Intertester LOA were higher for the noninvolved vs. the involved for IR (22° vs. 14°, $p < 0.05$). Intertester LOA were not different between test 1 and test 2 or between the goniometer and the digital level.

It was apparent that intratester ICCs were higher than intertester ICCs (0.96 vs. 0.8, $p < 0.001$) and intratester LOA were lower than intertester LOA (6° vs 15°, $p < 0.001$).

TABLE 1 Goniometric and digital level range of motion measurements for test 1 and test 2 for tester 2 and tester 2 (I = involved, N = noninvolved; Flex = shoulder flexion, ER = external rotation, IR = internal rotation). *ROM greater for digital level vs. goniometer ($p < 0.05$)

		Goniometer		Digital level	
		Tester 1		Tester 1	
		Test 1	Test 2	Test 1	Test 2
Flex	I	156 ± 14	155 ± 13	155 ± 14	156 ± 13
	N	172 ± 6	172 ± 7	172 ± 6	172 ± 7
ER	I	69 ± 20	69 ± 21	72 ± 21*	72 ± 21*
	N	92 ± 13	93 ± 12	96 ± 17*	97 ± 16
IR	I	55 ± 13	55 ± 13	59 ± 13*	59 ± 13*
	N	62 ± 14	62 ± 13	66 ± 14*	67 ± 13*
		Tester 2		Tester 2	
		Test 1	Test 2	Test 1	Test 2
Flex	I	154 ± 15	155 ± 14	156 ± 14	156 ± 13
	N	172 ± 5	172 ± 6	173 ± 6	172 ± 6
ER	I	70 ± 22	69 ± 22	72 ± 26	71 ± 24
	N	94 ± 13	94 ± 13	99 ± 14*	98 ± 15*
IR	I	54 ± 15	53 ± 14	59 ± 10*	57 ± 11*
	N	60 ± 9	59 ± 10	65 ± 9*	64 ± 8*

Agreement between the goniometer and the digital level (interinstrument reliability)

For the comparison of goniometric vs. digital level ROM, ICCs ranged from 0.71 to 0.98 (Table 4). There

TABLE 2 Intratester reliability (ICC and LOA) for shoulder flexion, external rotation, and internal rotation (I = involved, N = noninvolved; Flex = shoulder flexion, ER = external rotation, IR = internal rotation)

		Intratester reliability			
		Goniometer		Digital level	
		Tester 1 ICC	Tester 2 ICC	Tester 1 ICC	Tester 2 ICC
Flex	I	0.96	0.97	0.97	0.98
	N	0.97	0.91	0.94	0.96
ER	I	0.99	0.99	0.99	0.98
	N	0.97	0.98	0.99	0.98
IR	I	0.94	0.95	0.97	0.91
	N	0.96	0.91	0.98	0.92
		LOA	LOA	LOA	LOA
Flex	I	7°	8°	6°	6°
	N	3°	5°	5°	3°
ER	I	6°	7°	6°	9°
	N	6°	5°	5°	5°
IR	I	9°	8°	6°	9°
	N	7°	8°	5°	7°

TABLE 3 Intertester reliability (ICC and LOA) for shoulder flexion, external rotation, and internal rotation (I = involved, N = noninvolved; Flex = shoulder flexion, ER = external rotation, IR = internal rotation)

		Intertester reliability			
		Goniometer		Digital level	
		Test 1 ICC	Test 2 ICC	Test 1 ICC	Test 2 ICC
Flex	I	0.88	0.93	0.91	0.91
	N	0.79	0.74	0.87	0.86
ER	I	0.95	0.92	0.91	0.93
	N	0.76	0.79	0.82	0.85
IR	I	0.87	0.82	0.82	0.84
	N	0.63	0.62	0.47	0.31
		LOA	LOA	LOA	LOA
Flex	I	14°	10°	11°	11°
	N	7°	9°	6°	7°
ER	I	13°	17°	20°	17°
	N	17°	16°	18°	16°
IR	I	14°	16°	14°	13°
	N	20°	23°	20°	25°

TABLE 4 Interinstrument reliability (ICC and LOA) for shoulder flexion, external rotation, and internal rotation (I = involved, N = noninvolved; Flex = shoulder flexion, ER = external rotation, IR = internal rotation). See Table 1 for magnitude of bias

		Inter-Instrument reliability			
		Tester 1		Tester 2	
		Test 1 ICC	Test 2 ICC	Test 1 ICC	Test 2 ICC
Flex	I	0.93	0.91	0.93	0.95
	N	0.81	0.86	0.81	0.84
ER	I	0.98	0.98	0.98	0.96
	N	0.71	0.81	0.93	0.94
IR	I	0.92	0.96	0.82	0.86
	N	0.93	0.92	0.88	0.83
		LOA	LOA	LOA	LOA
Flex	I	10°	10°	11°	9°
	N	7°	7°	7°	6°
ER	I	Bias	Bias	10°	13°
	N	22°	Bias	Bias	Bias
IR	I	Bias	Bias	Bias	Bias
	N	Bias	Bias	Bias	Bias

was measurement bias for 13 of 16 interinstrument comparisons of glenohumeral rotation ROM (Table 4). ER and IR ROM was 3–5° greater with the digital level vs. the goniometer (Table 1). Goniometric vs. digital level LOA ranged from 6° to 11° for shoulder flexion with no measurement bias (Table 4).

DISCUSSION

This study established reliability for active assisted shoulder flexion and shoulder rotation ROM at 90° of abduction. The major findings in this study were the following: 1) intratester reliability was excellent for shoulder ROM measurements made with the goniometer and digital level; therefore, both instruments can be used to reliably measure shoulder ROM; 2) intertester reliability was significantly worse than intratester reliability; 3) reliability was similar between the goniometer and digital level; and 4) glenohumeral internal and external rotation ROM were significantly greater when measured with the digital level vs. the goniometer; therefore, the instruments should not be used interchangeably.

It has been suggested that ICCs should exceed 0.9 for clinical measurements and that values below 0.75 indicate poor to moderate reliability (Portney and Watkins, 1993). All 24 intratester values (2 sides, 3 motions, 2 instruments) exceeded 0.9, but only 7 of 24 intertester values exceeded 0.9 and 5 were below 0.75 (noninvolved flexion with the goniometer and all measurement of IR on the noninvolved side). However, it is important to note that intertester ICCs were better for the involved vs. the noninvolved, and 7 of 12 values on the involved side exceeded 0.9 while zero of 12 exceeded 0.9 on the noninvolved side.

The LOA values provide quantitative results that may aid clinical interpretation of reliability data. The 95% LOA represent the range within which 95% of repeated measures can be expected to fall. If a clinician is trying to detect a change in ROM due to deterioration of the clinical condition or improvement with treatment, and the observed change is greater than the 95% LOA, then it is likely that it is a real change and not due to measurement error. Because the LOA for all intratester comparisons were less that 10°, a change of greater than that for shoulder flexion IR or ER ROM can be reliably detected by an examiner. For many of the tests much smaller changes can be detected (Table 2). However, if different testers make a repeated ROM measurement on a patient to document clinical change, the change in ROM would have to exceed 14° for flexion, 20° for ER, and 25° for IR (Table 3) to be reliably interpreted as a real change. This measurement error is likely in excess of what would be deemed a clinically relevant change in ROM. However, it was apparent that LOA varied

between the involved and noninvolved sides for the different motions tested. Of interest to the clinician is the reliability of repeated measures on the involved side for a particular motion. Average intertester LOA on the involved side was 12° for shoulder flexion, 17° for ER, and 14° for IR. Given that intratester LOA were significantly lower than intertester values, it would be ideal to have the same clinician make repeated ROM measurements on patient to maximize the ability to detect a change in ROM.

Considering that LOA are not commonly used to assess measurement error in physical therapy, it is important to understand how this approach relates to more common approaches such as the standard error of the measurement (SEM) or the minimal detectable change (MDC). As stated previously the 95% LOA are calculated as:

$$\text{LOA} = \text{SDd} * 1.96 \quad (1)$$

where *SDd* is the standard deviation of the difference between two repeated measures on a sample.

The equation for 95% MDC is:

$$\text{MDC} = \text{SEM} * 1.96 * \sqrt{2} \quad (2)$$

The standard error of the measurement (SEM) is typically estimated from the equation:

$$\text{SEM} = \text{SD} * \sqrt{1 - 1\text{CC}} \quad (3)$$

where *SD* is the standard deviation of the scores for all subjects in the sample (Weir, 2005). Because ICCs are affected by sample heterogeneity independent of actual measurement error (Atkinson and Nevill, 1998; Weir, 2005), the resultant MDC is also affected by sample heterogeneity. In addition, inclusion of ICC values in calculation of the SEM means that the measure of absolute reliability is calculated by using a measure of relative reliability. Therefore, Weir (2005) has suggested that SEM should be calculated from the standard deviation of the difference in test-retest measures (SDd):

$$\text{SEM} = \frac{\text{SDd}}{\sqrt{2}} \quad (4)$$

This approach derives a 95% MDC that equals the 95% LOA as can be seen when comparing Eq. (2), Eq. (4), and Eq. (1):

$$\text{Eq. 2} \quad \text{MDC} = \text{SEM} * 1.96 * \sqrt{2}$$

$$\text{Eq. 4} \quad \text{SEM} = \frac{\text{SDd}}{\sqrt{2}}$$

Therefore, $MDC = \frac{SDd}{\sqrt{2}} * 1.96 * \sqrt{2} \rightarrow \text{MDC} = \text{SDd} * 1.96 = LOA$ (Eq. (1))

Thus, the 95% LOA provide an MDC that is not affected by the ICC value derived from the same set of data.

The digital level may be a useful alternative to the standard goniometer in various situations: 1) where the reference arm of the goniometer is held horizontal or vertical and has no bony prominence to reference, such as with glenohumeral internal and external rotation ROM measured in 90° abduction with the elbow flexed 90°; and 2) in many passive and active range of motion measures when the examiner may not have two free hands to specifically align a goniometer, such as a passive straight-leg measurement. It is noteworthy that glenohumeral ROM values were significantly higher for the digital level than for the goniometer. This may be due to difficulty in maintaining the free arm of the goniometer in a horizontal position while rotating the other arm of the goniometer with the subject's forearm. However, four alternative, more plausible explanations may include 1) that manually placing the digital level on the anterior aspect of the forearm for external rotation testing and the posterior aspect of the forearm for internal rotation testing added a stretching force in the direction of the movement being measured; 2) placing the digital level on soft tissue vs. placing it parallel to the bone landmark (like in goniometry) may increase the reading based on the soft tissue mass of the subject; 3) that by having a free hand the tester is able to apply more force when moving the joint to end ROM; and 4) while recording digital measurements, the "HOLD" button was pressed during each measurement to capture a specific measurement. The digital level is a highly sensitive instrument that may have been influenced by the pressing of the "HOLD" button while capturing a particular measurement. These factors may have been sufficient to increase the ROM by 3–5°. If these are the reasons for the observed measurement bias, it may not be apparent for active ROM testing since the examiner would not be applying the stretch force. Active assisted ROM was examined here and hence the examiner was applying a force.

An obvious limitation in this study was including only patients who could achieve at least 90° shoulder abduction. Obviously, many patients on whom ROM measurements are important will not be able to achieve this amount of abduction. However, the goal was to establish reliability data for testing glenohumeral rotation in 90° of abduction because this is a common clinical measure, especially with overhead athletes. It is important to note that indices of reliability were similar for the involved and noninvolved arms, or for some tests reliability was better on the involved side. This finding indicates that a pathological loss of ROM does not adversely affect reliability of the measurement.

CONCLUSION

In conclusion, both the goniometer and digital level had excellent intratester reliability, but for intertester reliability reliability ICCs were 20% lower and LOA were 2.3 times higher than intratester values. A digital level may be a viable alternative to a goniometer for assessing shoulder range of motion. However, the lack agreement between the two methods of measurement for glenohumeral internal and external rotation indicates that they cannot be used interchangeably. On the basis of the average intratester LOA, an experienced physical therapist using the same instrument for repeated measures (standard goniometer or digital level) should be able to detect a change in shoulder ROM in a given patient of at least 6°, but for comparison of measures made by two different therapists, the detectable change is 15°.

REFERENCES

Atkinson G, Nevill AM 1998 Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Medicine 26: 217–238

Boone DC, Azen SP, Lin CM, Spence C, Baron C, Lee L 1978 Reliability of goniometric measurements. Physical Therapy 58: 1355–1360

Ekstrand J, Wiktorsson M, Oberg B, Gillquist J 1982 Lower extremity goniometric measurements: A study to determine their reliability. Archives of Physical Medicine and Rehabilitation 63: 171–175

Gajdosik RL, Bohannon RW 1987 Clinical measurements of range of motion: Review of goniometry emphasizing reliability and validity. Physical Therapy 67: 1867–1872

Hayes K, Walton JR, Szomor ZL, Murrell GAC 2001 Reliability of five methods for assessing shoulder range of motion. Australian Journal of Physiotherapy 47: 289–294

Hoving JL, Pool JJ, van Mameren H, Deville WJ, Assendelft WJ, de Vet HC, de Winter AF, Koes BW, Bouter LM 2005 Reproducibility of cervical range of motion in patients with neck pain. BMC Musculoskeletal Disorders 13: 59

Jones MA, Stratton G, Reilly T, Unnithan VB 2002 Measurement error associated with spinal mobility measures in children with and without low-back pain. Acta Paediatrica 91: 1339–1343

McHugh MP, Lexell J, Bangsbo J 2003 Training and performance tests in sports and rehabilitation. In: Kjaer M, Magnusson SP, Krogsgaard M, Roos H, Engbretsen L, Woo S, Takala T (eds) Textbook of sports medicine, pp 201–225. Cambridge, MA, Blackwell Science

Norkin CC, White DJ 1988 Measurement of joint motion: A guide to goniometry, pp 15–22. Philadelphia, F.A., Davis Co

Portney, LG., Watkins, MP., 1993. Statistical measures of reliability. In: Foundations of clinical research: Applications to practice, pp 505–528. Englewood Cliffs, NJ, Prentice Hall

Riddle DL, Rothstein JM, Lamb RL 1987 Goniometric reliability in a clinical setting. Physical Therapy 67: 668–673

Rome K, Cowieson F 1996 A reliability study of the universal goniometer, fluid goniometer, and electrogoniometer for the measurement of ankle dorsiflexion. Foot and Ankle International 17: 28–32

Rothstien JM, Miller PJ, Roettger RF 1983 Goniometric reliability in a clinical setting. Physical Therapy 63: 1611–1615

Somers DL, Hanson JA, Kedzierski CM, Nestor KL, Quinlivan KY 1997 The influence of experience on the reliability of goniometric and visual measurement of forefoot position. Journal of Sports and Orthopedic Physical Therapy 25: 192–202

Watkins MA, Riddle DL, Lamb RL, Personius WJ 1991 Reliability of goniometric measurements and visual estimates of knee range of motion obtained in a clinical setting. Physical Therapy 71: 90–96

Weir JP 2005 Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. Journal of Strength and Conditioning Research 19: 231–240

Youdas JW, Bogard CL, Suman VJ 1993 Reliability of goniometric measurements and visual estimates of ankle joint active range of motion obtained in a clinical setting. Archives of Physical Medicine and Rehabilitation 74: 1113–1118