Clinical Technical Note

The reliability and concurrent validity of scapular plane shoulder elevation measurements using a digital inclinometer and goniometer

Morey J. Kolber, PT, PhD, OCS, CSCS,1 Cydne Fuller, BA,1 Jessica Marshall, BS,1 Amanda Wright, MPH, CSCS,1 and William J. Hanney, DPT, ATC, CSCS2

1Nova Southeastern University, Department of Physical Therapy, Fort Lauderdale, Florida, USA
2University of Central Florida, Program in Physical Therapy, Orlando, Florida, USA

ABSTRACT

This study investigated the reliability and concurrent validity of active shoulder elevation in the scapular plane (scaption) using a digital inclinometer and goniometer. Two investigators used a goniometer and digital inclinometer to measure scaption on 30 asymptomatic participants in a blinded repeated measures design. Good reliability was present with intraclass correlation coefficients (ICCs) for intrarater reliability of goniometry = 0.87, intrarater digital inclinometry = 0.88, interrater goniometry = 0.92, and interrater digital inclinometry = 0.89. The minimal detectable change (MDC95) for the interrater analysis indicated that a change equal to or greater than 8 degrees for goniometry and 9 degrees for inclinometry is required to be 95% certain that the change is not due to intertrial variability or measurement error. The concurrent validity between goniometry and digital inclinometry was excellent with an ICC value of 0.94 for both raters. The 95% limits of agreement suggest that the difference between these two measurement instruments can be expected to vary by up to ±11 degrees. The results support the interchangeable use of goniometry and digital inclinometer for measuring scaption. Clinicians and researchers should consider the MDC values presented when interpreting change during subsequent measurement sessions.

INTRODUCTION

The assessment of mobility is an integral component of a physical examination. The examination of joint integrity and mobility is necessary to select appropriate physical therapy interventions (American Physical Therapy Association, 2003). Recognizing impairments in joint mobility may assist clinicians in making diagnoses, measuring improvements or deteriorations in mobility, and in determining functional limitations. Therefore, it is essential for clinicians to have reliable and valid measurement instruments to objectively monitor disease progression, outcomes, and mobility impairments.

The examination of shoulder mobility may be accomplished by using a number of instruments, including visual observation, goniometry, linear measures, and inclinometry (Clarkson, 2005). The method and type of assessment will vary among clinicians and institutions on the basis of factors such as time, the educational inclination of the clinician, availability of equipment, and the specific movement or tissue being assessed. Goniometry has been used widely because of its portability and low cost (Gajdosik and Bohannon, 1987; Laupattarakasem et al, 1990). However, a limitation of goniometry is that it requires the clinician to use both hands, making stabilization of the extremity more difficult and thus increasing the risk of error in reading the instrument (Gajdosik and Bohannon, 1987). Inclinometry is another practical alternative that incorporates the use of constant gravity as a reference point to assess joint mobility (Bovens et al, 1990; de Jong, Nieuwboer, and Aufdemkampe, 2007; Laupattarakasem et al, 1990). Digital inclinometers are portable, lightweight, and require training similar to that of goniometry. However, a disadvantage of digital inclinometry may lie in the fact that it is more costly than conventional
goniometers and requires the examiner to establish the zero point of the digital inclinometer accurately and consistently prior to use to minimize the risk of measurement errors.

Shoulder elevation in the scapular plane (hereafter referred to as scaption) involves elevation of the humerus in an intermediate plane between pure shoulder flexion and abduction. Most upper extremity functional activities of daily living are rarely performed solely in pure cardinal planes. Scaption is an important movement to consider when treating shoulder pathologies because it incorporates functional movement patterns that are a part of the normal biomechanics of the shoulder complex (Youdas, Carey, Garrett, and Suman, 1994). In addition, following labral repairs and reverse shoulder arthroplasty, rehabilitation protocols often dictate movement to take place in scaption because it provides the least amount of capsuloligamentous tension of the surgically repaired structures (Blackburn and Guido, 2000; Boudreau, Boudreau, Higgins, and Wilcox, 2007). Lastly, incorporation of treatment interventions in the scapular plane are often advocated; thus, it is essential that clinicians and researchers possess a reliable means to quantify scaption.

To our knowledge, a paucity of research exists to describe the intrarater reliability, interrater reliability, and concurrent validity of goniometric and inclinometric measurements of scaption, despite its widespread use in examinations and interventions. To date, only one study has been conducted on the reliability of goniometric measurements of scaption. Youdas, Carey, Garrett, and Suman (1994) investigated the reliability of goniometric measurements of active scaption in 43 subjects (32 females and 11 males; age range 27–82 years old) with orthopedic disorders of the shoulder or neck. Ten physical therapists with 1–29 years of clinical experience obtained repeated measurements of 43 subjects tested in standing with a handheld scapulohumeral goniometer (SHG) positioned over the subject’s scapula and posterior shoulder. The study results concluded that for both scapular and glenohumeral rotation, 50% of the time, the first and second measurements made by the same therapist differed by 3 degrees or more and were considered to be too wide of a distribution to be clinically reliable. Ten percent of the time, these measurements differed by 8 degrees or more (Youdas, Carey, Garrett, and Suman, 1994). The researchers of the study concluded that the intrarater reliability of active scaption achieved when measured with the SHG was poor because many of the data points on the x-y coordinate system lay a vertical distance of at least 10 degrees from the line of equality (Youdas, Carey, Garrett, and Suman, 1994). A limitation of the study is that the methods are not clinically reproducible and reliability coefficients were not provided.

Given the lack of available research investigating the reliability and concurrent validity of goniometry and inclinometry measurements of scaption, further investigation is essential to provide clinicians and researchers with the necessary information needed to make clinical decisions about their interchangeability. Whether goniometry and inclinometry can be used interchangeably for measuring scaption has yet to be investigated. Therefore, the purpose of this study was to investigate the intrarater reliability, interrater reliability, and the concurrent validity of digital inclinometry and goniometry for measuring active shoulder scaption.

**METHODS**

**Subjects**

Thirty asymptomatic adult participants, 9 males and 21 females, were recruited from a local university setting. Participants who met study requirements were provided with an informed consent document approved by the Institutional Review Board at Nova Southeastern University, and all questions were answered to their satisfaction prior to commencing data collection.

Participants completed a questionnaire to report age, height, body mass, and arm dominance. Exclusion criteria consisted of reported cervical spine or upper extremity pain at the time of data collection or recent shoulder surgery on the dominant arm for which the subject was still receiving care. The mean age, height, body mass, and arm dominance. Exclusion criteria consisted of reported cervical spine or upper extremity pain at the time of data collection or recent shoulder surgery on the dominant arm for which the subject was still receiving care. The mean and standard deviation (SD) for the participants’ age, body mass, and height were 26 (4.2) years, 70 (11.3) kilograms, and 170 (8.1) centimeters, respectively. Testing was conducted on the dominant arm. The right arm was dominant in 26 of the 30 subjects.

**Instruments**

A standard plinth and Acumar™ digital inclinometer (model ACU 360), Lafayette Instrument Company (Lafayette, IN) was used for all inclinometric measurements (Figure 1). The manufacturer’s specifications indicate that this instrument is capable of measuring a range up to 180 degrees with an accuracy of ±1 degree. A 12-inch plastic BASELINE® goniometer (model 12-1000), Fabrication Enterprises (White Plains, NY) was used for all goniometric measurements (Figure 2).
Procedures

Following completion of paperwork and consent, individuals who agreed to participate were brought into a private testing laboratory where they performed a standardized warm-up supervised by doctoral physical therapy students in their 3rd year, having completed all measurement-related coursework. Both students taking measurements performed 6 hours of practice trials with the digital inclinometer before beginning the experiment in addition to exposure they received in their curriculum.

Measurements were performed in an intra- and intersession design. Rater A and rater B each took both goniometric and inclinometric measurements on day 1 of data collection for the interrater component of the study, whereas only rater B took measurements on the 2nd day for the intrarater component of the study. The raters were blinded to the results, as an independent third student with similar experience in goniometry and inclinometry recorded all data. Prior to performing measurements, the digital inclinometer was zeroed by using a fixed vertical reference point to ensure accuracy. For the validity component of the study, measurements were taken with the goniometer and digital inclinometer simultaneously in one motion of scaption and then repeated with the mean value used for analysis.

Prior to obtaining measurements participants completed a warm-up that required approximately 3 minutes to complete and consisted of pendulums and standing shoulder blade squeezes. Each participant was required to perform the same warm-up for consistency; however, to our knowledge there is no benefit or detriment to performing the warm-up. Following the warm-up, participants were seated in a chair with back supported below the scapula, feet flat on the floor, and trunk upright. Prior to taking measurements, rater B passively moved the subject’s dominant arm through one repetition of scaption with the thumb pointed up toward the ceiling to full end range. The angle of scaption was visually estimated at 30–45 degrees anterior to the coronal plane by the rater to honestly replicate procedures that occur in the clinical setting. The purpose of the passive trial repetition was to familiarize the participant with the requested motion.

After passive motion by rater B, participants were asked to actively move through the motion to end range. Verbal and manual cues were used to correct motion on a strictly as-needed basis. Once the participant reached end range in correct form, rater B started the first series of measurements. The order in which the measurements were taken was goniometer first and digital inclinometer second. The goniometer was positioned with the fulcrum placed at the midpoint of the central aspect of the glenohumeral joint, the stable arm parallel to the trunk, and the moving arm parallel to the longitudinal axis of the humerus (Figure 3). Once a goniometric measurement was taken, the participant was asked to hold the position and the digital inclinometer was placed in position for measurement. The digital inclinometer was positioned on the superior portion of humeral shaft proximal to the elbow (Figure 4).

After completion of the measurements, the participant was asked to return the arm to the side. From the resting position, the participant was then asked to raise the arm again in the scapular plane to full end range where goniometric and inclinometric measurements were repeated by rater B for a total of two measurements. This concluded the first set of measurements and was followed by a 5-minute rest for the participant. After the 5-minute rest, rater A took two series of measurements in the same order as rater B. The measurements from session 1 took approximately 30

![FIGURE 1 Acumar digital inclinometer (model ACU 360), Lafayette Instrument Company, Lafayette, IN.](image1)

![FIGURE 2. Standard BASELINE 12-inch plastic goniometer (model 12-1000), Fabrication Enterprises, Inc., White Plains, NY.](image2)
minutes from the initiation of the warm-up to completion. Participants were asked to return in 24 hours for the 2nd day of testing by rater B to complete the intrarater reliability measurements. Raters remained blinded to both their results as well as the other rater’s results throughout the investigation.

**Statistical methods**

Data analysis was performed with SPSS version 15.0 for Windows statistical program. Descriptive data including mean measurement angles with standard deviations (SD) were calculated for each session. The reliability of scaption was determined by the intraclass correlation coefficient (ICC) model 3, k for the intrarater component of analysis and model 2, k for the interrater analysis. The mean value from each testing session was used for the analysis. Model 3, k was used for the intrarater analysis because rater B was the only tester of interest. Model 2, k was used for the interrater analysis to determine if the instrument of choice can be used with confidence and reliability among equally trained clinicians (Portney and Watkins, 2009; Shrout and Fleiss, 1979).

Interpretation of ICC values was based on guidelines offered by Portney and Watkins (2009), where a value above 0.75 was classified as good reliability and those below 0.75 were classified as moderate to poor reliability. ICC values may be influenced by intersubject variability of scores, because a large ICC may be reported despite poor trial-to-trial consistency if the intersubject variability is too high (Portney and Watkins, 2009; Weir, 2005). The standard error of measurement (SEM) is not affected by intersubject variability (Weir, 2005). Therefore, SEM was reported in conjunction with the ICCs using the formula: $SEM = SD \sqrt{1 - r}$ (Portney and Watkins, 2009). The minimal detectable change (MDC) was calculated by using the formula: $MDC_{95} = 1.96 \times SEM \times \sqrt{2}$ to determine the magnitude of change that would exceed the threshold of measurement error at 95% confidence level (Portney and Watkins, 2009). MDC values were rounded to the nearest degree on the tables to reflect the smallest unit of measurement on the goniometer and digital inclinometer.

An ICC model 3, k was used in the concurrent reliability analysis to determine if both methods of measurement analysis produced comparable results. ICC value interpretations were based on the
aforementioned guidelines established by Portney and Watkins (2009). Bland-Altman plots were used to provide a visual illustration of the relationships and 95% limits of agreement between the goniometric and inclinometric measurements (Portney and Watkins, 2009) for each rater. The 95% limits of agreement were calculated by using the formula: 95% limits of agreement = mean difference +/- 2SD (Portney and Watkins, 2009).

### RESULTS

Measurements obtained from the intrarater and interrater reliability analysis of scaption, including the mean, SD, ICC value with 95% CI, SEM, and MDC95, are presented in Tables 1 and 2, respectively. The concurrent validity between goniometry and digital inclinometer measurements of scaption are presented in Table 3. Bland-Altman plots of the measurements suggest no bias for raters A (Figure 5a) and B (Figure 5b) because difference scores are equally represented above and below the zero difference line. The 95% limits of agreement suggest that the difference between measurement of scaption using goniometry and inclinometry varied between −10 degrees and 8 degrees for rater A and −11 degrees and 11 degrees for rater B.

### DISCUSSION

When adhering to the procedures outlined in this investigation, measurements taken for scaption using both the inclinometer and goniometer possessed good intrarater and interrater reliability. Moreover, measurements with a digital inclinometer were found to be comparable to those taken with the standard 12-inch plastic goniometer. For agreement, clinicians and researchers should recognize that the difference between these two measurement instruments can be expected to vary by up to ±11 degrees. Unfortunately, no previous research has been reported in the area of concurrent validity between goniometric and digital inclinometric measurements in the scapular plane for us to compare our results. Research has been performed by using a gravity inclinometry to assess other shoulder motions. However, the majority of existing research only assessed shoulder flexion, abduction, external rotation (ER), and internal rotation (IR). Only flexion and abduction will be used for reliability comparisons because they are both overhead shoulder motions and are therefore more closely related to scaption.

Hoving et al (2002) evaluated shoulder flexion AROM with a gravity inclinometer on six participants with pain; intrarater reliability of shoulder flexion was found to be good with ICC = 0.83, whereas interrater

| Table 1 Intrarater reliability of goniometer and inclinometer (rater B).  
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Session 1 Mean angle(^\circ) (SD)</th>
<th>Session 2 Mean angle(^\circ) (SD)</th>
<th>ICC 3,k (95% CI)</th>
<th>SEM(^\circ)</th>
<th>MDC(^\circ)(_{95})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goniometer</td>
<td>159 (9.1)</td>
<td>159 (11.1)</td>
<td>0.87 (0.74–0.94)</td>
<td>3.6</td>
<td>10</td>
</tr>
<tr>
<td>Inclinometer</td>
<td>158 (9.2)</td>
<td>159 (10.5)</td>
<td>0.88 (0.75–0.94)</td>
<td>3.4</td>
<td>9</td>
</tr>
</tbody>
</table>

SD: standard deviation; ICC: intraclass coefficient; SEM: standard error of measurement; and MDC\(_{95}\): minimum detectable change at the 95% confidence level.

| Table 2 Interrater reliability of goniometer and inclinometer.  
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Rater A Mean angle(^\circ) (SD)</th>
<th>Rater B Mean angle(^\circ) (SD)</th>
<th>ICC 2,k (95% CI)</th>
<th>SEM(^\circ)</th>
<th>MDC(^\circ)(_{95})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goniometer</td>
<td>160 (11.4)</td>
<td>159 (9.1)</td>
<td>0.92 (0.83–0.96)</td>
<td>2.9</td>
<td>8</td>
</tr>
<tr>
<td>Inclinometer</td>
<td>160 (10.9)</td>
<td>158 (9.2)</td>
<td>0.89 (0.77–0.95)</td>
<td>3.4</td>
<td>9</td>
</tr>
</tbody>
</table>

SD: standard deviation; ICC: intraclass correlation coefficient; SEM: standard error of measurement; and MDC\(_{95}\): minimum detectable change at the 95% confidence level.

| Table 3 Concurrent reliability of goniometry and digital inclinometry.  
<table>
<thead>
<tr>
<th>Rater</th>
<th>ICC(3,k)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.94</td>
<td>0.87–0.97</td>
</tr>
<tr>
<td>B</td>
<td>0.94</td>
<td>0.86–0.97</td>
</tr>
</tbody>
</table>

ICC: intraclass correlation coefficient; CI: confidence interval.
reliability was found to be moderate at ICC = 0.72. All other shoulder motions tested (abduction, IR, and ER) were shown to have moderate to low interrater and intrarater reliability. The authors suggest the raters’ limited experience with gravity inclinometry may have negatively affected the outcomes of the study (Hoving et al, 2002). Valentine and Lewis (2006) reported reliability coefficients ranging from ICC = 0.82–0.91 for shoulder flexion to ICC = 0.88–0.91 for shoulder abduction using gravity inclinometry.

Several studies have looked at the reliability of goniometry and digital inclinometry separately to measure PROM of the shoulder. Intrarater reliability values for goniometric measurements of passive shoulder flexion and abduction ranged from ICC = 0.82–0.98 to ICC = 0.84–0.98, respectively (Pandya et al, 1985; Riddle, Rothstein, and Lamb, 1987; Sabari et al, 1998). Interrater reliability of digital inclinometry and goniometry of passive shoulder flexion and abduction had a larger range than previously mentioned. Reliability coefficients ranged from ICC = 0.73–0.99 for passive shoulder flexion to ICC = 0.15–0.87 for passive shoulder abduction (de Jong, Nieuwboer, and Aufdemkampe, 2007; de Winter et al, 2004; Pandya et al, 1985; Riddle, Rothstein, and Lamb, 1987). Because of the large variance of reliability coefficients reported for passive shoulder flexion and abduction, it is difficult to determine whether the results from these studies can be compared to those produced by studies analyzing AROM measurements that are in a different movement plane.

Another issue that warrants further discussion is the concern over the higher interrater reliability vs. intrarater reliability, as was found in this study. Generally speaking, one might expect intrarater reliability to exceed interrater reliability. However, this was not the case in this study. Goniometry interrater reliability of scapular plane elevation (ICC = 0.92), as established in this study, was higher than goniometry intrarater reliability (ICC = 0.88). Similarly, digital inclinometry interrater reliability of scapular plane elevation (ICC = 0.89) was higher than digital inclinometry intrarater reliability (ICC = 0.87). The intrarater component of the investigation was completed by using an intersession design; beyond this we cannot currently make an assumption about why interrater reliability was greater than intrarater reliability for both measurement techniques.

For measurement error and what might constitute true change, the MDC$_{95}$ for intrarater analysis indicated that a change equal to or greater than 10 degrees for goniometry and 9 degrees for inclinometry is required to be 95% certain change is not due to intratrial variability or measurement error. The MDC$_{95}$ for the interrater analysis indicated that a change equal to or greater than 8 degrees for goniometry and 9 degrees for inclinometry is required to be 95% certain that the change is not due to intertrial variability or measurement error. Similar to our discussion of reliability and validity, there are no previous studies for which our results may be compared.

This study was the first to analyze the interrater and intrarater reliability, MDC, and concurrent validity of goniometric and digital inclinometric measurements of scaption. Because of the lack of research in this area, a comparison between our study and previous research cannot be made. However, this study does set the groundwork for further research in this area to evaluate the interchangeability, reliability, and MDC.
of goniometric and digital inclinometric measurements of scapular plane elevation.

Limitations and future research

When interpreting the reliability values in our investigation, one must recognize that range of motion consistency in individuals with healthy shoulders may not correlate with those who have shoulder pathology. Triffitt, Wildin, and Hajioff (1999) assessed the reproducibility limits of inclinometric shoulder abduction and external rotation in symptomatic and asymptomatic subjects. Asymptomatic subjects had a difference ranging from 24 to 33 degrees for all measurements, compared to 24 to 41 degrees in symptomatic subjects, suggesting a greater variance among those with a painful shoulder. Although we cannot state with certainly that this would be the case with scaption, it is an issue requiring consideration.

In addition, when considering the symptomatic population, the ability of an individual to achieve and maintain his or her arm in a specific plane may be compromised secondary to restrictions in joint mobility and/or pain. This may result in changes in the actual motion being assessed (i.e., shoulder flexion vs. scaption vs. abduction) on the basis of the patient’s ability to maintain static positioning in the desired plane. Therefore, further research in the interchangeability of goniometry and digital inclinometry for measuring elevation in the scapular plane among symptomatic patients is warranted.

The subjects in this investigation consisted of a young, college-aged population (mean age = 26). The average age of individuals seen in the clinical setting is 44 years (Di Fabio and Boissonnault, 1998). Therefore, our results may not necessarily be generalized to a subgroup with increasing or decreasing age.

When comparing instruments such as goniometry and inclinometry, it is important to consider limitations to both instruments. The inclinometer uses a fixed vertical reference point realized by gravity, thus is stable, provided the zero point is accurately calibrated and established. Traditional goniometry requires visualization of the vertical reference point, which may compromise measurement reproducibility. Another issue that warrants discussion is the effect of body types (ectomorphs vs. endomorphs and mesomorphs) on digital inclinometer measurements. The inclinometer measurements were taken at the midline of the humerus on the superior portion of the biceps brachii. Ectomorphic body types tend to have a linear frame with sparse muscular development, whereas mesomorphs and endomorphs have a greater amount of tissue surrounding their frame (Thomas, 1993). Digital inclinometer measurements may differ from goniometric measurements because of the differences in tissue mass. In our opinion, for those who desire greater portability and require time efficiency, using a digital inclinometer may be worth the greater cost and time required to familiarize oneself with the instrument.

When interpreting change scores, it should be recognized that the MDC is not the same as the minimum clinically important difference (MCID) (Portney and Watkins, 2009). The MCID is the amount of change that is clinically meaningful and is typically associated with an external criterion that indicates when meaningful change has occurred (Portney and Watkins, 2009). The MDC reported in this investigation is the smallest amount of change that can be considered above the threshold of error (Portney and Watkins, 2009). However, one must not make the assumption that this change has reached the threshold of clinically meaningful improvement.

CONCLUSION

This investigation is the first of its kind to evaluate the reliability and concurrent validity of digital inclinometric and goniometric measurements of arm elevation in the scapular plane. When used with the measurement procedures outlined in this investigation, both techniques are reliable, as evidenced by reliability coefficients that approach 0.90 (the threshold recommended for making clinical decisions) (Portney and Watkins, 2009). Excellent concurrent validity statistics were produced as well; however, one should recognize the ranges of disagreement presented. Clinicians should consider the MDC values presented when interpreting change values during subsequent measurement sessions to be certain that the change is not due to intertrial variability or measurement error.

Declaration of Interest: The author has no conflicts of interest. The author alone is responsible for the content and writing of the article.

REFERENCES

Clarkson HM 2005 Joint motion and function assessment: A research based practical guide. Philadelphia, PA, Lippincott Williams & Wilkins


de Winter AF, Heemskerk MA, Terwee CB, Jans MP, Deville W, van Schaardenburg DJ, Scholten RJ, Bouter LM 2004 Inter-observer reproducibility of measurements of range of motion in patients with shoulder pain using a digital inclinometer. BMC Musculoskeletal Disorders 5: 18


Copyright © Informa Healthcare USA Inc.