Intertester Reliability and Concurrent Validity of Fluid-based and Universal Goniometers for Active Knee Flexion

WENDY RHEAULT, MICHELLE MILLER, PAULA NOTHNAGEL, JOHN STRAESSLE, and DENISE URBAN

This study was designed 1) to determine the intertester reliability of the universal goniometer and the fluid-based goniometer and 2) to establish the concurrent validity of the fluid-based goniometer. Two testers measured active knee flexion of 20 healthy subjects (15 women, 5 men) using both instruments. The subjects had a mean age of 24.8 years (s = 5.5) and reported no previous history of pathological conditions of the right lower extremity. Using a correlational design, high intertester reliability was established for each instrument (universal goniometer, r = .87; fluid-based goniometer, r = .83). The concurrent validity of the fluid-based goniometer was also good for both testers (tester A, r = .83; tester B, r = .82). When the data were subjected to t tests, significant differences were found between the instruments for each tester (p < .05). The results suggest that similar measurements will be obtained between therapists using the universal and fluid-based goniometers; however, the two instruments cannot be used interchangeably in the clinic.

Key Words: Knee joint; Lower extremity, knee; Muscle performance, measurement; Tests and measurements, range of motion.

Goniometry has been used for years by physical therapists in an attempt to quantify joint range of motion. Unfortunately, these measurements are often inconsistent. These inconsistencies can be partially due to intertester differences in measurement technique, including differences in patient and goniometric positioning.1-3

The universal (standard) goniometer is currently used in many physical therapy clinics. A less commonly used instrument, the fluid-based goniometer, works on the principle of the carpenter's level with a fluid-filled circular chamber. The fluid-based goniometer has advantages over the universal goniometer in that it is not subject to errors caused by mistaking anatomical landmarks and requires only one hand for use.4 Disadvantages of the fluid-based goniometer include its inability to measure outside the fluid's straight-plane movement in gravity. Variation in measurement can also occur because of the placement of the instrument on soft tissue.5 For the fluid-based goniometer to be useful clinically, its reliability and concurrent validity must be established.

LITERATURE REVIEW

Many articles have been published on the intratester and intertester reliability of the universal goniometer. Boone and Azen found that intratester reliability was greater than intertester reliability.2 Mitchell et al, however, found that the intertester and intratester reliability were both high when using a standardized procedure for measurement.6 Some evidence also exists that the reliability of upper extremity measurements may be greater than lower extremity measurements.2

The validity of the universal goniometer for the knee has been established by Gogia et al.6 They compared goniometric measurements of the knee to measurements derived from roentgenograms and reported high correlation coefficients for validity (r = .97-.98).6

With the introduction of an alternate goniometric instrument, such as the fluid-based goniometer, it is necessary to compare the new instrument to the accepted standard. A recent study by Petherick et al was conducted to determine the intertester reliability and concurrent validity of the fluid-based goniometer for the elbow.7 Although a correlation of .83 existed for concurrent validity, there was a significant difference between the two instruments. Intertester reliability for the fluid-based goniometer was high (R = .92), whereas intertester reliability of the universal goniometer was poor (R = .53).

This particular study by Petherick et al7 prompted us to conduct a similar investigation at the knee joint. The purpose of our study was 1) to determine the intertester reliability of the universal and fluid-based goniometers for active knee flexion and 2) to establish the concurrent validity of the fluid-based goniometer. We measured knee flexion because it is a commonly measured joint

W. Rheault, MA, PT, is Chairman and Associate Professor, Department of Physical Therapy, University of Health Sciences-The Chicago Medical School, 3333 Greenbay Rd, Bldg 51, North Chicago, IL 60064 (USA).

Ms. Miller, BS, Ms. Northnagel, BS, Ms. Urban, BS, and Mr. Straessle, BS, were physical therapy students at the University of Health Sciences-The Chicago Medical School when this study was conducted.

This article was submitted September 24, 1987; was with the authors for revision 30 weeks; and was accepted May 23, 1988. Potential Conflict of Interest: 4.

PHYSICAL THERAPY
in physical therapy. Knee extension was not measured because we wanted to restrict the scope of the study to one joint motion. We hypothesized that there would be no difference between the universal and fluid-based goniometers in measuring active knee flexion and that the intertester reliability of these two instruments would be high.

METHOD

Subjects

Our subjects were 20 volunteers (15 women, 5 men) with a mean age of 24.8 years (s = 5.5). It should be noted that all of our subjects were healthy individuals and, thus, the generalizability of our results to a patient population may be limited. Because our aim was to explore the actual characteristics of the two instruments, however, we felt that an available sample of healthy students would be appropriate. The subjects had no previous history of musculoskeletal or neurological injury to the right lower extremity. Each subject signed an informed consent statement that was approved by the university’s institutional review board.

Instrumentation

The universal goniometer was a plastic 360-degree goniometer* with 10-in movable arms. The fluid-based goniometer had a plastic fluid-filled circular chamber marked in 1-degree increments. The fluid contained within this goniometer responded to changes in motion. We hypothesized that there would be no difference between the universal and fluid-based goniometers in measuring active knee flexion and that the intertester reliability of these two instruments would be high.

Alignment of the universal goniometer with the knee in full extension was performed according to standard procedure. The stationary arm was aligned along the shaft of the femur, pointing in the direction of the greater trochanter. The movable arm was aligned parallel to the fibula, pointing to the lateral malleolus. The axis of the goniometer was placed at the joint line. Although the center of knee motion is not at the joint line, we put the axis on the joint line to facilitate measurement.

The fluid-based goniometer was positioned 10 in distal to the middle of the popliteal crease. The scale was set at 0 degrees in full knee extension and read at the bottom of the fluid meniscus in full knee flexion. This position was chosen by trial and error. Both instruments were checked to ensure that they were in working order. Testers agreed on their measuring technique before the study, without engaging in extensive practice.

Subjects were asked to actively flex and extend the right knee five times before measurement as a warm-up. Each of the two investigators (J.S., P.N.) recorded four measurements of active knee flexion on each subject, two measurements with the fluid-based goniometer and two with the universal goniometer.

To prevent bias, all test sessions were completed separately from the other investigator. All data were collected in the same laboratory and at the same time of day over a three-day period.

Data Analysis

The average of each subject’s measurements was computed for each instrument and tester. These values were then used to compute Pearson product-moment correlation coefficients. The same data were used for the paired t tests, because correlations may be high even if there are significant differences between the instruments and testers.

RESULTS

Means and standard deviations were calculated for both goniometers (Tab. 1). Pearson product-moment correlations for intertester reliability and concurrent validity are shown in Table 2. The results of the t tests demonstrated a significant difference between testers (Tab. 3) and between the two instruments for each tester (Tab. 4).

DISCUSSION

According to Currier, a correlation of .80 to 1.00 is considered very reliable. Although high correlation coefficients were established for both intertester reliability and concurrent validity when using the Pearson product-moment correl...

---

* Distributed by Chattanooga Corp, PO Box 4287, Chattanooga, TN 37405.
† 1 in = 2.54 cm.
relations, significant differences ($p < .05$) were noted when the raw data were subjected to the $t$ tests. The Pearson product-moment correlation coefficient can be high, even if scores are inconsistent, as long as they vary together.

The measurements obtained by the two testers with each device were significantly different according to the $t$-test results; however, the mean difference in degrees between tester A and tester B for the universal goniometer was only $2.48$. According to a study by Boone and Azen, a change of 5 degrees for upper extremity measurement and 6 degrees for lower extremity measurement is acceptable when more than one tester is taking measurements; however, our results suggest that a difference of 2 to 3 degrees may be statistically significant.

The intertester reliability of the universal and fluid-based goniometers was .87 and .83, respectively. Rothstein et al also reported good intertester reliability for the standard goniometer. Petherick et al, however, concluded that although the fluid-based goniometer had high intertester reliability ($R = .92$), the universal goniometer had poor reliability ($R = .53$) when measuring elbow ROM in healthy subjects. Although our results are inconsistent with those of Petherick et al, it should be noted that both studies were conducted in the same laboratory and with the same instruments. It is evident that replications of both studies are needed.

Petherick and associates found that the Pearson product-moment correlation between the fluid-based goniometer and the universal goniometer was .83. This value is in excellent agreement with our results. We found that the concurrent validity for tester A and tester B was .83 and .82, respectively.

A "practiced" technique for goniometric measurement was not used in our study. The anatomical landmarks were predetermined, but the testers were permitted to use their own technique. Rothstein et al did not use a standardized technique in their study and found both intertester and intratester reliability to be high. Other researchers, however, have found that intertester reliability for normal hip ROM measurement improved because of standardized measurement.

We have several suggestions for future research. First, we suggest that this study be conducted on patients with known pathological conditions. We also suggest that different joints be used in a study comparing the universal and fluid-based goniometers. Replication with more than two testers taking the joint measurements may provide interesting information. Last, because few studies have investigated the validity of goniometric tools, we feel that this is also an important area of study.

**CONCLUSION**

Although the Pearson product-moment correlation coefficients demonstrated a high correlation between measurements from the two goniometers, paired $t$ tests failed to establish concurrent validity between the fluid-based and universal goniometers. Thus, we conclude that therapists cannot use these two goniometric instruments interchangeably. Although the intertester reliability was high for both instruments, a statistical difference may exist between testers. The intertester reliability was found to be slightly higher with the universal goniometer than with the fluid-based goniometer; however, it could be argued that the fluid-based goniometer is more convenient to use. The fluid-based goniometer is portable, lightweight, and easy to apply, allowing one hand free for moving the patient's limb.

It is important to note that this study was conducted with healthy subjects, thus limiting the generalizability of the results to the clinical setting. To determine the clinical value of the fluid-based goniometer, further research with patients must be conducted to establish its reliability and validity.

**REFERENCES**