Research Report

Reliability of Measurements of Cervical Spine Range of Motion—Comparison of Three Methods

To determine reliabilities within and between persons measuring cervical active range of motion (AROM), three methods were examined: use of a cervical-range-of-motion (CROM) instrument, use of a universal goniometer (UG), and visual estimation (VE). Measurements were made on 60 patients with orthopedic disorders of the cervical spine who were divided into three groups of 20 subjects each. All subjects were tested in a standardized seated position using operationally defined goniometric placements and nongoniometric estimation techniques. Cervical flexion and extension, lateral flexion, and rotation were measured. Intraclass correlation coefficients (ICCs) were used to quantify within-tester and between-tester reliability. We found that goniometric measurements of AROM of the cervical spine made by the same physical therapist had ICCs greater than .80 when made with the CROM device or the UG. When different physical therapists measured the same patient's cervical AROM, the CROM device had ICCs greater than .80, whereas the UG and VE generally had ICCs less than .80. [Youdas JW, Carey JR, Garrett TR. Reliability of measurements of cervical spine range of motion—comparison of three methods. Phys Ther. 1991;71:98-106.]

Key Words: Cervical vertebrae; Spine; Tests and measurements, range of motion.

Disorders of the cervical spine often alter the normal active range of motion (AROM) of the neck. The response of a patient with neck pain to therapeutic intervention is often documented clinically by measuring or visually estimating changes in the AROM of the cervical spine. Although rehabilitation personnel frequently measure patients' available AROM, the quality of the technique and the consistency of the measurements often are taken for granted.1

Cole2 argued that clinical measurements of cervical spine motion are the least accurate of the common measurements of the mobility of the body's joints. He attributed this lack of accuracy to the lack of bony landmarks on the head and to the thickness of soft tissues overlying segments of the cervical spine. Moll and Wright3 stated that visual estimation (VE) is the most popular technique for recording spine mobility, although this was their personal opinion and they provided no supporting data. Moore,4 however, discouraged use of VE because she believed it did not lead to reliable measurements. Although VE has been maligned as a subjective measurement of joint motion, to our knowledge no evidence has been published to document between-tester reliability of the VE technique for measuring cervical range of motion (ROM) in patients with neck pain.

The full-circle goniometer, or universal goniometer (UG), is a versatile device for recording measurements of peripheral joint ROM in healthy sub-

James W Youdas
James R Carey
Tom R Garrett

1 Youdas, MS, PT, is Physical Therapist, Physical Therapy Program, Mayo School of Health-Related Sciences, and Assistant Professor of Physical Therapy, Mayo Medical School, 200 First St SW, Rochester, MN 55905. Address all correspondence to Mr Youdas at Mayo Clinic, 200 First St SW, Rochester, MN 55905 (USA).

2 Carey, PhD, PT, is Senior Associate Consultant, Physical Therapy Program, Mayo School of Health-Related Sciences, and Assistant Professor of Physical Therapy, Mayo Medical School.

3 T Garrett, BS, PT, is Physical Therapist, Physical Therapy Program, Mayo School of Health-Related Sciences, and Assistant Professor of Physical Therapy, Mayo Medical School.

This research project was approved by the Mayo Institutional Review Board.

This article was submitted January 23, 1990, and was accepted September 24, 1990.

Physical Therapy/Volume 71, Number 2/February 1991 98/23
The purposes of this study were (1) to examine, in patients, the within-tester and between-tester reliabilities of measurements of AROM of the cervical spine in the three cardinal planes using the CROM device, the UG, and VE and (2) to compare the measurement errors of the three techniques.

**Method**

**Subjects**

The subjects were 60 patients (39 women, 21 men) referred to the Mayo Clinic Department of Physical Medicine and Rehabilitation (Rochester, Minn) with orthopedic disorders (Tab. 1). The patients' ages ranged between 21 and 84 years (X=59.1, SD=15.7), and the most frequent reason for physical therapy evaluation was cervical muscle pain (n=27, 45%). Descriptive information regarding the AROM of the patients for the six cervical motions is given in Table 2. Criteria for admission to the study were (1) that assessment of cervical AROM was an appropriate component of the patient's routine physical therapy evaluation and (2) that the patient was 18 years of age or older. Any patient with the diagnosis of spasmodic torticollis was excluded. According to the patients' self-reports, repeated cervical motions did not exacerbate their clinical signs or symptoms. Informed oral consent was obtained from all subjects.

Using previously reported studies as a model, we divided the data-collection process into three parts to minimize patient discomfort and to maintain the physical therapy department's productivity. Each part consisted of an assessment of cervical AROM in 20 subjects in one cardinal plane of motion (Tab. 1).

**Testers**

The CROM, UG, and VE measurements were made by 11 volunteer physical therapists employed at the Mayo Clinic. Their clinical experience ranged from 2 to 27 years (X=7, SD=7), and they had graduated from three different schools of physical therapy. Prior to the study, the therapists' self-reports indicated that they generally used VE...
arms and sturdy pivot joints. The measurement scales of the UGs were marked in 1-degree increments. One side of each goniometer's scale was covered with white adhesive paper so that the physical therapist could not see the measurement, but the recorder could read the numbers from the reverse side of the goniometer and record them on prepared recording sheets. The UGs were not tested for their individual measurement accuracy.

The CROM device (Fig. 1) consists of a plastic frame that is mounted over the subject's nose bridge and ears and secured to the head by a Velcro\textsuperscript{TM} strap. Three dial angle meters attached to the frame and arranged orthogonally to one another indicate the subject's cervical ROM. Neck flexion, extension, and lateral flexion movements are recorded by gravity goniometers. The cervical rotation (transverse plane) meter is a compass goniometer and operates in conjunction with a shoulder-mounted magnetic yoke. The dial meters were marked in 2-degree increments. The same CROM instrument was used throughout the study. The device was not checked for measurement accuracy.

**Figure 1. Cervical-range-of-motion (CROM) instrument with magnetic yoke.** Dial meters A and B are gravity goniometers that indicate cervical motion in the sagittal and frontal planes, respectively. Dial C is a compass that operates in conjunction with the shoulder-mounted magnetic yoke to indicate motion in the transverse plane.

for measuring cervical ROM and the UG for measuring ROM of the extremities; they had no previous experience with the CROM device. Before the study began, we conducted a 60-minute training session using a written protocol that described each method of measurement. The standardized measurement procedures were demonstrated, and the therapists practiced measuring AROM on themselves by using each of the three methods. The measurement protocol for the UG was modified from a previously reported technique.\textsuperscript{12}

**Instrumentation**

We used large, plastic, 360-degree UGs\textsuperscript{4} with 30.5-cm (12-in) movable bases, and sturdy pivot joints. The measurement scales of the UGs were marked in 1-degree increments. One side of each goniometer's scale was covered with white adhesive paper so that the physical therapist could not see the measurement, but the recorder could read the numbers from the reverse side of the goniometer and record them on prepared recording sheets. The UGs were not tested for their individual measurement accuracy.

The CROM device (Fig. 1) consists of a plastic frame that is mounted over the subject's nose bridge and ears and secured to the head by a Velcro\textsuperscript{TM} strap. Three dial angle meters attached to the frame and arranged orthogonally to one another indicate the subject's cervical ROM. Neck flexion, extension, and lateral flexion movements are recorded by gravity goniometers. The cervical rotation (transverse plane) meter is a compass goniometer and operates in conjunction with a shoulder-mounted magnetic yoke. The dial meters were marked in 2-degree increments. The same CROM instrument was used throughout the study. The device was not checked for measurement accuracy.

**Procedure**

To decrease between-tester variability, we standardized subject position and placement of the measurement devices. All subjects sat in a standard metal-frame chair so that their thoracic spine maintained contact with the chair's backrest and their lumbosacral spine filled the gap between the seat and the backrest. Their feet were positioned flat on the floor, and their arms rested freely at their sides. As instructed by the tester, each subject performed three repetitions of neck AROM (warm-ups) in each direction within a designated cardinal plane to increase compliance of the neck's soft tissues. The tester then measured the subject's cervical AROM, in both directions within a cardinal plane, by each of the three techniques (CROM, UG, and VE) used in a randomized order. Immediately after the first six measurements, the subject repeated the same movements, providing two sets of six measurements for each tester. Except for the warm-ups, the subject completed the same movement sequence for a second tester, who likewise made two sets of six measurements. Thus, each subject completed a total of 30 cervical AROM movements (6 warm-ups + 24 measurements).

**Part 1.** This part of the study investigated cervical flexion with chin tuck and extension with chin elevation. Tester and recorder stood on opposite sides of the subject, with the recorder always lateral to the subject's left shoulder. From this location, the recorder read numbers from the CROM device's sagittal plane GG or from the reverse side of the UG's protractor scale. The starting position for both cervical flexion and extension was assumed after the tester manually adjusted the subject's neck so that the external acoustic meatus-to-base of nares reference line was parallel to the floor. Placement of the UG followed the technique previously described by Norkin and White.\textsuperscript{12} The UG's axis was centered over the external acoustic meatus (Fig. 2 [Top]); the fixed arm was held vertical, while the movable arm was aligned with the meatus-to-base of nares reference line as the subject actively flexed and extended the neck. For the CROM, the tester positioned the subject at the appropriate starting position by using manual and verbal cues, without the aid of the sagittal plane GG. The recorder wrote down both start and end points of the cervical AROM for the CROM device and the UG. For VE, the tester verbally reported to the recorder the cervical AROM to the nearest 5 degrees, based on the landmarks described for placement of the UG.

**Part 2.** We studied AROM of cervical lateral flexion. Each subject bent his or her head and cervical spine first left and then right without elevating his or her shoulder. For the VE and
Figure 2. Alignment of universal goniometer (UG): (Top) Procedure for measuring cervical flexion and extension. The examiner positions the UG axis (A) at the center of the subject's external auditory meatus. The fixed arm (B) is vertical, and the movable arm (C) is aligned parallel to the imaginary line between the external auditory meatus and the base of the nares. The recorder writes down both start and end points of the cervical active range of motion (AROM) from the reverse side of the UG protractor scale. (Middle) Procedure for measuring cervical lateral flexion. The examiner positions the UG axis (A) over the center of the subject's sternal notch. The UG fixed arm (B) is aligned parallel to an imaginary line between the subject's acromion processes; the movable arm (C) is aligned with the center of the subject's nose. The recorder writes down both start and end points of the cervical AROM from the reverse side of the UG protractor scale. (Bottom) Procedure for measuring cervical rotation. The examiner positions the UG axis (A) over the center of the subject's head. The fixed arm (B) of UG is aligned with an imaginary line between the subject's acromion processes; the movable arm (C) is aligned with the tip of the subject's nose. The recorder writes down both start and end points of the cervical AROM from the reverse side of the UG protractor scale.

UG techniques, the subject sat between the tester and the recorder; the tester was in front of the patient (Fig. 2 [Middle]). The tester aligned the fixed arm of the UG parallel with a horizontal reference line between the patient's sternal notch and acromion process; the movable arm was aligned with the midline of the patient's nose. The starting or neutral position was with the arms of the UG perpendicular. After the CROM device was mounted, the tester stood behind the subject and offered instructions and corrected substitution patterns; the recorder stood in front of the subject and recorded readings from the frontal plane angle meter at the start and end points of movement.

Part 3. We studied AROM of cervical rotation. Each subject rotated his or her head first left and then right. For VE and UG measurements, the tester stood behind the subject and gazed at the top of his or her head; the recorder faced the subject. The placement procedure for the UG has been
described by Norkin and White.12 The UG axis was centered on the top of the subject’s head (Fig. 2 [Bottom]); the fixed arm was aligned parallel to an imaginary line between the subject’s acromion processes, and the movable arm was aligned with the subject’s nose. After the CROM device was mounted with its magnetic needle and shoulder yoke, the tester faced the subject and gave instructions without looking at the magnetic dial meter. The recorder remained behind the subject, and recorded the readings of the magnetic needle at the start and end points of the motion.

Data Analysis

Intraclass correlation coefficients (ICC [1,1])13 were calculated to express the reliability of the measurements. We calculated ICCs for within-tester reliability of the UG and CROM device by comparing the first and second measurements made by each tester with the same device (ie, 20 subjects were measured by two testers; thus, 40 paired measurements were recorded for each motion). Because the tester reported his or her measurements to the recorder, we did not evaluate within-tester reliability for the VE technique, because we believe the second measurement was biased by the first measurement.

We calculated ICCs for between-device (parallel-forms) reliability by comparing the first measurements made by each tester with a given instrument with their first measurements made with one of the other instruments (40 paired measurements for each motion). We calculated ICCs for between-tester reliability by comparing the first measurements made by each pair of testers (20 paired measurements were taken for each motion).

No universally acceptable levels have been adopted for correlation coefficients for the purpose of describing the reliability of measurements.14 In the following sections, we use a previously reported scheme for defining the amount of reliability based on our ICC values: .90 to .99, high reliability; .80 to .89, good reliability; .70 to .79, fair reliability; and .69 and below, poor reliability.15

Results

Part 1

With the CROM device, ICCs for within-tester and between-tester reliability were larger compared with the UG for both flexion and extension motions (Tab. 3). The between-tester ICCs with both the CROM device and the UG were larger than the VE ICC for neck flexion and extension. The ICC values for interdevice comparisons (Tab. 4) among the three methods generally demonstrated poor to fair between-device reliability for flexion and extension.

Part 2

For left lateral flexion, within-tester reliability was equal for the CROM device and the UG, whereas the CROM ICC was higher than the UG ICC for right lateral flexion (Tab. 3). The ICC for between-tester reliability was higher for the UG than for the CROM device in left lateral flexion, whereas the CROM ICC was larger than the UG ICC with right lateral flexion. Both the CROM and the UG ICCs were larger than the VE ICC for left and right neck flexion. The ICC values for interdevice comparisons (Tab. 4) among the three methods generally demonstrated poor to fair between-device reliability for both neck lateral flexion motions.

Part 3

The ICCs for within-tester and between-tester reliability were larger with the CROM device than with the UG for both left and right neck rotations (Tab. 3). The CROM ICC was larger than the VE ICC for left and right rotation, whereas the UG ICC was lower than the VE ICC for both rotation motions. The ICC values for interdevice comparisons (Tab. 4) among the three methods generally demonstrated poor to fair between-device reliability for both neck rotation movements.

Discussion

Part 1

The high ICC values indicate that AROM measurements of neck flexion (ICC= .95) and extension (ICC= .90) obtained with the CROM device are highly reproducible when repeated by the same physical therapist (Tab. 3). The UG measurements demonstrated good reliability for both flexion (ICC= .86) and extension (ICC= .83).

We believe that between-tester reliability for AROM measurements of neck flexion and extension (ICC= .86) was good for the CROM device (Tab. 3), whereas the UG demonstrated poor reliability for flexion (ICC= .57) and fair reliability (ICC= .79) for extension. Tucci et al19 likewise reported poor between-tester reliability for neck flexion (ICC= .08), although they reported good reliability for neck extension (ICC= .82) when healthy subjects were measured by two experienced examiners using a UG. With VE, the between-tester reliability for AROM measurements was poor (ICC= .42) for both flexion and extension.

Between-device reliability for AROM measurements of neck flexion (ICC= .65) and extension (ICC= .46) was poor when the CROM device and the UG (Tab. 4) were compared. Generally, we also found poor to fair interdevice reliability for AROM measurements of neck flexion and extension when comparing either of the goniometric devices with the VE.

Part 2

The AROM measurements of left neck lateral flexion demonstrated good reliability (ICCs = .84) when repeated by the same tester using either the CROM device or the UG; however, within-tester reliability for right neck flexion was high (ICC= .92) with the CROM and good (ICC= .85) with the UG (Tab. 3).

Measurements obtained with both goniometric devices demonstrated...
between-tester reliability for both left (ICC = .35) and right (ICC = .21) lateral flexion. With the VE method, between-tester reliability for AROM measurements of neck lateral flexion was poor to fair for both left (ICC = .63) and right (ICC = .70) movements.

We found poor between-device reliability for AROM measurements of left (ICC = .66) and right (ICC = .60) lateral flexion with the CROM device and the UG (Tab. 4). Generally, we also found poor between-device reliability when each goniometric device was compared with the VE (Tab. 4).

### Part 3

The AROM measurements of left (ICC = .90) and right (ICC = .93) neck rotation were highly reliable when repeated by the same physical therapist (Tab. 3). With the VE, within-tester reliability was fair for left rotation (ICC = .78) and high for right rotation (ICC = .90).

Between-tester reliability for AROM measurements of neck rotation with the CROM device ranged from good (ICC = .82) for left rotation to high for right rotation (ICC = .92) (Tab. 3). However, the UG demonstrated poor between-tester reliability for both left (ICC = .54) and right (ICC = .62) rotations. Tucci et al. likewise reported poor between-tester reliability for left (ICC = .60) and right rotation (ICC = .52) with the UG. Cervical AROM measurements obtained by VE had poor reliability for left rotation (ICC = .69) but good reliability for right rotation (ICC = .82).

We found fair between-device reliability (ICC = .72) for AROM measurements of left rotation and good reliability (ICC = .81) for right rotation with the CROM device and the UG (Tab. 4). Generally, repeated AROM measurements of cervical rotation obtained by replacing VE with either the CROM device or the UG demonstrated poor to fair reliability.

### General Comments

When used by the same physical therapist, both the CROM device and the UG demonstrated good to high reliability for AROM measurements of cervical motion in patients with orthopedic disorders. According to our data, the CROM device is preferable to the UG when two physical therapists take repeated measurements of cervical AROM on the same patient.
appears that the CROM device can be mounted consistently by two different therapists without need for locating specific anatomic landmarks. Both the UG and the VE techniques generally gave poor between-tester reliability, even though the therapists used a standard measurement procedure. Additionally, because between-device reliability was poor, a physical therapist should avoid interchanging the CROM device and the UG when performing repeated measurements of cervical AROM on the same patient.

This study was conducted in one clinical outpatient department, so the results may not be generalized to all clinical departments. All patients studied had orthopedic disorders, and repeated cervical movements did not aggravate their neck pain, according to their self-reports. Like other researchers, we found that determination of goniometric reliability is possible even in a busy outpatient setting. We used a previously reported design as a model for our project, however, our design differed from the previous design in that we used a standardized method for testing. By attempting to decrease individual variation in testing method, we sought to account for a large amount of the measurement error. Despite these efforts, within-tester and between-tester reliabilities of goniometric measurements of cervical AROM generally were lower than previously reported passive ROM measurements at the elbow, knee, and shoulder. Such differences indicate that the reliability of measuring joint ROM is specific to the movement measured and to the regional anatomy and the biomechanics. For example, repeated measurements of AROM of the elbow generally would show less day-to-day variability than would repeated measurements of AROM of the cervical spine. The cervical spine consists of a complex series of multiaxial joints in which movements are controlled by numerous muscles that act across several joints simultaneously.

Physical therapists have been urged to avoid the VE technique for measurement of joint ROM, presumably because of the subjective nature of this measure. Recently, Rothstein defined a subjective measurement operationally as lacking a reasonable level of between-tester reliability. To our knowledge, our study is the first to document the between-tester reliability of VE for measuring AROM of the cervical spine in patients. Compared with goniometric techniques, the between-tester reliability of VE is poor overall, with moderate measurement error. We urge physical therapists to avoid using VE when two or more therapists take repeated AROM measurements of the cervical spine in the same patient. Clinicians may erroneously conclude that a patient's AROM has indeed changed because of treatment effects when the change could probably be attributed to inherent measurement error.

**Conclusion**

Based on our clinical study of 60 patients with orthopedic disorders in a physical therapy outpatient department, we conclude that AROM measurements on the cervical spine made by the same physical therapist have good to high reliability, regardless of whether the therapist used the CROM device or the UG. When different physical therapists measured the same subject's AROM, the CROM device was the most reliable testing instrument. Repeated measurements with the UG and VE had poor to fair between-tester reliability, even though the subject's body position was controlled and the testers used operationally defined measurement techniques. Based on low overall ICC values, we believe that goniometric and nongoniometric devices should not be interchanged when taking repeated measurements of cervical AROM in patients with neck pain.

**Acknowledgments**

We thank the following physical therapists for their assistance in this study: W. Jean Baron, Scott W Bandel, Ronald J Cervantes, Jay C Goetting, Margaret D Lane, Peter A Lommen, Timothy J Madson, Jenny L Miller, Daynee A Orte, Susan M Piper, and Lee D Shibley. We also thank Robert D Tryggestad, PT, for his cooperation and encouragement during this study.