The Reliability of Goniometric Measurements of Active and Passive Wrist Motions

Margaret M. Horger

Key Words: goniometry • measurement scales • wrist

A reliability study was conducted to determine (a) the intrarater and interrater reliability of goniometric measurement of active and passive wrist motions under clinical conditions and (b) the effect of a therapist's specialization on the reliability of measurement. Randomly paired therapists performed repeated measurements of active and passive wrist motions in 48 subjects who had been referred to one of four occupational therapy or hand management clinics for evaluation and treatment. The data were analyzed with an intraclass correlation coefficient. A posteriori data analyses were performed to determine the effects of identified sources of error on the reliability of measurement.

The results indicated that measurement of wrist motion by individual therapists is highly reliable and that intrarater reliability is higher than interrater reliability for all active and passive motions. Interrater reliability was generally higher among specialized therapists for reasons not immediately apparent from this study. With the exception of pain, identified sources of error were found to have surprisingly little effect on the reliability of measurement.

Margaret M. Horger, MS, OTR/L, is Assistant Professor, School of Occupational Therapy, Saint Ambrose University, 518 West Locust Street, Davenport, Iowa 52803.

This article was accepted for publication May 5, 1989.
motions were found. Therefore, a reliability study was conducted to answer the following questions:

1. What is the intrarater reliability of the goniometric range of motion measurement of active and passive wrist motions in a clinical setting?
2. What is the interrater reliability of the goniometric range of motion measurement of active and passive wrist motions in a clinical setting?
3. What is the effect of a therapist's specialization on the reliability of goniometric measurement of active and passive range of motion of the wrist?

The study design was a modification of the method originally described by Rothstein et al. (1983).

**Method**

*Definition of Terms*

The following terms were operationally defined:

- **Active range of motion**—The maximum amount of joint motion attained by a subject during the performance of voluntary joint motion.
- **Passive range of motion**—The maximum amount of joint motion attained by a subject during the application of external force in the absence of voluntary muscle contraction.
- **Specialized therapist**—An occupational therapist or physical therapist at least 50% of whose treatment time was spent in the treatment of upper extremity injuries.
- **Nonspecialized therapist**—An occupational therapist or physical therapist less than 50% of whose treatment time was spent in the treatment of upper extremity injuries.

*Subjects*

The subjects were drawn from two occupational therapy departments and two hand management clinics. Criteria for inclusion in the study were that the subjects had been referred to the clinic for evaluation, treatment, or both and that the measurement of wrist range of motion would have normally been included in their assessments. Subjects undergoing periodic re-evaluation were eligible for inclusion, but repeated measurement of the same extremity was not permitted. All subjects were at least 18 years old and signed a statement of informed consent.

Fifty wrist joints were measured in a sample of 48 subjects, 33 men and 15 women. The subjects ranged in age from 18 to 71 years, with a mean of 38.8 years. Forty-four subjects were of working age, 23 of whom had been injured in the course of their employment. Twenty-four subjects had injuries of the dominant hand, and 36 had undergone at least one surgical intervention. None of the subjects presented brain injury or upper motor neuron disorders.

*Raters*

Eleven registered occupational therapists and two licensed physical therapists participated in this study. The clinical experience of the 13 raters ranged from 2 months to 17 years, with a mean of 8.67 years. Six raters were specialists in hand therapy, with a mean of 3.75 years of specialty practice, and were employed at hand management clinics. The 7 nonspecialized raters worked in general physical disability settings. The 6 specialized raters had a mean of 8.08 years of clinical practice; the 7 nonspecialized raters had a mean of 9.41 years of clinical practice. In their daily work, the specialized raters measured wrist motions an average of five times per day; the nonspecialized raters measured wrist motions an average of three times per month.

*Instrumentation*

Active and passive range of motion measurements were performed with 8-in. blinded goniometers to which Vernier scales had been added (see Figure 1). A Vernier scale is a device used to increase the level of accuracy to which a linear or angular scale may be read; it eliminates the need for estimation of readings over a 5° interval. The accuracy of the goniometers and the Vernier scales was tested through the measurement of known angles. The goniometers used in this study were accurate to within 1°.

*Procedure*

The following procedure was used for the measurement of subjects by randomly paired sets of raters. For

Figure 1. Blinded goniometer with Vernier scale.

adduction. No techniques were specified, but the subject's elbow, forearm, and finger positions were noted. The measurement was read and recorded, and the goniometer was reset to 0° before it was returned to the rater. The sequence was repeated so that there were two measurements of each active motion. After the first rater completed these measurements, the second rater measured the subject's active wrist range of motion in the same sequence. Upon completion, the first rater returned and, using the same sequence, repeated the measurements for passive wrist range of motion, followed again by the second rater doing the same. To reduce the effect of muscle fatigue, a 2-min rest period was provided between the measurement of active motion by the first and second raters. After all of the data were collected for each subject, the raters were asked to note factors that might have affected the reliability of measurement for that subject.

The measurements required for this study preceded other evaluation procedures, and no treatment was provided until all of the data had been collected for the subject. None of the raters observed the measurement technique used by their paired raters during the data collection procedure. The raters were instructed to refrain from discussing the study until after its completion.

Data Analysis
A data analysis was performed with an intraclass correlation coefficient (ICC) (1, 1) (Shrout & Fleiss, 1979), which is based on the analysis of variance and provides an estimate of the agreement between ratings. The between-raters variance is included in the calculation of the ICC (1, 1), thus permitting generalization of the results beyond the raters in the present study. The standard error of measurement, which reflects the amount of error associated with individual measurements, was calculated for measurement of all active and passive wrist motions.

Results
Reliability
Descriptive statistics for each variable are displayed in Table 1. Intrarater reliability for the measurement of active wrist motions was determined through a comparison of the first and second measurements of each motion by each rater (see Table 2). Reliability coefficients for the measurement of active wrist range of motion exceeded .900 for all wrist motions. Intrarater reliability coefficients for passive wrist range of motion were similarly high. Agreement between measurements of motions in the sagittal plane (flexion and extension) was in both cases higher than that for measurements of motions in the frontal plane (abduction and adduction).

Interrater reliability was determined by a comparison of the first measurement of the first and second raters for each subject (see Table 2). Reliability coefficients for interrater reliability of measurement of active wrist motions ranged from .783 for adduction to .905 for wrist flexion. Interrater reliability for measurement of passive wrist motions was generally lower, ranging from .662 for abduction to .855 for flexion. With the exception of adduction, the ICCs were lower for passive than for active measurement of all motions.

Specialty Practice
Intrarater reliability coefficients for the measurement of active wrist motions were higher among specialized than among nonspecialized raters (see Table 3). For active range of motion measurements, the coefficients for the specialized raters ranged from .909 for abduction to .974 for extension; those for the nonspecialized raters ranged from .860 for abduction to .890 for extension. Intrarater reliability coefficients for measurement of passive range of motion were slightly higher for the specialized than for the nonspecialized raters for flexion and extension. The reverse was true,
Table 2
Reliability for Measurement of Wrist Range of Motion

<table>
<thead>
<tr>
<th>Wrist Motion</th>
<th>Active</th>
<th></th>
<th></th>
<th>Passive</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC</td>
<td>960</td>
<td>942</td>
<td>3.652</td>
<td>958</td>
<td>939</td>
<td>3.537</td>
</tr>
<tr>
<td>95% CI</td>
<td>958</td>
<td>940</td>
<td>4.528</td>
<td>958</td>
<td>940</td>
<td>4.436</td>
</tr>
<tr>
<td>SEM</td>
<td>4.528</td>
<td>4.557</td>
<td>908</td>
<td>868</td>
<td>2.746</td>
<td></td>
</tr>
<tr>
<td></td>
<td>935</td>
<td>830</td>
<td>3.485</td>
<td></td>
<td>2.007</td>
<td></td>
</tr>
<tr>
<td>ICC</td>
<td>853</td>
<td>774</td>
<td>7.000</td>
<td>853</td>
<td>747</td>
<td>7.012</td>
</tr>
<tr>
<td>95% CI</td>
<td>853</td>
<td>851</td>
<td>6.579</td>
<td>853</td>
<td>777</td>
<td>5.242</td>
</tr>
<tr>
<td>SEM</td>
<td>853</td>
<td>799</td>
<td>3.009</td>
<td>853</td>
<td>5.077</td>
<td>5.263</td>
</tr>
<tr>
<td></td>
<td>783</td>
<td>673</td>
<td>5.777</td>
<td></td>
<td>7.42</td>
<td>4.849</td>
</tr>
</tbody>
</table>

Note. ICC = intraclass correlation coefficient; CI = confidence interval; SEM = standard error of measurement.

Table 3
Intra-rater Reliability for Measurement of Wrist Range of Motion as a Function of Occupational Therapists' Specialty Practice

<table>
<thead>
<tr>
<th>Wrist Motion</th>
<th>Active</th>
<th></th>
<th></th>
<th>Passive</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC</td>
<td>974</td>
<td>958</td>
<td>909</td>
<td>932</td>
<td>860</td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>958</td>
<td>955</td>
<td>900</td>
<td>933</td>
<td>933</td>
<td></td>
</tr>
<tr>
<td>SEM</td>
<td>946</td>
<td>935</td>
<td>855</td>
<td>903</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Inter-rater Reliability for Measurement of Wrist Range of Motion as a Function of Occupational Therapists' Specialty Practice

<table>
<thead>
<tr>
<th>Wrist Motion</th>
<th>Active</th>
<th></th>
<th></th>
<th>Passive</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC</td>
<td>.919</td>
<td>.938</td>
<td>.837</td>
<td>.826</td>
<td>.527</td>
<td>.678</td>
</tr>
<tr>
<td>95% CI</td>
<td>.872</td>
<td>.896</td>
<td>.734</td>
<td>.717</td>
<td>.009</td>
<td>.094</td>
</tr>
<tr>
<td>SEM</td>
<td>.901</td>
<td>.902</td>
<td>.637</td>
<td>.808</td>
<td>.545</td>
<td>.594</td>
</tr>
</tbody>
</table>

Note. ICC = intraclass correlation coefficient; CI = confidence interval.
Reliability

Intrarater reliability for the measurement of both active and passive wrist motion was determined to be excellent for all motions. Reliability coefficients for active and passive measurements were similar for each motion, with the exception of adduction. The lower bounds of the ICCs were also in the excellent range. Thus, repeated measurements of wrist motion by the same therapist can be expected to be highly reliable under clinical conditions.

Intrarater reliability coefficients were lower than those for intrarater reliability, although coefficients for all motions except passive abduction remained in the excellent range, as defined by Fleiss (1986). The lower bounds of the intrarater reliability coefficients, however, indicate that repeated measurements made by different therapists should be interpreted more conservatively than should measurements made by the same therapist. Thus, sequential measurements made by different therapists must show a greater disparity before the presence of clinical change in a patient may be assumed.

Specialty Practice

Although specialty practice was found to affect both intrarater and interrater reliability, the effect was greater in the case of interrater reliability. Intrarater reliability among specialized therapists was excellent for all motions except passive abduction. Intrarater reliability among nonspecialized therapists, however, was only fair for active extension, flexion, and adduction and for passive extension and flexion. Furthermore, the lower bounds of the ICCs for all motions except active abduction and passive adduction were close to 0, indicating that sequential measurements made by different therapists should be compared cautiously.

The source of lower interrater reliability among the nonspecialized therapists was difficult to determine with the data from this study. Three sources of error may have contributed to the discrepancies in reliability found between specialized and nonspecialized therapists. First, a practice effect may have resulted from the more frequent measurement of wrist motion by specialized than by nonspecialized therapists. Second, differences in goniometer alignment, particularly during measurement of flexion and extension, may have affected the reliability of measurement. The literature supports alignment of the goniometer along either the radial (Moore, 1984; Scott & Trombly, 1983) or the ulnar (Esch & Lepley, 1973; Moore, 1984; Norkin & White, 1985) aspect of the wrist. All nonspecialized raters measured on the radial aspect of the wrist; all specialized raters measured on the ulnar aspect of the wrist. Thus, the present study could not include a separate analysis of the factors of specialization and goniometer alignment.

Finally, the specialized therapists may have been more knowledgeable of and thus more consistent in the location of anatomical landmarks used in alignment of the goniometer. Fish and Wingate (1985) found that reliability was enhanced by consistent use and location of anatomical landmarks during measurement.

Sources of Error

Potential sources of error in goniometric measurement cited in the literature included location of the axis of motion (Hellebrandt et al., 1949), inconsistent patient positioning and goniometric alignment (Fish & Wingate, 1985; Moore, 1949), inability (Cobe, 1928) or disinclination (Salter, 1955) of the patient to accurately reproduce active joint motion, and inconsistent identification (Fish & Wingate, 1985) or alteration (Hamilton & Lachenbruch, 1969) of bony landmarks used to guide measurement. The application of external force is an additional variable during the measurement of passive range of motion (Fish & Wingate, 1985).

Pain; variability in the application of external force during measurement of passive range of motion; and disarrangement of anatomical landmarks due to trauma, scarring, or edema were most often cited as factors influencing the reliability of measurement in this study. A discussion of their effects on the reliability of measurement of wrist range of motion follows.

Pain. Pain, cited as a factor for 28 subjects (41 measurements), was found to systematically reduce intrarater reliability of both active and passive range of motion measurements, although overall intrarater reliability remained excellent for all motions. The presence of pain was associated with more marked reduction in the intrarater reliability of measurement of active, as opposed to passive, wrist range of motion. Intrarater reliability of passive range of motion measurements did not appear to be systematically affected by the presence of wrist pain. This effect may have been the result of the order in which measurements were performed. The exacerbation of pain symptoms during measurement of active motion by the first rater may have affected measurements obtained by the second rater. Discrepancy might be expected between measurement of the asymptomatic joint by the first rater and the symptomatic joint by the second rater. Passive range of motion measurements followed active range of motion measurements in all cases. Pain levels during passive range of motion measurement of the first and second raters might then be expected to be more constant than during active motion. The presence or anticipation of pain by the subject, the rater, or both may have provided a more consistent end
Application of external force. The rater typically applied external force during measurement of passive range of motion. In 15 subjects (21 measurements), however, the subjects themselves applied external force, according to the rater's instructions. Intrarater agreement for passive wrist extension, flexion, and adduction was excellent regardless of the method of force application, as indicated by ICCs exceeding .920. Passive abduction was found to be measured less reliably when external force was applied by the subject. The ICCs for measurement of passive abduction, however, were generally lower than for those of other motions. Thus, with the exception of abduction, external force required for the measurement of passive wrist range of motion may be applied to the hand by the therapist or by the patient, according to the therapist's preference.

Anatomical changes. Anatomical landmarks used for goniometer alignment are easily obscured after injury to the hand or wrist. Hand edema is typically manifested dorsally, obscuring the contours of the metacarpals. Arthritic changes or posttraumatic wrist injuries may alter the orientation of the carpal bones, thus compromising the therapist's ability to locate the axis of motion. Finally, scarring may alter contours and obscure landmarks.

Although these features concerned the raters of 12 subjects in the present study (16 measurements), the presence of anatomical changes did not adversely affect intrarater reliability. In fact, intrarater reliability was higher for the subjects with observable anatomical changes. This associated improvement in reliability may be due to the raters' more careful goniometric alignment for what was presumed to be a difficult measurement situation.

Positioning. Riddle et al. (1987) found that among physical therapists, intrarater reliability of measurement of passive shoulder motion was improved when the subjects were positioned similarly. A similar analysis in the present study was impeded by the fact that little variation in the positioning of subjects was noted.

Clinical Implications. The results of this study suggest that despite a lack of standardization of technique, the goniometric measurement of wrist motions is highly reproducible when performed under clinical conditions. Because intrarater reliability was consistently lower than intrarater reliability, the same therapist should reevaluate patients whenever possible. These results are consistent with those of similar studies of measurement of the shoulder, elbow, and knee joints (Riddle et al., 1987; Rothstein et al., 1983).

The standard error of measurement provides an estimate of the amount of error associated with individual measurements; this is useful in determining the presence of meaningful clinical change in patients in response to clinical intervention. The standard error of measurement had not been presented in any studies on the reliability of goniometry that were reviewed.

The reliability of measurement of wrist motion during the present study was remarkably unaffected by such factors as variations in application of external force and anatomical changes due to trauma or deformity. Perhaps this was due to the raters' adequate control of these variables, which introduces implications for entry-level education. The control of such variables in changing clinical situations requires an understanding of the rationale for measurement techniques as well as competence in the psychomotor aspects of assessment. This level of understanding should be imparted in academic and fieldwork education programs.

The results of this study cannot be applied to measurements of wrist motion used for permanent disability ratings. The American Medical Association (AMA) stated that measurement techniques for the evaluation of permanent impairment should be "simple, practical, and scientifically sound" (AMA, 1984, p. 1). The AMA (1984) has provided detailed instructions for the measurement of each joint. None of the raters in this study used the AMA's suggested measurement techniques. Occupational therapists performing final disability ratings should therefore continue to use the techniques recommended by the AMA.

Only two reviewed studies addressed the amount of wrist motion that normal subjects required to perform functional activities (Brumfield & Champoux, 1984; Palmer, Werner, Murphy, & Glisson, 1985). The combined results of these two studies indicated that an arc of approximately 35° of motion in the sagittal plane and 25° in the frontal plane was sufficient for most activities of daily living. No studies have addressed the wrist range of motion requirements for occupational activities in an industrial environment, although the role of wrist positioning in the development of nerve compression and cumulative trauma disorders has been hypothesized (Armstrong, Fine, Goldstein, Lifshitz, & Silverstein, 1987; Gelberman, Szabo, & Mortenson, 1984; Silverstein, Fine, & Stetson, 1987).

An effort was made to sample a broad range of clinical experiences and situations to make the results of this study more useful to occupational therapists in clinical practice. Nevertheless, the subjects in this study represent a limited sample, so the study results are not necessarily generalizable beyond the clinics.
and situations represented. In addition, the use of a date recorder eliminated the potential error of a therapist's inability to accurately read and record measurements from a goniometer scale.

Summary

The goniometric measurement of active and passive wrist motions was found to be highly reliable when conducted under clinical conditions. Intrarater reliability was found to be consistently higher than interrater reliability, although overall reliability remained excellent. Specialty practice was found to affect interrater more than intrarater reliability. Interrater reliability may have been influenced by frequency of measurement, the knowledge base associated with specialization of practice, or the method of wrist motions was found to affect intrarater reliability. Subject-related variables such as anatomical changes due to trauma or deformity and variations in the application of external force were found to have surprisingly little effect on reliability. Pain was associated with a greater reduction in the reliability of measurement of active rather than passive range of motion; this effect may have been an artifact of the data collection procedure. Further research is recommended to determine the effects of specific patient- and therapist-related variables on the reliability of measurements and the functional requirements for wrist motion during work activities.

Acknowledgments

I wish to thank the staff at the following facilities for their assistance in data collection: Hand Management Specialists, Medical College of Virginia Hand Management Center, and Hunter Holmes McGuire Veterans Administration Medical Center, Richmond, Virginia; and Riverside Regional Medical Center, Newport News, Virginia. I also wish to thank Patti Maurer, PhD, OTT, Craig E. Nelson, MS, OTT, and David Bauer, PhD, for their editorial and statistical assistance.

This article was based on a thesis submitted as partial fulfillment of the requirements for the master of science degree from the Department of Occupational Therapy, Virginia Commonwealth University, Richmond, Virginia.

References


