Accuracy and reliability testing of two methods to measure internal rotation of the glenohumeral joint

Justin M. Hall, MD, Frederick M. Azar, MD, Robert H. Miller III, MD, Richard Smith, PhD, Thomas W. Throckmorton, MD*

Department of Orthopaedic Surgery and Biomedical Engineering, University of Tennessee–Campbell Clinic, Memphis, TN, USA

Background: We compared accuracy and reliability of a traditional method of measurement (most cephalad vertebral spinous process that can be reached by a patient with the extended thumb) to estimates made with the shoulder in abduction to determine if there were differences between the two methods.

Methods: Six physicians with fellowship training in sports medicine or shoulder surgery estimated measurements in 48 healthy volunteers. Three were randomly chosen to make estimates of both internal rotation measurements for each volunteer. An independent observer made objective measurements on lateral scoliosis films (spinous process method) or with a goniometer (abduction method). Examiners were blinded to objective measurements as well as to previous estimates.

Results: Intraclass coefficients for interobserver reliability for the traditional method averaged 0.75, indicating good agreement among observers. The difference in vertebral level estimated by the examiner and the actual radiographic level averaged 1.8 levels. The intraclass coefficient for interobserver reliability for the abduction method averaged 0.81 for all examiners, indicating near-perfect agreement. Confidence intervals indicated that estimates were an average of 8° different from the objective goniometer measurements. Pearson correlation coefficients of intraobserver reliability for the abduction method averaged 0.94, indicating near-perfect agreement within observers. Confidence intervals demonstrated repeated estimates between 5° and 10° of the original.

Conclusions: Internal rotation estimates made with the shoulder abducted demonstrated interobserver reliability superior to that of spinous process estimates, and reproducibility was high. On the basis of this finding, we now take glenohumeral internal rotation measurements with the shoulder in abduction and use a goniometer to maximize accuracy and objectivity.

Level of evidence: Level III, Diagnostic Study.

Keywords: Glenohumeral joint; internal rotation; measurement; spinous process method; abduction method; accuracy

Internal rotation of the glenohumeral joint is an important part of the clinical examination of the shoulder, reflecting an ability to perform several activities of daily living including self-care. Several shoulder outcomes measures use glenohumeral internal rotation as part of their...
each of the 3 examiners then recorded his estimate of the vertebral midline location on the back by the independent observer (Fig. 1). was internally rotated and the thumb was placed at a random point between 0° and 90°. This method has been shown to have poor interobserver reliability and only good intraobserver reliability. In addition, this traditional method can easily be skewed by abnormal anatomy and limited or increased motion of the elbow and scapulothoracic articulation.

Measurement of internal rotation of the glenohumeral joint with goniometers or inclinometers with the shoulder in 90° of abduction and the elbow flexed to 90° has been shown to have good interobserver and intraobserver reliability. Scapulothoracic motion has been shown to contribute up to approximately 12% to the measured internal rotation of the shoulder. Boon and Smith determined that a more accurate assessment of true glenohumeral internal rotation can be made when the scapula is manually stabilized by the examiner and the shoulder is in 90° of abduction.

Although spinous process estimation has been proved to be unreliable, it is still the most commonly used method to report internal rotation of the glenohumeral joint. No studies to date have compared this method with measurements taken with the shoulder in abduction. The purpose of this study was to compare the accuracy and reliability of the traditional spinous process estimation method to estimates of measurements taken in abduction. We hypothesized that there would be no significant differences in reliability or accuracy between the two methods.

**Materials and methods**

This prospective comparative study recruited 61 healthy volunteers between 18 and 65 years of age. Excluded were women of childbearing age, patients with previous shoulder surgery, and those with current shoulder pain. Women of childbearing age were excluded because of the radiation from the lateral scoliosis view taken of each patient as a part of the study. Patients with elbow pain or prior elbow surgery also were excluded because elbow range of motion has been shown to affect spinous process estimates. Thirteen patients were excluded because of poor-quality radiographs, leaving a total of 48 subjects, 38 men and 10 women.

Six physicians with fellowship-level training in sports medicine or shoulder and elbow surgery composed a pool of observers to estimate measurements. All examiners were familiar with both internal rotation measurement techniques. For each patient in the study, 3 of the 6 physicians were randomly chosen to make estimates of both internal rotation measurements.

The side examined was chosen at random by an independent observer (27 right and 21 left). The traditional spinous process estimation method was used first. After a small radiopaque metal sphere was taped to the tip of the thumb, the volunteer’s shoulder was internally rotated and the thumb was placed at a random midline location on the back by the independent observer (Fig. 1). Each of the 3 examiners then recorded his estimate of the vertebral midline that would maximize the angle between horizontal and the axis of the volunteer’s forearm and recorded this measurement (Fig. 3). With the goniometer removed, the 3 examiners then recorded their estimates of internal rotation. This process was repeated for each patient after an interval of at least 10 minutes, after which the independent observer placed the volunteer’s shoulder back in its initial internal rotation using the goniometer as a guide, and the 3 examiners again estimated the internal rotation of the glenohumeral joint. All data were collected by the independent observer. The examiners were blinded to goniometric measurements as well as to their previous estimates.

Statistical analyses for interobserver agreement were conducted for both methods by the two-tailed Pearson correlation coefficient for intraclass correlation with 95% confidence intervals. Statistical analysis for intraobserver reliability also was performed for the abduction/internal rotation method by the Spearman correlation coefficient to find intraclass correlation with 95% confidence intervals. Intraclass coefficients of 0.6 to 0.8 were...
Results

Review of the radiographs demonstrated that vertebral levels from T6 to S4 were used in the study group for the traditional spinous process estimation method. Intraclass coefficients for interobserver reliability for the traditional method averaged 0.75 (range, 0.62-0.94), indicating good agreement among observers (Table I). Confidence intervals placed estimates at an average of 1 or 2 levels different from the objective radiographic findings. The difference in vertebral level estimated by the examiner and the actual radiographic level averaged 1.8 levels (range, 1.0-2.4). There were no statistically significant differences in estimate accuracy between observers for this method (P > .05).

Internal rotation estimates by the abduction/internal rotation method ranged from 10° to 68°, with an average of 38°. The intraclass coefficient for interobserver reliability for this method averaged 0.81 (range, 0.7-0.91) for all examiners, indicating near-perfect agreement (Table I). Confidence intervals indicated that estimates were 6° to 11° (average, 8°) different from the objective goniometer measurements. The intraclass coefficient for intraobserver reliability for this method averaged 0.94 for all observers (range, 0.89-0.96) within individual examiners, again indicating near-perfect agreement. Confidence intervals of repeated measurements indicated 5° to 10° differences among observers on repeated examination. There were no statistically significant differences in estimate accuracy between or within observers for this method (P > .05).

Discussion

Internal rotation of the glenohumeral joint is commonly reported in outcomes studies and in the clinical setting and is important for shoulder function associated with self-care, such as toileting and dressing. Outcomes of shoulder injuries and treatment often are reported with shoulder outcomes measures, many of which, including the Rowe, Constant, and American Shoulder and Elbow Surgeons scores, use internal rotation as a component of their scoring systems. Internal rotation is most commonly reported as the spinous process estimate as described in this study;
measurement of glenohumeral internal rotation

No statistically significant differences were demonstrated among examiners. Interobserver and intraobserver reliability for the abduction/internal rotation method were superior to spinous process estimates.

<table>
<thead>
<tr>
<th>Examiner</th>
<th>Interobserver ICC for spinous process estimates</th>
<th>Interobserver ICC for abduction/internal rotation estimates</th>
<th>Intraobserver ICC for abduction/internal rotation estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.82</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>2</td>
<td>0.86</td>
<td>0.78</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>0.57</td>
<td>0.83</td>
<td>0.95</td>
</tr>
<tr>
<td>4</td>
<td>0.62</td>
<td>0.89</td>
<td>0.95</td>
</tr>
<tr>
<td>5</td>
<td>0.94</td>
<td>0.74</td>
<td>0.90</td>
</tr>
<tr>
<td>6</td>
<td>0.71</td>
<td>0.70</td>
<td>0.94</td>
</tr>
<tr>
<td>Average</td>
<td>0.75</td>
<td>0.81</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Several studies have shown that spinous process estimates are an imprecise method of estimating glenohumeral internal rotation. Edwards et al. showed that measurement of internal rotation by vertebral-level estimation was not readily reproducible between observers, and Mallon et al. showed in a radiographic analysis that elbow motion contributes greatly to the vertebral level reached by the patient. Therefore, not only do spinous process estimates fail to achieve strong interobserver reliability, they also are affected by motion or compensation in other upper extremity joints. Ginn et al. studied spinous process estimates of active internal rotation with the shoulder in 90° of abduction and found poor correlation between active internal rotation and vertebral-level measurements for patients with shoulder pain; however, we did not include patients with shoulder pain in this study.

Aims of our study were to compare spinous process estimates with those obtained with the abduction/internal rotation method to see if any differences existed in interobserver reliability and to further investigate the interobserver reliability of the abduction/internal rotation method. Awan et al. found intraobserver reliability of the abduction/internal rotation method to be good according to the intraclass correlation coefficient (ICC). Although the premise of their study was similar to ours, they differed in that they examined the patient supine and used digital inclinometers. We used standard goniometers, which may be more practical in the clinical setting, and examined the patient upright with manual scapular stabilization. Whereas Awan et al. found good intraobserver reliability for the abduction/internal rotation method, our study showed an even better ICC of 0.94, indicating near-perfect agreement. We also found a higher interobserver reliability, with an average ICC of 0.81, demonstrating near-perfect agreement, than the ICCs of 0.41 to 0.61 reported by Awan et al. Our study correlated well with that of Edwards et al., who also found suboptimal interobserver reliability of spinous process estimates with ICC less than 0.4.

Our data indicate that estimates with use of the abduction/internal rotation method are more accurate than those with the spinous process method. Spinous process estimates did show good intraclass correlation for interobserver reliability, but at 0.75 it was less than desired for outcomes reporting. Also, 95% confidence intervals showed that examiner estimates were 1 or 2 vertebral levels off from the radiographically confirmed level. The results of the abduction/internal rotation method were clearly better by comparison, in which a near-perfect interobserver ICC of 0.81 was found; however, 95% confidence intervals showed examiner estimates to be 6° to 11° off from the objective goniometer measurements. Intraobserver reliability was even better for this method, with an ICC of 0.94 and 95% confidence intervals demonstrating an average difference of 5° to 10° between examiner estimates and goniometer measurements. The examiners in this study used visual inspection only for the abduction/internal rotation method, and we suspect that had they used goniometers to make their assessments, intraobserver and interobserver reliability would be even higher.

There were several limitations to our study. First, not every patient was examined by every examiner. Although our data could have been further solidified if every examiner had examined every patient, logistical concerns made it more practical to conduct the study as described. Having 6 examiners, however, was a strength because we were able to collect interobserver data using multiple data points. Also, all 6 examiners had near-perfect intraobserver reliability for the abduction/internal rotation method, showing it to be reproducible. Another weakness was that examiners were given only 10 minutes between examinations using the abduction/internal rotation method. Although examiners were blinded to their original estimates, this may have been a short enough time that the examiners could recall their initial estimation and simply record the same value, leading to artificial inflation of ICC for intraobserver reliability for this method. We believed that with the large number of patients being examined, 10 minutes was a long enough time to allow a true reassessment of the internal rotation angle. Whereas there is no literature showing a
minimum amount of time needed between repeated examinations, Edwards et al, in a similar study with fewer patients, used a time of 20 minutes between repeated examination. Another potential limitation is that we used asymptomatic volunteers in this study, all of whom could attain 90° of passive abduction. We acknowledge that in the clinical setting, some patients may not be able to obtain this level of passive abduction because of either pain or stiffness from their shoulder disorder. Finally, all examiners in our study had advanced training beyond residency in sports medicine or shoulder surgery. Therefore, these examiners may be more familiar with physical examination of the shoulder than the general orthopaedic physician, which could make it difficult to extrapolate these data to non-fellowship-trained physicians.

### Conclusion

The results of this study confirmed that spino...