Intrarater and interrater reliability of a manual technique to assess anterior humeral head translation of the glenohumeral joint

Todd S. Ellenbecker, MS, PT, SCS, OCS, CSCS,[a] David S. Bailie, MD,[b] Angelo J. Mattalino, MD,[c] David G. Carfagno, DO,[c] Michael W. Wolff, MD,[d] Scott W. Brown, PhD,[e] and Jonna M. Kulikowich, PhD,[e] *Scottsdale and Phoenix, Ariz, and Storrs, Connecticut*

The purpose of this study was to determine the intrarater and interrater reliability of a manual anterior humeral head translation test. Fifteen subjects were positioned lying in a supine position with their identity shielded from examiners. A standard manual anterior humeral head translation test was performed and repeated with the glenohumeral joint in 90° of elevation in the scapular plane, with use of the grading method proposed by Altchek and Dines in 1993. Reliability was assessed with the coefficient of agreement and κ statistic. Intrarater reliability was 81.4% comparing grade I and II translation. This decreased to 54% when examiners distinguished between grades I, I+, II, and II+. Interrater reliability for the same comparisons was 70.4%, decreasing to 37.3%. On the basis of these data, the technique of manually assessing anterior humeral head translation studied has poor overall interrater reliability and only fair intrarater reliability. The test-retest accuracy of humeral head translation is enhanced when examiners only determine the relationship of the humeral head relative to the glenoid rim. (J Shoulder Elbow Surg 2002;11:470–5.)

Clinically identifying anterior humeral head translation in the shoulder is an integral part of the comprehensive evaluation of shoulder pathology.[1,4,5,10-12] Examination techniques for detecting anterior humeral head translation of the shoulder, particularly in throwing or overhead athletes, are typically performed with 90° of elevation in the scapular plane with an anteriorly directed force.[4,8] A complete examination of the glenohumeral joint for anterior stability includes assessment of humeral head translation in 0°, 45°, and 90° of glenohumeral joint abduction to test portions of the anterior capsule selectively.[13] The position of 90° of glenohumeral joint abduction closely approximates the abduction position used during functional overhead activities such as tennis and throwing a baseball or football.[4]

The presence of increased anterior translation of the humeral head, as a result of anterior capsular laxity in the dominant shoulder of the throwing arm, has been reported previously.[8,11,16] Increased humeral head translation has been detected by manual testing in patients diagnosed with glenohumeral joint instability.[16,17] Conversely, a decrease in humeral head translation has been reported in patients diagnosed with glenohumeral joint impingement.[17] It is, therefore, important to establish a standard clinical test to detect glenohumeral translation reliably. The presence of increased humeral head translation without symptoms, termed laxity, does not necessarily equate with instability.[4]

Errors inherent with manual examination include the subjective nature of the interpretation of the test result, failure to standardize the extremity position, and variation in the forces imparted to produce movement, resulting in a lack of interrater reliability.[3,10] Despite their widespread use, there is limited objective research reporting the intrarater or interrater reliability of the clinical tests used to detect anterior glenohumeral joint translation.

The purpose of this study was to assess the intrarater and interrater reliability of a clinically accepted manual technique to assess anterior humeral head translation of the humeral head relative to the glenoid.

MATERIAL AND METHODS

Fifteen human subjects (6 male and 9 female), without a history of significant glenohumeral joint injury in the past

Figure 1 Experimental setup demonstrating protection of the subjects' identities from the examiners.

year and no history of shoulder surgery, were included in this study. A significant glenohumeral joint injury was defined as pain or limitation of function that altered either daily activity or recreational activity for 3 consecutive days. Several individuals with increased glenohumeral joint laxity were selected to ensure variability among the subject population with respect to humeral head translation. Subjects selected had a variety of levels of glenohumeral mobility and were considered representative of the population of individuals examined clinically in orthopaedic and sports medicine. An informed consent form was completed by all subjects before participation in this study. The research protocol was approved by an institutional review board (Stryker Physiotherapy Associates, Memphis, Tenn).

Humeral head translation testing was performed by 4 physicians and 3 physical therapists. The physicians consisted of 2 sports medicine fellowship–trained orthopaedic surgeons, 1 primary care physician who completed a sports medicine fellowship, and 1 physical medicine and rehabilitation specialist. The physical therapists all worked in orthopaedic and sports physical therapy. All examiners were experienced in the use of manual anterior humeral head translation examination techniques and used them in their daily practice.

Subjects' identities were concealed from the examiners throughout the study. All subjects wore surgical scrub tops and bottoms and were covered with a standard hospital sheet while lying in a supine position on a plinth. Identical paper grocery bags were worn over the head of each subject with a towel serving as a swathe to conceal the neck (Figure 1).

Using the same testing sequence, each examiner performed a clinical anterior humeral head translation test, using an anterior drawer or anterior glide.[5] This technique placed the humeral head in 90° of glenohumeral joint elevation in the scapular plane (20°-30° anterior to the coronal plane), with neutral humeral rotation (Figure 2). An anterior force was then exerted directly along the line of the glenohumeral joint.

Anterior humeral head translation was graded by the system proposed by Altchek and Dines[1] (Table I). This system classifies translation as follows: grade I, translation of the humeral head contained within the glenoid; grade II, translation of the humeral head onto the glenoid rim with immediate relocation upon removal of the anterior stress; and grade III, translation of the humeral head over the glenoid rim without relocation with removal of stress. In addition, a plus sign to designate a soft, compliant end-feel was used to allow the clinician to describe humeral head translation further.[2,3] The end-feel has been defined as the feeling transmitted to the examiner's hands at the extreme range of passive motion.[3] There are 3 normal end-feels: bone-to-bone, soft-tissue approximation, and tissue stretch. A tissue stretch end-feel is expected at end range during the manual anterior humeral head translation test. This hard or springy type of end-feel has a slight give normally, as compared with the compliant or softer end-feel characteristic of a hypermobile glenohumeral joint. The examiner recorded the appropriate grade of humeral head translation for the corresponding extremity, with an identical method of testing applied to the contralateral extremity.

After initial testing by all examiners, the locations of the subjects on the plinths were scrambled by an assistant. Testing was repeated to assess each examiner's intrarater reliability with the manual test. To determine the interrater reliability, each examiner's degree of agreement was compared with that of all other examiners for the 15 subjects.

Data analysis

Intrarater reliability for each examiner was expressed as the percent agreement and was calculated as the number of identical matches divided by the total number of possible matches. This coefficient of agreement is the simplest index of agreement and measures how often raters agree on scores given to individual subjects. Interrater reliability was calculated by a similar method with the number of identical matches among the 7 examiners serving as the numerator.

Figure 2 Manual anterior humeral head translation used in this investigation. Placement of the glenohumeral joint in the scapular plane with 90° of elevation is noted.

Table I Grading of manual tests for anterior humeral head translation

Grade	Glenohumeral joint translation
I	Humeral head rides up the glenoid slope but not over the rim
II	Humeral head rides up and over the glenoid rim but spontaneously reduces when stress is removed
III	Humeral head rides up and over the glenoid rim and remains dislocated on removal of stress
+	Designates the presence of a soft, compliant endfeel at the end range of humeral head excursion in any of the 3 designated grades of movement

Table II Intrarater reliability of the 7 examiners for manual anterior humeral head translation

	% Agreement	κ	SE	95% CI
Grades I, I+, II, and II+*				
Examiner				
1	60	0.284	0.154	0.284 ± 0.302
2	33	0.085	0.117	0.085 ± 0.229
3	66	0.552	0.116	0.552 ± 0.227
4	63	0.465	0.142	0.465 ± 0.278
5	60	‡	—	—
6	46	0.273	0.120	0.273 ± 0.235
7	50	0.091	0.144	0.091 ± 0.282
Total	54	0.342	0.050	0.342 ± 0.098
Grades I and II†				
Examiner				
1	90	0.516	0.243	0.516 ± 0.476
2	83	0.645	0.136	0.645 ± 0.266
3	90	0.798	0.110	0.798 ± 0.216
4	73	0.459	0.162	0.459 ± 0.317
5	80	–0.059	0.051	–0.059 ± 0.100
6	66	0.286	0.178	0.286 ± 0.348
7	86	–0.053	0.041	–0.053 ± 0.080
Total	81.4	0.529	0.066	0.529 ± 0.129

Examiners 1 and 2 are orthopaedic surgeons, 3, 4, and 5 are physical therapists, and 6 and 7 are non-orthopaedic physicians.
*Data analysis with examiners distinguishing among all 4 grades.
†Data analysis with examiners distinguishing between only 2 grades.
‡Insufficient data for calculation of κ for examiner 5 for grades I, I+, II, and II+.

The initial data analysis was performed with 4 possible translation grades: I, I+, II, and II+. Statistical analysis was also performed with a κ coefficient. The κ coefficient is a chance-corrected measure of agreement and, in addition to looking at the proportion of observed agreements, also considers the proportion of agreements that would be expected simply by chance. SE and 95% CIs were also calculated and expressed for the κ coefficients.

Additional analysis of the test-retest data was repeated examining the ability of the examiners to interpret only grade I or II without the addition of the plus sign for end-feel.

RESULTS

Intrarater reliability

The coefficient of agreement, κ coefficient, SE, and 95% CIs from all 7 examiners are displayed in Table II for both conditions (for determination of grades I, I+, II, and II+ and determination solely between grades I and II). The overall percent agreement when interpreting humeral head translation between grades I, I+, II, and II+, was 54%, with a κ statistic of 0.342. The overall coefficient of agreement when interpreting between grades I and II was 81.4%, with a κ statistic of 0.529.

Table III Interrater reliability of manual anterior humeral head translation among the 7 examiners for grades with end-feel designation

Grade	Grade				
	I	I+	II	II+	Total*
I	**111**	113	40	12	276
I+	69	**88**	32	12	201
II	24	29	**20**	11	84
II+	6	31	16	**16**	69

Bold numbers indicate the number or correct matches in translation. The overall coefficient of agreement between all examiners was 37.3% (κ = 0.091).
*The total matches possible is 630.

Table IV Interrater reliability of manual anterior head translation among the 7 examiners for grades without end-feel designation

Grade	Grade		
	I	II	Total*
I	**381**	96	477
II	90	**63**	153

Bold numbers indicate the number or correct matches in translation. The overall coefficient of agreement between all examiners is 70.4% (κ = 0.208).
*The total matches possible is 630.

Interrater reliability

Table III contains interrater reliability data for all grades of translation. Evaluation with all 4 possible grades (I, I+, II, and II+) resulted in a coefficient of agreement of 37.3% (κ, 0.091). When end-feel was not considered, the coefficient of agreement increased to 70.4% (κ, 0.208) (Table IV).

DISCUSSION

The results of this study indicate that both intrarater reliability and interrater reliability of a clinically relevant manual examination technique assessing humeral head translation were greatly improved when only the relationship of the humeral head to the glenoid rim was measured. The addition of an end-feel designation to the traditional grading system resulted in a decrease in reliability both within and among examiners.

Quantification of manual humeral head translation has been studied previously. In vivo assessment of glenohumeral joint translation was performed by Harryman et al.[6] In the normal shoulder, a mean of 7.8 ± 4.0 mm of anterior translation of the humeral head was found, with 7.9 ± 5.6 mm of posterior translation. The relative motion of the humeral head and scapula was determined with an electromagnetic spatial tracker. This device was pinned percutaneously to the humerus and scapula of 8 normal subjects, with manual humeral head translation tests performed by 1 investigator. Despite the detailed objective data accumulated in their research, this method of humeral head translation measurement is not clinically applicable.

Levy et al[10] performed a test-retest examination of manual humeral head translation using anterior/posterior drawer and sulcus tests on asymptomatic division I collegiate athletes over a 3-month period. They measured anterior, posterior, and inferior translation. Glenohumeral laxity was graded from 0 to 3+ by 4 physicians, with an overall intraobserver reproducibility of 46%. Interobserver reliability was 47%. They reported improved intrarater and interrater reliability when combining grades 0 and 1 of the grading system of Altchek and Dines.[10] The combination of grades 0 and 1 in their study was similar to the combination of grades I and I+ used in this investigation. They also concluded that caution should be used when diagnosis and treatment are being determined on the basis of this type of manual examination because of the compromised intrarater and interrater reliability.[10] This caution is particularly applicable when examining unanesthetized shoulders with anterior laxity while interpreting results between examiners.[10]

Hawkins et al[7] quantified glenohumeral translation as a percent of humeral head width in 3 groups of anesthetized patients using a manual load-and-shift test in the supine position. Specific patterns of translation were identified among normal subjects, those with unidirectional glenohumeral joint instability, and those with multidirectional instability. However, no measure of intrarater or interrater reliability was reported.

Using an ultrasonic measurement technique, Krarup et al[9] documented anterior humeral head translation in normal asymptomatic subjects and in patients diagnosed with anterior glenohumeral joint instability. Bilateral differences of 1.9 mm in healthy shoulders and 4.9 mm in unstable shoulders were reliably documented with patients in the seated position.

Ellenbecker et al[4] studied the relationship between manual anterior humeral head translation tests and stress radiography in professional baseball pitchers. No statistically significant correlation was found between the amount of humeral head displacement documented radiographically and the manual grading of humeral head translation.

Because the classification scheme proposed by Altchek and Dines[1] was originally designed for preoperative examination with patients under anesthesia, most clinical evaluations use only grades I, I+, II, and II+. Only grades I, I+, II, and II+ were expected and ultimately encountered during this investigation,

as grade III or III+ involves complete dislocation of the glenohumeral joint. This finding was also reported by Lintner et al[11] in a study of 76 asymptomatic division I collegiate athletes, in which no athlete was found to have III or III+ glenohumeral joint translation in any direction. Their study also indicated that translations of up to grade II+ in any direction may be a normal finding.[11] This was supported in our study, in which asymptomatic subjects were found to have up to grade II+ anterior humeral head translation.

The challenge of documenting and describing humeral head translation is evident in the orthopaedic literature. Typical clinical application of humeral head translation tests involves comparison of the involved extremity to the uninvolved extremity.[5] Unilateral increases in anterior humeral head translation in symptomatic patients are one of the clinical signs that lead clinicians to the diagnosis of anterior glenohumeral joint instability.[5,11] However, the presence of multiple grading scales, such as the system of Altchek and Dines,[1] percent width estimation of the humeral head,[6] and others,[9,14,15] as well as numerous manual and instrumented techniques,[4-6,14,15,17] has complicated the interpretation of manual anterior humeral head translation tests. Reproduction of the patient's symptoms is pivotal during anterior humeral head translation and ultimately aids the clinician in diagnosing anterior glenohumeral joint instability.[8]

Varying degrees of intrarater and interrater reliability have been reported by experienced examiners using manual humeral head translation tests.[10,15] One clear finding is the enhancement of reliability when combining the grading levels used to differentiate humeral head translation. It appears that the most reliable aspect of the manual anterior humeral head translation test lies in the determination of whether the humeral head traverses the glenoid rim.[10] This is evident from this study by both the lower percent agreement and the lower κ coefficients found with the classification using grades I, I+, II, and II+ versus the classification using grades I and II. Both the study of Levy et al[10] and the current study support this finding.

On the basis of the findings of this study, the use of the simple grades (ie, I and II) of humeral head translation, as proposed by Altchek and Dines,[1] appears to be warranted. The additional delineation of humeral head translation through the use of end-feel classification cannot be recommended.

Conclusion

Intrarater and interrater reliability of a standard manual anterior humeral head translation test was higher when grading classifications did not require a distinction for end-feel. The addition of end-feel produced fewer test-retest translation matches and lower

coefficients of agreement for both intratester and intertester paradigms. Overall, poor interrater reliability and only fair intrarater reliability were measured for the manual technique used to assess anterior humeral head translation. Enhancement of reliability can be expected by using only grades of interpretation that focus on the relationship of the humeral head to the glenoid rim, in the presence of an anteriorly directed force. Further research is clearly needed to further delineate and identify the role of manual humeral head translation testing in the complete evaluation and treatment of the patient with shoulder pathology.[16]

REFERENCES

1. Altchek DW, Dines DW. The surgical treatment of anterior instability: selective capsular repair. Op Tech Sports Med 1993;1: 285-92.
2. Chesworth BM, MacDermid JC, Roth JH, Patterson SD. Movement diagram and "end-feel" reliability when measuring passive lateral rotation of the shoulder in patients with shoulder pathology. Phys Ther 1998;78:593-601.
3. Cyriax JH, Cyriax PJ. Illustrated manual of orthopaedic medicine. London: Butterworths; 1983.
4. Ellenbecker TS, Mattalino AJ, Elam E, Caplinger R. Quantification of anterior translation of the humeral head in the throwing shoulder: manual assessment versus stress radiography. Am J Sports Med 2000;28:161-7.
5. Gerber C, Ganz R. Clinical assessment of instability of the shoulder (with special reference to anterior and posterior drawer tests). J Bone Joint Surg Br 1984;66:551-6.
6. Hawkins RJ, Schutte JP, Janda DH, Huckell GH. Translation of the glenohumeral joint with the patient under anesthesia. J Shoulder Elbow Surg 1996;5:286-92.
7. Harryman DT, Sidles JA, Harris SL, Matsen FA. Laxity of the normal glenohumeral joint: a quantitative in vivo assessment. J Shoulder Elbow Surg 1992;1:66-76.
8. Jobe FW, Kvitne RS. Shoulder pain in the overhand or throwing athlete: the relationship of anterior instability and rotator cuff impingement. Orthop Rev 1989;18:963-75.
9. Krarup AL, Court-Payen M, Skjoldbye B, Lausten GS. Ultrasonic measurement of the anterior translation in the shoulder joint. J Shoulder Elbow Surg 1999;8:136-41.
10. Levy AS, Lintner S, Kenter K, Speer KP. Intra- and interobserver reproducibility of the shoulder laxity examination. Am J Sports Med 1999;27:460-3.
11. Lintner SA, Levy A, Kenter K, Speer KP. Glenohumeral translation in the asymptomatic athlete's shoulder and its relationship to other clinically measurable anthropometric variables. Am J Sports Med 1996;24:716-20.
12. McFarland EG, Campbell G, McDowell J. Posterior shoulder laxity in asymptomatic athletes. Am J Sports Med 1996;24:468-71.
13. O'Brien SJ, Neves MC, Arnoczky SJ, Rozbruck SR, DeCarlo EF, Warren RF, et al. The anatomy and histology of the inferior glenohumeral ligament complex of the shoulder. Am J Sports Med 1990;18:449-56.

14. Oliashirazi A, Mansat P, Cofield RH, Rowland CM. Examination under anesthesia for evaluation of anterior shoulder instability. Am J Sports Med 1999;27:464-8.
15. Van Duijn AJ, Meers MB, Jensen RH. Reliability of accessory joint mobility testing of the glenohumeral joint [abstract]. Phys Ther 1999;79:S25.
16. Walch G, Boileau P, Noel E, Donell ST. Impingement of the deep surface of the supraspinatus tendon on the posterosuperior glenoid rim: an arthroscopic study. J Shoulder Elbow Surg 1992;1:238-45.
17. Warner JJP, Micheli LJ, Arslanian LE, et al. Patterns of flexibility, laxity, and Strength in normal shoulders and shoulders with instability and impingement. Am J Sports Med 1990;18:366-75.