ORIGINAL ARTICLE

Acta Orthop Traumatol Turc 2014;48(2):169-174
doi: 10.3944/AOTT.2014.3184

Clinical evaluation of the anterior translation of glenohumeral joint using ultrasonography: an intra- and inter-rater reliability study

Leonard H. JOSEPH[1,2], Rizuana I. HUSSAIN[3], Ubon PIRUNSAN[1], Amaramalar S. NAICKER[4], Ohnmar HTWE[4], Aatit PAUNGMALI[1]

[1]*Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand;*
[2]*Physiotherapy Program, School of Rehabilitation Sciences, Faculty of Allied Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia;*
[3]*Department of Radiology, Faculty of Medicine, Universiti Kebangsaan Malaysia Medical Centre, Kuala Lumpur, Malaysia;*
[4]*Department of Orthopedics, Faculty of Medicine, Universiti Kebangsaan Malaysia Medical Centre, Kuala Lumpur, Malaysia*

Objective: The aim of this study was to investigate the intra- and inter-rater reliability of ultrasonography (US) to measure anterior translation of the humeral head (ATHH) among healthy subjects and patients with sacroiliac joint dysfunction.

Methods: The study included a total of 22 shoulder joints from 11 subjects. Six subjects were healthy and 5 had sacroiliac joint dysfunction. Anterior translation of the humeral head was measured twice using US by two different investigators. Intraclass correlation coefficient (ICC3,1), standard error of measurements (SEMs), coefficient of variations (CVs) and Bland-Altman plot were used as analytical tests to investigate intra- and inter-rater reliability, amount of error and agreeability of the measurements between investigators.

Results: Intraclass correlation coefficient was 0.94, showing a high level of intra-rater reliability of the first investigator with SEMs (0.01 cm) and CV (5.1%) in measuring ATHH. Intra-rater reliability of the second investigator was 0.84 with SEMs (0.03 cms) and CV (9.6%), indicating a high level of reliability. Inter-rater reliability was high, with an ICC value of 0.92 with SEMs (0.02 cms) and CV (5.9%).

Conclusion: The use of US as a measurement of ATHH has good levels of intra- and inter-rater reliability in clinical practice.

Key words: Examination; humeral head; reliability; shoulder; translation; ultrasound.

Quantification of humeral head (HH) translation during an applied translator force may be useful in the diagnosis of glenohumeral joint (GHJ) pathologies.[1,2] Additionally, measuring the translation of HH helps to understand the pathomechanics of the GHJ and assist in the choice of conservative and/or surgical intervention for GHJ disorders.[3] In clinical practice, clinicians measure the anterior translation of the humeral head

Correspondence: Aatit Paungmali, PT, Assoc. Prof. Department of Physical Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai 50200, Thailand.
Tel: +66-53 949 246 e-mail: aatit.p@cmu.ac.th

Submitted: September 02, 2013 **Accepted:** January 06, 2014

Available online at
www.aott.org.tr
doi: 10.3944/AOTT.2014.3184
QR (Quick Response) Code

(ATHH) in order to diagnose GHJ anterior translation dysfunction and other shoulder problems.[1,4]

Previous studies have reported the use of several instruments, including X-ray, fluoroscopy, magnetic resonance imaging and electromagnetic tracking device to measure ATHH in both normal and symptomatic subjects.[1,3,5,6] These techniques have limitations such as exposure to radiation, expense, and unavailability in all clinics.[1,3,5-7] On the other hand, manual palpation techniques used to grade ATHH are subjective in nature and yielded poor reproducibility of findings.[4]

In the current study, real-time ultrasonography (US) was used to quantify ATHH as an objective measure among subjects. A limited number of studies have been reported in the literature investigating the reliability of US in quantifying ATHH in GHJ.[8-10] Thus, it is of practical significance to determine the reliability of US before its use in research studies and clinical practice. When using a new method to quantify the ATHH of the GHJ, it is important to evaluate the reliability of the measurement procedure for the trustworthiness of the data.[11]

In addition, the reliability characteristics of a measure and investigators are influenced by the characteristics of the subjects from a specific disease population.[12] A recently published guideline on reliability studies (Guidelines for Reporting Reliability and Agreement Studies, GRRAS) stated that the measurement of reliability in a larger longitudinal study design must be tested among the subjects of the original study.[12] The US evaluation of ATHH is a variable of interest in an ongoing original study on sacroiliac joint dysfunction. Hence in the current study, the reliability of US in measuring the ATHH was determined in a population of normal individuals and subjects with sacroiliac joint dysfunction.

The main aim of this study was to investigate the intra- and inter-rater reliability of US to measure ATHH objectively in a population of healthy subjects and patients with sacroiliac joint dysfunction.

Patients and methods

The study included a total of 22 shoulder joints of 11 subjects (mean age: 41.7±6.2 years; mean weight: 66.4±6.3 kg; mean height: 1.68±0.5 cm). Six subjects were healthy and 5 had sacroiliac joint dysfunction. All subjects with sacroiliac joint dysfunction were recruited from a University tertiary hospital. The diagnosis of sacroiliac joint dysfunction was made if a subject showed positive response to at least four of five clinical mobilization tests, including the Gillet test, standing flexion test, prone knee flexion test, supine long sitting test and palpation of posterior iliac spine asymmetry on sitting.[13-15] Healthy subjects were recruited as a matched group from among patient caregivers who accompanied other patients to the hospital. All subjects had full range of shoulder motion with no history of any symptoms in shoulder joint and the ability to walk independently without ambulation aid. Subjects with shoulder pathology, presence of pain in the shoulder, shoulder injury occurring in the previous 3 months or with any past history of shoulder surgery were excluded. Subjects were briefed about the study details and written informed consent was obtained prior to participation. The University Hospital Ethical Committee provided ethical approval for this study with ethical code NN-181-2011.

The ATHH was measured using US (The Philips iU22 xMATRIX Ultrasound System; Philips Healthcare, Best, the Netherlands) on B mode through a linear transducer of 3.5 MHz based on the established protocol.[8-10] Shoulders were exposed and subjects positioned on a chair with backrest with both legs resting on the floor to ensure maximum stability during testing. The shoulder was kept in medial rotation in an adducted position alongside the trunk with the forearm facing to body and elbow flexed and supported with another hand. The investigator stood in a walk stance position behind the subject. Ultrasonography imaging of the shoulder translation was performed by a qualified radiologist.

During testing, the linear transducer with an Aquasonic gel was placed on the anterior aspect on the shoulder and three well-defined bony landmarks, the greater tubercle of the humerus, the coracoids process of the scapula and the anterior-superior part of the neck of the scapula, were identified and captured by the radiologist. In this position, the placement of the transducer on the skin was marked. The resting position of the HH was measured by placing the cursor on the coracoids process of the scapula, the neck of the scapula and the top of the greater tubercle in the captured image. The distance between the scapular neck and the top of greater tubercle was measured as the shoulder resting distance (d_1). A total of 3 trials were carried out and the average of the 3 readings was taken as the final measurement.

On the tested side, the acromion process and HH were palpated and the joint line was identified by the first investigator. The shoulder girdle along with the scapula was stabilized by the investigator using one hand while the other hand applied a manual posteroanterior translation of the HH to find the best angle of translation. Using a push-pull dynamometer, the investigator applied a translator force of 80 N to the posterior part of the HH to passively translate the HH anteriorly to the point of end feel. The bony landmarks of the shoulder post translation were measured by placing the cursor on the coracoid process of the scapula, the neck of the scapula and the top of the greater tubercle. The distance between the neck of the scapula and the top of the greater tubercle after translator force was measured and recorded as post-translation distance (d_2). An average of three measurements was taken for final reading of d_2. The ATHH was calculated as the difference between the distance measured during passive anterior translation (d_2) and at rest (d_1) (Fig. 1). The measurement of the ATHH was performed first on one shoulder and then the other. The ATHH was measured for both shoulders at the first testing and the second measure was repeated after 48 hours using the same method. The complete procedure was performed by the first investigator (LJ) who is a senior musculoskeletal physiotherapist with 14 years of experience. The whole procedure was then repeated by the second investigator (RH), a junior physiotherapist. The order of measurements in the shoulders and the investigators were randomized in order to prevent any possible memory bias. Investigators were blinded on the first reading during the second occasion of measurement. The same radiologist captured the US images of the ATHH for both investigators throughout the study.

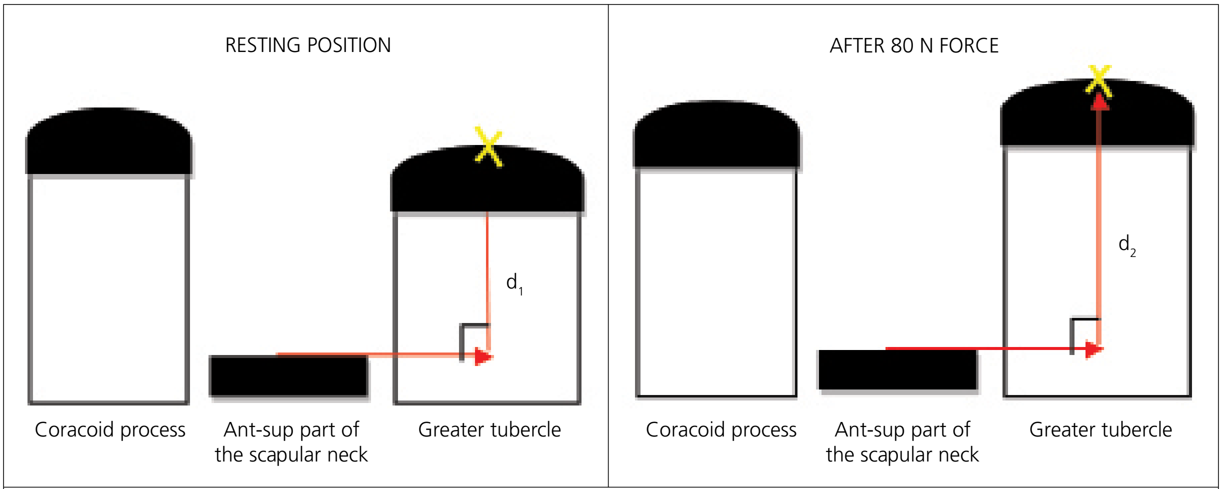
Data were analyzed using the statistical software package SPSS for Windows v20.0 (SPSS Inc., Chicago, IL, USA). Intraclass correlation coefficient (ICC3,1), standard error of measurements (SEMs), coefficient of variations (CVs) and Bland-Altman plot were used to investigate the intra- and inter-rater reliability, the amount of error and agreeability of the measurements between the investigators.[11,12,16-18] The differences in the first and second measures were analyzed using the paired sample t-test with the level of significance as 0.05. The SEMs were calculated using the formula below, where SD is the standard deviation and ICC is the intraclass correlation coefficient:

$$SEMs = SD \times sqrt(1-ICC)$$

The SEMs were used as a measure of absolute reliability of the measurements.[11] The estimation of error measurement was calculated through CV by SD of the data divided by the mean and multiplied by 100 to give a percentage score.[18] Intraclass correlation coefficient values of <0.50 were determined as 'low', 0.50 to 0.75 'moderate', and >0.75 'high'.[19,20] The Bland-Altman plot was depicted to measure the limits of agreement between the measurements of the two investigators. The limits of agreement equaling two SD of the mean difference above and below the mean were plotted and the variance in measurements outside ±2SD was examined through visual analysis in the Bland-Altman plot.[17,18]

Results

The mean±SD ATHH values of all measurements taken by both investigators are shown in Table 1. Intraclass cor-



Anterior translation of the humeral head = d_2 - d_1

d_1: distance between highest part of the greater tubercle perpendicular to imaginary line of anterior-superior neck of the scapula during resting.

d_2: distance between highest part of greater tubercle perpendicular to imaginary line of anterior-superior neck of the scapula during application of 80 N of force.

Fig. 1. Picture description of anterior translation of the humeral head (adapted from Court-Payen et al.'s[10] study). [Color figure can be viewed in the online issue, which is available at www.aott.org.tr]

Table 1. Intra-rater reliability with mean and standard deviation of the ATHH measures.

Investigators	Measurement of ATHH Mean±SD (cm)		Intraclass Correlation Coefficient (ICC) with 95% CI	95% Confidence Interval	Standard Error of Measurements (cm)	Coefficient of Variations (%)
	1st Measure	2nd Measure				
1st investigator	1.01±0.24	0.98±0.22	0.94	0.86-0.97	0.01	5.1
2nd investigator	1.02±0.24	0.98±0.21	0.84	0.49-0.95	0.03	9.6

CI: Confidence interval; SD: Standard deviation.

relation coefficient was 0.94, showing high intra-rater reliability of the first investigator with SEMs of 0.01 cm and CV of 5.1%. Intra-rater reliability of the second investigator was 0.84 with SEMs of 0.03 cms and CV of 9.6%, indicating a high level of repeatability. The inter-rater reliability between the first and second investigator was considered high with an ICC value greater than 0.92 with SEMs of 0.02 cms and CV of 5.9% for all measurements.

No statistical significant difference was found between the two measures (t[21]=0.44, p=0.66, 95% CI -0.27 to 0.04) taken by the first observer. Similarly, there was no statistical significance between the two measurements taken by the second investigator (t[21]=1.54, p=0.13, 95% CI -0.01 to 0.06). Visual analysis of the Bland-Altman plot showed that 95% of all measurement differences were in between the ±2SD, indicating that the agreement between measurements taken by both investigators were acceptable (Fig. 2).

Discussion

This study investigated the intra- and inter-rater reliability of an US technique used to measure the ATHH in

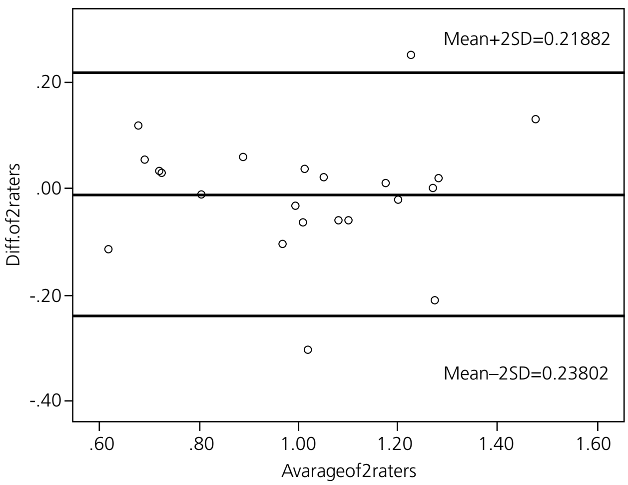


Fig. 2. Bland-Altman plot of the data presenting differences between the two measures against mean for the two measures indicating good limits of agreement between measurements, with 95% of the measures lying within ±2SD of the mean.

the GHJ. Limited studies are available reporting on this US technique and the evidence on the acceptable level of reliability is inconclusive. Hence, a separate study on the reliability of the US technique to quantify ATHH is essential before it can be applied to clinical and research practice. The results showed that the methods and measurements used in the current study are reliable between and within the investigators.

An US method with 90 N of force used to quantify ATHH was first proposed by Court-Payen et al. although reliability was not reported for this technique in their study.[10] Later, Krarup et al. investigated the ATHH using US with a force of 90 N among normal individuals and subjects with traumatic anterior shoulder dislocation in 3 different shoulder positions and concluded that the intra-observer reliability was considerable.[9] Yeap et al. used US to evaluate ATHH with two different forces of 90 N and 60 N in different positions of the shoulder and reported poor reliability.[8] In our opinion, the results of the above studies cannot be considered comparable to the current study because of the difference in force levels (90 N and 60 N) compared to the 80 N of force used in the current study. In addition, the differences in the shoulder positions, procedural differences and different subject populations were not comparable between the studies. Moreover, none of the above studies reported reliability analysis as a separate process and the study design was not meant to test the reliability of the technique and investigators. Thus, the current study was designed to evaluate the reliability of the US technique to quantify ATHH among a specific population of patient with sacroiliac joint dysfunction and normal subjects.

The current report is part of a main ongoing research study on the sacroiliac joint investigating the effects of sacroiliac joint dysfunction on global muscles, local muscles and the locomotive system. This original study is a biomechanical study in which the altered biomechanical effects due to sacroiliac joint dysfunction are evaluated in shoulder joint due to global muscle connection. In particular, one of the objectives of the main study is to examine the myofascial force transmission from the sacroiliac joint to the contralateral shoulder joint for which ATHH is one of the variables of interest. As per the recently published GRRAS guidelines for reliability, any report on reliability issues must be dealt with as a separate scientific entity.[12] The GRRAS guidelines recommend that the examination of reliability should not be reported as part of the main study, but separately from the main study on a particular specific population of interest. As reliability of the measurement procedures might be influenced by subject characteristics of the specific population of interest, reliability has to be examined in a specific population.[12] Thus, the current study aimed to report the reliability of US for the evaluation of the ATHH among a population of participants with sacroiliac joint dysfunction and healthy subjects.

Anterior translation of the humeral head measurements have been investigated by various studies in the past using different methods. Harryman et al.[21] evaluated the ATHH using an electromagnetic tracker and induced 7.8±4.0 mm of anterior translation in healthy subjects. Nevertheless, this procedure was invasive and not generally applicable in a clinical situation. Levy et al.[22] assessed the ATHH through a manual technique and reported an overall intra-observer reliability of 46% and inter-observer reliability of 47%. Yet, due to the subjective nature of the grading used to assess the AHTT and low level of reliability, the findings should be interpreted with caution. Hawkins et al.[7] quantified ATHH using a manual load and shift technique in anesthetized patients. However, there was no report of reliability and patients were anesthetized which is not applicable in clinical situations. Due to the differences in methods, equipment and techniques of measurements as well as patient population, it is very difficult to compare data on the ATHH. Nevertheless, the technique and methods used in the current study were shown to be reliable.

Several factors have been reported to influence US measurement of the ATHH in the GHJ. Such factors include; bulky tissues surrounding the bony landmarks in the shoulder, differences in the position and size of the shoulder, difference in the location of the placement of measurement tool (US transducer/tracker), fixation of the scapula, different rates of force application to cause ATHH, differences in the applied forces, and level of experience of the investigator.[8,23-25] Additionally, US imaging of the shoulder requires technical expertise and prior training.[23] In the current study, the above factors were taken into account prior to the measurement technique and addressed meticulously. Collective training between the investigators including the radiologist prior to the commencement of the study were performed to identify the bony landmarks manually, angle of translation with and without force and for identification of bony landmarks precisely with US over several sessions. In order to prevent placement errors of the transducer on the bony landmark, the placement spots were marked on the skin of each subject. Investigators were trained to stabilize the shoulder and posture of the subjects during application of the 80 N force in order to prevent postural buckling during the force application and measurement.

No single statistical measure is sufficient to comment on the reliability of measurements.[12] Therefore, the use of various measures of reliability such as ICC, SEMs, CV and the Bland-Altman plot may be considered the strength of the study. There were some limitations in this study. The small number of participants with the rare specific condition of sacroiliac joint dysfunction included in our study may be considered as one of the main limitations and future studies with a larger sample size should be considered. The reliability reported is applicable only to the population of interest and to the specific procedure mentioned in the current study. Hence, the application of the technique to other populations with revisions in methodology may alter the repeatability and in such cases, the reliability should be examined separately. However, the procedure to measure ATHH proposed in this study may serve as a guide for clinicians and researchers in their further research.

In conclusion, the intra- and inter-rater reliability of the US method used in this study for the measurement of the ATHH was acceptable and reliable. Quantification of the ATHH using US may be considered a useful method to investigate GHJ problems.

Acknowledgement: The authors wish to thank Ms. Foong Yee KUAN, Physiotherapist, for her assistance in this study.

Conflicts of Interest: No conflicts declared.

References

1. 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.
2. 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.
3. von Eisenhart-Rothe R, Müller-Gerbl M, Wiedemann E, Englmeier KH, Graichen H. Functional malcentering of the humeral head and asymmetric long-term stress

on the glenoid: potential reasons for glenoid loosening in total shoulder arthroplasty. J Shoulder Elbow Surg 2008;17:695-702. CrossRef
4. Taylor JD, Bandy WD. Intrarater reliability of the KT1000 arthrometer in determining anterior translation of the glenohumeral joint. Arch Phys Med Rehabil 2005;86:826-9. CrossRef
5. Norris TR. C-arm fluoroscopic evaluation under anesthesia for gleno-humeral subluxation. In: Bateman JE, Welsh RP, editors. Surgery of the shoulder. Philadelphia, PA: BC Becker, Inc; 1984. p. 22-5.
6. Lin HT, Hsu AT, Chang GL, Chang Chien JR, An KN, Su FC. Determining the resting position of the glenohumeral joint in subjects who are healthy. Phys Ther 2007;87:1669-82. CrossRef
7. 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. CrossRef
8. Yeap JS, McGregor AH, Humphries K, Wallace AL. Ultrasonic evaluation of anterior shoulder translation in normal shoulders. J Musculoskeletal Res 2003;7:125-34. CrossRef
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. CrossRef
10. Court-Payen M, Krarup AL, Skjoldbye B, Lausten GS. Real-time sonography of anterior translation of the shoulder: an anterior approach. EJU 1995;2:283-7.
11. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217-38. CrossRef
12. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011;64:96-106. CrossRef
13. Arab AM, Abdollahi I, Joghataei MT, Golafshani Z, Kazemnejad A. Inter- and intra-examiner reliability of single and composites of selected motion palpation and pain provocation tests for sacroiliac joint. Man Ther 2009;14:213-21. CrossRef
14. Tong HC, Heyman OG, Lado DA, Isser MM. Interexaminer reliability of three methods of combining test results to determine side of sacral restriction, sacral base position, and innominate bone position. J Am Osteopath Assoc 2006;106:464-8.
15. Cibulka MT, Koldehoff R. Clinical usefulness of a cluster of sacroiliac joint tests in patients with and without low back pain. J Orthop Sports Phys Ther 1999;29:83-92. CrossRef
16. Khamwong P, Nosaka K, Pirunsan U, Paungmali A. Reliability of muscle function and sensory perception measurements of the wrist extensors. Physiother Theory Pract 2010;26:408-15. CrossRef
17. Bruton A, Conway JH, Holgate ST. Reliability: what is it and how is it measured. Physiotherapy 2000;86:94-9. CrossRef
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10. CrossRef
19. Numanoğlu A, Günel MK. Intraobserver reliability of modified Ashworth scale and modified Tardieu scale in the assessment of spasticity in children with cerebral palsy. Acta Orthop Traumatol Turc 2012;46:196-200. CrossRef
20. Günel MK, Tarsuslu T, Mutlu A, Livanelioğlu A. Investigation of interobserver reliability of the Gillette Functional Assessment Questionnaire in children with spastic diparetic cerebral palsy. Acta Orthop Traumatol Turc 2010;44:63-9. CrossRef
21. Harryman DT 2nd, Sidles JA, Harris SL, Matsen FA 3rd. Laxity of the normal glenohumeral joint: A quantitative in vivo assessment. J Shoulder Elbow Surg 1992;1:66-76. CrossRef
22. 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.
23. Bahk M, Keyurapan E, Tasaki A, Sauers EL, McFarland EG. Laxity testing of the shoulder: a review. Am J Sports Med 2007;35:131-44. CrossRef
24. Sauers EL, Borsa PA, Herling DE, Stanley RD. Instrumented measurement of glenohumeral joint laxity: reliability and normative data. Knee Surg Sports Traumatol Arthrosc 2001;9:34-41. CrossRef
25. Borsa PA, Sauers EL, Herling DE, Manzour WF. In vivo quantification of capsular end-point in the nonimpaired glenohumeral joint using an instrumented measurement system. J Orthop Sports Phys Ther 2001;31:419-31. CrossRef