RESEARCH ARTICLE

Computed Tomography and Magnetic Resonance Imaging are Similarly Reliable in the Assessment of Glenohumeral Arthritis and Glenoid Version

Christopher M. Hopkins, MD[1]; Frederick M. Azar, MD[1]; Ryan P. Mulligan, MD[2]; Anthony M. Hollins, MD[3]; Richard A. Smith, PhD[1]; Thomas W. Throckmorton, MD[1]

Research performed at University of Tennessee-Campbell Clinic Department of Orthopaedic Surgery, Memphis, TN, USA

Received: 11 March 2020 *Accepted: 14 November 2020*

Abstract

Background: The purpose of this study was to compare the intraobserver and interobserver reliability of CT and T2-weighted MRI for evaluation of the severity of glenoid wear, glenohumeral subluxation, and glenoid version.

Methods: Sixty-one shoulders with primary osteoarthritis had CT and MRI scans before shoulder arthroplasty. All slices were blinded and randomized before evaluation. Two fellowship-trained shoulder surgeons and three orthopaedic surgery trainees reviewed the images to classify glenoid wear (Walch and Mayo classifications) and glenohumeral subluxation (Mayo classification). Glenoid version was measured using Friedman's technique. After a minimum two-week interval, the process was repeated.

Results: Intraobserver reliability was good for the CT group and fair-to-good for the MRI group for the Walch, Mayo glenoid, and Mayo subluxation classifications; interobserver reliability was poor for the CT and fair-to-poor for the MRI group. For the measurement of glenoid version, intraobserver reliability was good for the CT and substantial for the MRI group; interobserver agreement was good for both groups. There were no significant differences in reliability between staff surgeons and trainees for any of the classifications or measurements.

Conclusion: CT and MRI appear similarly reliable for the classification of glenohumeral wear patterns. For the measurement of glenoid version, MRI was slightly more reliable than CT within observers. Differences in training level did not produce substantial differences in agreement, suggesting these systems can be applied by observers of different experience levels with similar reliability.

Level of evidence: III

Keywords: Assessment, Computed tomography, Glenoid version, Glenohumeral arthritis, Magnetic resonance imaging

Introduction

Glenohumeral arthritis is a common condition that has been reported to affect more than 25% of individuals over the age of 60 years (1, 2). The use of shoulder arthroplasty to manage glenohumeral arthritis has well-documented success, with 84% survivorship at 20 years (3). As with any surgical procedure, preoperative planning is critical to anticipate intraoperative problems and to be prepared to modify the surgical procedure if needed. In shoulder arthroplasty, accurate placement of glenoid components has been shown to be important for stability and to decease loosening. Studies have demonstrated reduced survival of glenoid components

Corresponding Author: Thomas W. Throckmorton, University of Tennessee-Campbell Clinic Department of Orthopaedic Surgery and Biomedical Engineering, Memphis, Tennessee, USA
Email: tthrockmorton@campbellclinic.com

THE ONLINE VERSION OF THIS ARTICLE ABJS.MUMS.AC.IR

Table 1. Mayo classification for glenohumeral subluxation and glenoid erosion

	None	Mild	Moderate	Severe
Glenohumeral subluxation	0	<25%	25-50%	>50%
Glenoid erosion	0	To subchondral bone	Medialization of glenoid, hemispheric deformation	Bone loss to coracoid base

when the glenoid component is malpositioned (4-6). The importance of accurately assessing glenoid wear and version was further emphasized in a cadaver study by Gillespie *et al*., who demonstrated that eccentric reaming to correct 15-degrees of glenoid retroversion resulted in the inability to place the glenoid component because of inadequate bone stock in four of the eight specimens tested (7). Correction of 20-degree deformities resulted in deficient bone stock in six of eight specimens.

Currently, the Walch classification is the most widely used to describe glenoid deformity. Although Walch reported substantial interobserver and intraobserver agreement in the original description of the classification, others have reported less reliable agreement, with only fair agreement (3, 8, 9). Sperling *et al*. proposed a different classification system (Mayo classification) to characterize glenohumeral subluxation and glenoid erosion [Table 1] (3). This has been found to have agreement similar to the Walch classification system in the assessment of glenohumeral arthritis (10).

Preoperative assessment of glenoid deformity has been done by axillary radiography (AXR), computed tomography (CT) and magnetic resonance imaging (MRI) (11-14). AXR, though the least expensive and easiest to obtain, has demonstrated poor intraobserver and interobserver reproducibility and has been shown to overestimate the degree of glenoid retroversion (13). CT scan allows excellent assessment of bony anatomy and has been shown to have moderate interobserver reliability and substantial intraobserver reliability in classifying glenohumeral arthritis (11). While many surgeons prefer CT as a preoperative planning tool because of the presumed advantage of better illustration of bony detail, it has been shown to compare favorably to MRI in demonstrating osseous detail and has the added benefit of imaging the soft-tissues structure, in particular the rotator cuff, that may influence the surgical technique (13, 15-17). It is unknown, however, how CT compares to MRI in classifying glenohumeral arthritis and glenoid version. The purpose of this study was to compare intraobserver and interobserver agreement for the assessment of glenoid wear (Walch and Mayo classifications), glenohumeral subluxation (Mayo), and measurements of glenoid version using CT and MRI.

Materials and Methods

After institutional review board approval, a retrospective review of electronic medical records at our institution identified 61 consecutive patients (61 shoulders) with primary glenohumeral osteoarthritis who had CT and MRI scanning as part of the pre-operative evaluation before shoulder arthroplasty between 2011-2014. Patients with incomplete imaging (no MRI and/or CT), revision arthroplasty, and a diagnosis other than primary shoulder osteoarthritis (e.g. inflammatory arthritis, post-traumatic arthritis) were excluded. All patients had CT and MRI scans as part of their preoperative evaluation. Of note, MRI was not part of the standard preoperative workup for all patients undergoing shoulder arthroplasty at our institution during the study period. Rather, CT was the routine diagnostic test of choice to assess glenoid deformity, and MRI typically was obtained by a referring physician or if there was suspicion of rotator cuff dysfunction.

A 64-detector CT scanner (Lightspeed VCT, GE, Fairfield, CT) was used for CT imaging. Axial cuts with 2.5-mm thickness were obtained. Patients were positioned supine with the arm at the side. MRI scans of the shoulder were obtained with a Siemens 3-Tesla MR scanner (Siemans AG, Erlangen, Germany) with the patient's hand at his or her side in neutral position. Axial slice thickness was 3 mm with a 16 cm field of view.

Demographic data, including age, sex, affected side, operative side, and body mass index (BMI) were recorded. CT and MRI images were then reviewed and single best-representative axial slices at the center of the glenoid (defined as 1 cm inferior to the tip of the coracoid process) were captured and saved into separate PowerPoint files for assessment of glenoid deformity (Walch and Mayo) and glenohumeral subluxation (Mayo). The same images were then uploaded to our radiographic archiving and communication system (PACS) to use the angle measurement function to assess glenoid version.

Glenoid version was measured using the method described by Friedman et al (18). The anterior and posterior edges of the glenoid were marked with one line and then the transverse axis of the scapula was marked with a line drawn from the center of the glenoid down the medial border of the scapular spine. A line was then drawn perpendicular to this to define neutral version [Figure 1]. If the posterior margin of the glenoid was medial to the line of neutral version and the line connecting the anterior and posterior margins of the glenoid, this was considered retroversion and was recorded as a negative number. Anteversion was recorded as a positive value.

Glenoid morphology was classified according the Walch method (19). This classification system classifies glenoids into five groups: A1,A2, B1, B2, and C. Type A1 glenoids are defined as those with minor central wear and type A2 has more severe central wear. B1 has mild posterior glenoid erosion with posterior subluxation. B2 glenoids have erosion of the posterior glenoid creating a biconcave appearance. Type C glenoids are dysplastic with retroversion exceeding 25 degrees [Figure 2].

THE ARCHIVES OF BONE AND JOINT SURGERY. ABJS.MUMS.AC.IR
VOLUME 9. NUMBER 1. JANUARY 2021

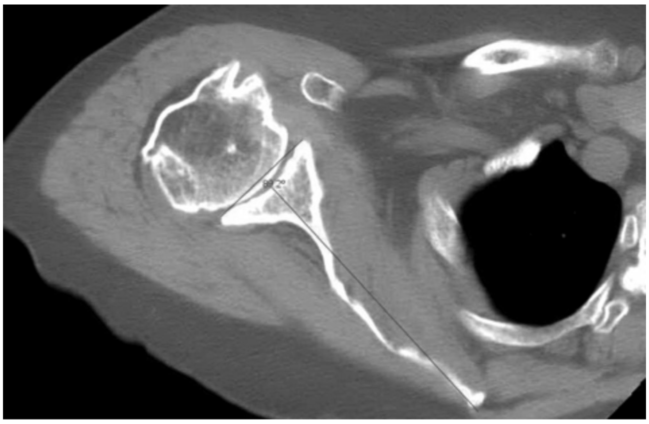

Figure 1. Measurement of glenoid version (18).

Glenoid erosion and subluxation were also assessed using the Mayo classification [Table 1] (3). All images were assessed by two staff physicians with fellowship training in shoulder surgery and two orthopaedic trainees. All images were then reassessed again at two weeks to determine intraobserver agreement. Study participants were blinded to patient identity to avoid bias.

Statistical analysis was performed SPSS version 22 (Armonk, NY: IBM Corp) for Mac. Intraobserver reliability for the Walch, Mayo glenoid, and Mayo subluxation classifications were determined using Spearman's correlation coefficient while interobserver agreement was determined using Congers kappa. Reliability of glenoid version measurements was determined using the Pearson correlation coefficient for both interobserver agreement and intraobserver reliability. Kappa values greater than 0.8 were considered to indicate substantial agreement, values between 0.6-0.8 good agreement, values between 0.4-0.6 fair agreement and values less than 0.4 were considered to indicate poor agreement.

Results

Overall average intraobserver reliability for the CT group was 0.71, 0.73, and 0.64 for the Walch, Mayo glenoid, and Mayo subluxation classifications, respectively, indicating good agreement. Intraobserver reliability for the MRI group was 0.71, 0.52, and 0.62 for the Walch, Mayo glenoid, and Mayo subluxation classifications, respectively, indicating fair-to-good agreement [Table 2]. There were no consistent differences in intraobserver agreement for any of the classification systems.

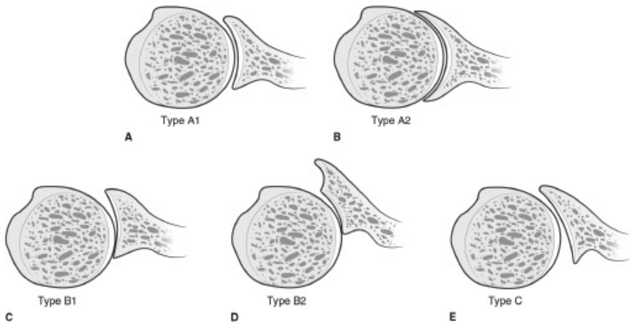

Figure 2. Walch classification of glenoid morphology, based on wear patterns and version. A, Type A1, centered humeral head with minor glenoid erosion. B, Type A2, centered humeral head with major glenoid erosion. C, Type B1, posterior subluxation with no erosion. D, Type B2, posterior erosion with a biconcave glenoid. E, Type C, severe retroversion. (Reproduced with permission from Sears BW, Johnston PS, Ramsey ML, Williams GR. Glenoid bone loss in primary total shoulder arthroplasty: evaluation and management. J Am Acad Orthop Surg 2012; 20:602-613.).

Table 2. Intraobserver reliability for Walch, Mayo (erosion), Mayo (subluxation) using Spearman's correlation coefficient

	CT	MRI
Walch		
Staff 1	.86	.86
Staff 2	.54	.39
Trainee 1	.68	.68
Trainee 2	.60	.75
Trainee 3	.85	.68
Overall Avg	.71	.71
Avg staff	.57	.62
Avg trainee	.71	.67
Mayo (erosion)		
Staff 1	.92	.59
Staff 2	.70	.73
Trainee 1	.74	.67
Trainee 2	.70	.35
Trainee 3	.58	.24
Overall Avg	.73	.51
Avg staff	.81	.66
Avg trainee	.67	.42
Mayo (subluxation)		
Staff 1	.74	.85
Staff 2	.43	.70
Trainee 1	.70	.74
Trainee 2	.59	.69
Trainee 3	.75	.58
Overall Avg	.64	.73
Avg staff	.59	.81
Avg trainee	.68	.67

Table 3. Interobserver agreement for CT and MRI assessment of glenoid deformity using Conger's kappa

	First assessment	Second assessment	Average
CT			
Walch	.28	.30	.29
Mayo (erosion)	.33	.38	.35
Mayo (subluxation)	.37	.29	.33
MRI			
Walch	.27	.41	.34
Mayo (erosion)	.19	.35	.27
Mayo (subluxation)	.37	.43	.40

Table 5. Interobserver agreement for CT and MRI measurement of glenoid version using Pearson's correlation coefficient

	First assessment	Second assessment
CT		
Avg staff	.62	.72
Avg trainee	.77	.72
Avg	.76	.73
Overall Avg	.75	
MRI		
Avg staff	.84	.82
Avg trainee	.79	.71
Avg	.82	.77
Overall Avg	.79	

Interobserver reliability for the CT group was 0.29, 0.35, and 0.33 for the Walch, Mayo glenoid and Mayo subluxation classifications, respectively, indicating poor agreement. Interobserver reliability for the MRI group was 0.34, 0.27, and 0.40 for the Walch, Mayo glenoid, and Mayo subluxation classifications, respectively, indicating poor-to-fair agreement [Table 3].

For the measurement of glenoid version, average intraobserver reliability was 0.79 for the CT group and 0.90 for the MRI group, indicating good agreement in the CT group and substantial agreement in the MRI group [Table 4]. Overall interobserver agreement was 0.75 for the CT group and 0.79 for the MRI group, indicating good agreement [Table 5]. There were no consistent differences in intraobserver and interobserver agreement for measurement of glenoid version between staff physicians and trainees.

Table 4. Intraobserver reliability of glenoid version measurement using Pearson correlation coefficient

	CT	MRI
Staff 1	.89	.82
Staff 2	.80	.88
Trainee 1	.67	.99
Trainee 2	.75	.86
Trainee 3	.82	.96
Overall Avg	.79	.9
Avg staff	.85	.85
Avg trainee	.75	.93

Discussion

To our knowledge, this is the first study to directly compare the intraobserver and interobserver agreement of CT and MRI in the assessment of glenoid erosion, glenohumeral subluxation, and measurement of glenoid version. Our results indicate that MRI is comparable to CT in assessing these variables. For example, CT had better intraobserver agreement for the Mayo glenoid classification but was similar to MRI for the Walch and Mayo subluxation schemes. MRI had slightly superior interobserver reliability for assessing glenohumeral subluxation, but was similar to CT for both the Walch and Mayo glenoid classifications. In the measurement of glenoid version, MRI showed an advantage over CT, demonstrating better intraobserver agreement and similar interobserver reliability. These findings indicate that both MRI and CT can be reliably used to assess glenoid morphology and glenohumeral subluxation by individual surgeons for pre-operative planning before shoulder arthroplasty. The overall fair-to-poor interobserver agreement shown in this study suggests that assessments compared between surgeons are less reliable using either of these pre-operative planning tools.

Other authors have compared different imaging modalities in the assessment of glenoid morphology. Nyffeler *et al*. compared CT and AXR in the assessment of glenoid version in 50 patients (25 with anterior instability and 25 with total shoulder arthroplasty) and found CT to provide excellent agreement and cautioned against using AXR in the assessment of glenoid version (12). Raymond *et al*. compared the reliability AXR to MRI in measuring glenoid version in 48 shoulders with primary osteoarthritis and found MRI to be more reproducible with intraobserver and interobserver reliability coefficients of .96 and .90, respectively (13). However, to date, no study has specifically compared MRI to CT in these assessments.

MRI scanning is commonly used to assess rotator cuff pathology before arthroplasty because these findings can change the operative plan. Edwards *et al*. looked at the effect of rotator cuff disease on the outcomes following TSA for osteoarthritis and found that fatty degeneration of the infraspinatus or subscapularis musculature adversely affected outcome scores (20). Further, MRI avoids the radiation exposure of CT scanning. Biswas *et al*. found an effective dose of 2.06 mSv during shoulder CT scanning (9), which was comparable to the effective dose received during a routine head CT scan (21). The authors also reported a 2-times increase in lifetime cancer risk at this exposure level.

There has been debate regarding the reproducibility of the Walch classification, with the original report demonstrating solid intra- and interobserver reliability with kappa indexes ranging from 0.65 to 0.70 (19). However, Scalise *et al*. reported interobserver agreement of 0.37 and intraobserver reliability of 0.34 using four raters with this system (22). While the inter-rater reliability in this study is consistent with ours, we found the Walch system to have good agreement within observers for both CT and MRI.

The use of three-dimensional CT has demonstrated promise in the literature, but wide-spread use is lacking because of cost and technology barriers (8, 23-25). Hoenecke *et al*. reported intraobserver and interobserver correlation coefficients of .91 to .99 and .95 to .99, respectively, when using three-dimensional CT to assess characteristics of glenoid morphology (8); however, this study did not directly assess the reliability of glenohumeral arthritis classification schemes.

The level of training did not influence the reliability data in this study. Specifically, though there were differences seen between staff physicians and trainees in individual data sets, there were no consistent trends indicating that one group was more reliable than the other. This finding is consistent with that of Scalise *et al*., who also failed to demonstrate a correlation between training level and intra-rater agreement (22).

A potential explanation for our lower agreement could be the process we used to present images to raters. We included a single axial cut for both the CT and MRI groups. Presumably, agreement would be better if the rater had access to the entire MRI or CT scan. Additionally, Bokor *et al*. examined scapular rotation and its effect on glenoid version and found rotation to have a significant effect on measured glenoid version (26). Our study could have been enhanced with a standard positioning protocol used on all imaging studies.

We conclude that MRI and CT are similarly reliable in the assessment of glenoid erosion, glenohumeral subluxation, and glenoid version, making both suitable for surgical planning before shoulder arthroplasty. Additionally, both studies can be reliably applied by observers of differing experience levels.

Christopher M. Hopkins MD[1]
Frederick M. Azar MD[1]
Ryan P. Mulligan MD[2]
Anthony M. Hollins MD[3]
Richard A. Smith PhD[1]
Thomas W. Throckmorton MD[1]
1 University of Tennessee-Campbell Clinic Department of Orthopaedic Surgery and Biomedical Engineering, Memphis, Tennessee, USA
2 Duke University Medical School, Department of Orthopaedic Surgery, Durham, North Carolina, USA
3 Memphis Shoulder and Orthopaedic Surgery, Memphis, Tennessee, USA

References

1. Lo IK, Griffin S, Kirkley A.The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: The Western Ontario Osteoarthritis of the Shoulder (WOOS) index. Osteoarthritis Cartilage 2001; 9(8):771-778.
2. Petersson CJ. Degeneration of the gleno-humeral joint. An anatomical study. Acta Orthop Scand 1983; 54(3):277-283.
3. Sperling JW, Cofield RH, Schleck CD, Harmsen WS. Total shoulder arthroplasty versus hemiarthroplasty for rheumatoid arthritis of the shoulder: results of 303 consecutive cases. J Shoulder Elbow Surg 2007; 16(6):683-690.
4. Farron A, Terrier A, Buchler P. Risks of loosening of a prosthetic glenoid implanted in retroversion. J Shoulder Elbow Surg 2006; 15(4):521-526.
5. Iannotti JP, Spencer EE, Winter U, Deffenbaugh D, Williams G. Prosthetic positioning in total shoulder arthroplasty. J Shoulder Elbow Surg 2005; 14(1 Suppl S):111S-21S.
6. Lazarus MD, Jensen KL, Southworth C, Matsen FA 3rd. The radiographic evaluation of keeled and pegged glenoid component insertion. J Bone Joint Surg Am 2002; 84-A (7):1174-1182.
7. Gillespie R1, Lyons R, Lazarus M. Eccentric reaming in total shoulder arthroplasty: a cadaveric study. Orthopedics 2009; 32(1):21.
8. Hoenecke HR Jr, Hermida JC, Flores-Hernandez C, D'Lima DD. Accuracy of CT-based measurements of glenoid version for total shoulder arthroplasty. J Shoulder Elbow Surg 2010; 19(2):166-171.
9. Biswas D, Bible JE, Bohan M, Simpson AK, Whang PG, Grauer JN. Radiation exposure from musculoskeletal computerized tomographic scans. J Bone Joint Surg Am 2009; 91(8):1882-1889.
10. Throckmorton TW, Zarkadas PC, Sperling JW, Cofield RH. Pegged versus keeled glenoid components in total shoulder arthroplasty. J Shoulder Elbow Surg 2010;

19(5):726-733.
11. Nowak DD, Gardner TR, Bigliani LU, Levine WN, Ahmad, CS. Interobserver and intraobserver reliability of the Walch classification in primary glenohumeral arthritis. J Shoulder Elbow Surg 2010; 19(2):180-183.
12. Nyffeler RW, Jost B, Pfirrmann CW, Gerber C. Measurement of glenoid version: conventional radiographs versus computed tomography scans. J Shoulder Elbow Surg 2003; 12(2):493-496.
13. Raymond AC, McCann PA, Sarangi PP. Magnetic resonance scanning vs axillary radiography in the assessment of glenoid version for osteoarthritis. J Shoulder Elbow Surg 2013; 22(8):1078-1083.
14. Walch G, Boulahia A, Boileau P, Kempf JF. Primary glenohumeral osteoarthritis: clinical and radiographic classification. The Aequalis Group. Acta Orthopaedica Belgica 1998; 64 Suppl 2:46-52.
15. Burgkart R, Glaser C, Hyhlik-Dürr A, Englmeier KH, Reiser M, Eckstein F. Magnetic resonance imaging-based assessment of cartilage loss in severe osteoarthritis: accuracy, precision, and diagnostic value. Arthritis Rheum 2001; 44(9):2072-2077.
16. Fucentese SF, von Roll A, Koch PP, Epari DR, Fuchs B, Schottle PB. The patella morphology in trochlear dysplasia--a comparative MRI study. Knee 2006; 13(2):145-150.
17. Inui H, Sugamoto K, Miyamoto T, Machida A, Hashimoto J, Nobuhara K. Evaluation of three-dimensional glenoid structure using MRI. J Anat 2001; 199(Pt 3):323-328.
18. Friedman RJ, Hawthorne KB, Genez BM. The use of computerized tomography measurement of glenoid version. J Bone Joint Surg Am 1992; 74(7):1032-1037.
19. Walch G, Badet R, Boulahia A, Khoury A. Morphologic study of the glenoid in primary glenohumeral osteoarthritis. J Arthroplasty 1999; 14(6):756-760.
20. Edwards TB, Boulahia A, Kempf JF, Boileau P, Nemoz C, Walch G. The influence of rotator cuff disease on the results of shoulder arthroplasty for primary osteoarthritis: results of a multicenter study. J Bone Joint Surg Am 2002; 84-A (12):2240-2248.
21. Smith-Bindman R, Lipson J, Marcus R, Kim KP, Mahesh M, Gould R, et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch Intern Med 2009; 169(22):2078-2086.
22. Scalise JJ, Codsi MJ, Brems JJ, Iannotti JP. Inter-rater reliability of an arthritic glenoid morphology classification system. J Shoulder Elbow Surg 2008; 17(4):575-577.
23. Budge MD, Lewis GS, Schaefer E, Coquia S, Flemming DJ, Armstrong AD. Comparison of standard two-dimensional and three dimensional corrected glenoid version measurements. J Shoulder Elbow Surg 2011; 20(4):577-83.
24. Kwon YW, Powell KA, Yum JK, Brems JJ, Iannotti JP. Use of three-dimensional computed tomography for the analysis of the glenoid anatomy. J Shoulder Elbow Surg 2005; 14(1):85-90.
25. Walch G, Vezeridis PS, Boileau P, Deransart P, Chaoui J. Three-dimensional planning and use of patient-specific guides improve glenoid component position: an in vitro study. J Shoulder Elbow Surg 2015; 24(2):302-309.
26. Bokor DJ, O'Sullivan MD, Hazan GJ. Variability of measurement of glenoid version on computed tomography scan. J Shoulder Elbow Surg 1999; 8(6):595-598.