American Journal of Sports Medicine

http://ajs.sagepub.com

Interobserver Agreement in the Classification of Rotator Cuff Tears

Multicenter Orthopaedic Outcomes Network-Shoulder (MOON Shoulder Group), John E. Kuhn, Warren R. Dunn, Benjamin Ma, Rick W. Wright, Grant Jones, Edwin E. Spencer, Brian Wolf, Marc Safran, Kurt P. Spindler, Eric McCarty, Brian Kelly and Brian Holloway

Am. J. Sports Med. 2007; 35; 437 originally published online Jan 31, 2007;
DOI: 10.1177/0363546506298108

Interobserver Agreement in the Classification of Rotator Cuff Tears

Multicenter Orthopaedic Outcomes Network-Shoulder (MOON Shoulder Group):
John E. Kuhn,*[†] MD, Warren R. Dunn,[†] MD, MPH, Benjamin Ma,[‡] MD, Rick W. Wright,[§] MD, Grant Jones,[||] MD, Edwin E. Spencer,[¶] MD, Brian Wolf,[#] MD, Marc Safran,[‡] MD, Kurt P. Spindler,[†] MD, Eric McCarty,** MD, Brian Kelly,[††] MD, and Brian Holloway,[¶] MD
*From the [†]Division of Sports Medicine and Shoulder Surgery, Department of Orthopaedics and Rehabilitation, Vanderbilt University Medical Center, Nashville, Tennessee, [‡]Sports Medicine Program, Department of Orthopaedic Surgery, University of California, San Francisco, California, [§]Sports Medicine Service, Department of Orthopaedic Surgery, Washington University, St. Louis, Missouri, [||]Division of Sports Medicine, Department of Orthopaedic Surgery, The Ohio State University, Columbus, Ohio, [¶]Shoulder and Elbow Service, Knoxville Orthopaedic Clinic, Knoxville, Tennessee, [#]Sports Medicine Center, Department of Orthopaedics and Rehabilitation, University of Iowa, Iowa City, Iowa, **Division of Sports Medicine, Department of Orthopaedics, University of Colorado, Boulder, Colorado, and the [††]Sports Medicine and Shoulder Service, Hospital for Special Surgery, New York, New York*

Background: Six classification systems have been proposed for describing rotator cuff tears designed to help understand their natural history and make treatment decisions.

Purpose: To assess the interobserver variation for these classification systems and identify the method with the best interobserver agreement.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: Six rotator cuff tear classification systems were identified in a literature search. The components of these systems included partial-thickness rotator cuff tears and classification by size, shape, configuration, number of tendons involved, and by extent, topography, and nature of the biceps. Twelve fellowship-trained orthopaedic surgeons who each perform at least 30 rotator cuff repairs per year reviewed arthroscopy videos from 30 patients with a random assortment of rotator cuff tears and classified them by the 6 classification systems. Interobserver variation was determined by a kappa analysis.

Results: Interobserver agreement was high when distinguishing between full-thickness and partial-thickness tears (0.95, $\kappa = 0.85$). The investigators agreed on the side (articular vs bursal) of involvement for partial-thickness tears (observed agreement 0.92, $\kappa = 0.85$) but could not agree when classifying the depth of the partial-thickness tear (observed agreement 0.49, $\kappa = 0.19$). The best agreement for full-thickness tears was seen when the tear was classified by topography (degree of retraction) in the frontal plane (observed agreement 0.70, $\kappa = 0.54$).

Conclusion: With the exception of distinguishing partial-thickness from full-thickness rotator cuff tears and identifying the side (articular vs bursal) of involvement with partial-thickness tears, currently described rotator cuff classification systems have little interobserver agreement among experienced shoulder surgeons. Researchers should consider describing full-thickness rotator cuff tears by topography (degree of retraction) in the frontal plane.

Keywords: rotator cuff; classification; interrater reliability; agreement

To understand the natural history of rotator cuff disease and its treatment, it is important to find a reliable method of classifying or describing rotator cuff tears. A number of authors have proposed systems to classify rotator cuff tears, yet little work has been done to determine which systems or components of systems are most reliable. The purpose of this study was to assess the interobserver agreement for a

*Address correspondence to John E. Kuhn, MD, Vanderbilt Sports Medicine, Suite 4200, Medical Center East, South Tower, 1215 21st Avenue South, Nashville, TN 37232 (e-mail: j.kuhn@vanderbilt.edu).

No potential conflict of interest declared.

The American Journal of Sports Medicine, Vol. 35, No. 3
DOI: 10.1177/0363546506298108

variety of classification systems to describe a series of rotator cuff tears. The hypothesis driving this investigation was that some methods of describing rotator cuff tears will have better interobserver agreement than others.

METHODS

A literature search using key words “rotator cuff” and “classification” was performed to identify recommended systems for classifying rotator cuff tears. Of 162 potential manuscripts, 12 were reviewed to identify proposed classification systems. In addition, textbook chapters were reviewed. This search identified 9 rotator cuff tear classification systems that have been recommended to describe partial- and full-thickness rotator cuff tears.[1,3,4,5,6,8,9,10] Of these, 3 were excluded[1,5,9] as they were very similar to the others. Bayne and Bateman[1] described rotator cuff tears by size and divided them into 4 groups, essentially describing tears like DeOrio and Cofield.[3] Habermeyer and Lehmann[5] described tears in the sagittal plane in a fashion similar to Patte.[8] Snyder[9] proposed a thorough and complex system that included many of the elements of the systems used in our study and did not add new information. To that end, we elected to study 6 different classification systems that have been proposed to classify rotator cuff tears in 5 references: DeOrio (1984),[3] Ellman (1995),[4] Harryman (1991),[6] Patte (1990),[8] and Wolfgang (1974).[10]

Arthroscopic videos of 30 rotator cuff tears from 30 patients, including partial tears and full-thickness tears of varying size, were collected and used for this study. Each video began with an intra-articular view using a 30° arthroscope placed in a standard posterior portal. A probe with 2-mm laser-etched lines was introduced through an anterior portal through the rotator interval. The probe was used to measure the anterior-posterior dimensions of the tear, and the depth of the partial-thickness tear by placing the probe into the tear until resistance was met with the intent to measure depth, and along the exposed greater tuberosity footprint. The subacromial video was taken after the burscectomy had been performed, with the arthroscope in the subacromial space introduced through a posterior portal. The probe was introduced into a direct lateral portal, and the medial-lateral dimension of the tear was measured. The arthroscope was then introduced into the subacromial space through a lateral portal, and the probe was placed in the posterior portal to measure the anterior-posterior dimension of the tear. Videos were made by 4 of the authors who had been instructed in the above technique before collecting videos. Two additional videos were obtained but did not follow this technique and were discarded. Videos were not distributed for review until 4 months after their collection. No patient identifiers were visible on the videos.

Twelve fellowship-trained shoulder surgeons who each repair more than 30 rotator cuffs each year and perform arthroscopic shoulder surgery regularly reviewed the videos. The investigators could rewind and replay and use the stop function as needed to review each video. The investigators then classified each of the 30 tears according to a form that was developed for this study (see Appendix). This form included identification and classification of partial-thickness tears with regard to side and depth[4]; full-thickness tears were classified by size,[3] shape,[10] configuration,[4] and the number of tendons involved.[6] The sixth classification system is more complex and was developed by Patte.[8] This system has 5 questions that describe rotator cuff tears by a) tear extent, b) sagittal plane topography (divided into segments which generally reflect the number of tendons involved), c) frontal plane topography (proximal tendon stump close to the bony insertion, tendon stump at the level of the humeral head, or tendon stump at the level of the glenoid), d) nature of the long head of the biceps tendon, and e) trophicity determined by fatty infiltration of the muscle on computed tomography. (See Appendix for detail regarding these classification systems.) For this study, we included descriptors that could be seen during arthroscopy and did not include evaluation of fatty infiltration on computed tomography. The data collection form included a cover page with instructions and the original illustrations from each published classification system to assist the surgeons in classifying the rotator cuff tears in the videos.

TABLE 1
Interobserver Agreement for Partial-Thickness Rotator Cuff Tears[a]

	Full- vs Partial-Thickness Tear	Articular vs Bursal Surface	Depth or Grade of Tear
Agreement	0.95	0.92	0.49
Kappa	0.85	0.85	0.19
Agreement beyond chance	Excellent	Excellent	Poor

[a](From Ellman.[4]) See Appendix for details regarding the classification systems.

After each surgeon reviewed and classified each of the 30 videos, the data from the forms were entered into a spreadsheet. Interobserver agreement was determined using kappa statistical methods. Kappa is a measure of agreement between 2 or more observers that takes into account agreement that could occur by chance. By convention, a kappa value of 0 to 0.19 is considered poor agreement; 0.20 to 0.39 is rated fair; 0.40 to 0.59 is moderate; 0.60 to 0.79 is good; and 0.80 to 1.00 is excellent.[7]

RESULTS

The experienced shoulder surgeons in this group could distinguish between partial- and full-thickness tears (observed agreement 0.95, $\kappa = 0.85$) (Table 1). With regard to partial-thickness tears, excellent interobserver agreement was found with regard to determining the side of involvement (observed agreement 0.92, $\kappa = 0.85$). However, determining the depth of the tear had poor interobserver agreement (observed agreement 0.49, $\kappa = 0.19$).

TABLE 2
Interobserver Agreement for Full-Thickness Rotator Cuff Tears[a]

	Size	Shape	Configuration	Number of Tendons	*Extent*	*Sagittal Topography*	*Frontal Topography*	*Biceps*
Agreement	0.56	0.64	0.40	0.62	0.60	0.61	0.70	0.83
Kappa	0.32	0.31	0.22	0.35	0.34	0.34	0.54	0.13
Agreement beyond chance	Fair	Fair	Fair	Fair	Fair	Fair	Moderate	Poor

[a]Classification systems rated full-thickness rotator cuff tears by size (DeOrio[3]), by shape (Wolfgang[10]), by configuration (Ellman[4]), and by the number of tendons involved (Harryman[6]). The Patte classification[8] (italicized in column headings) determined full-thickness rotator cuff tears by extent, topography in 2 planes, and by the nature of biceps tendon. See Appendix for details regarding the classification.

Figure 1. Grading rotator cuff tears by the topography in the frontal plane. (Reprinted from Patte.[8] Classification of rotator cuff lesion. *Clin Orthop Relat Res*. 1990;254:84 [figure 2].) This is 1 component of the Patte classification. Tears are classified as either stage I: close to the bony insertion, stage II: at the level of the humeral head, or stage III: at the level of the glenoid. This method of describing the size of a full-thickness rotator cuff tear had the best agreement and best kappa.

When evaluating full-thickness tears, interobserver agreement was fair for nearly all of the classification systems currently described in the literature, including the commonly used system of classifying tears by size (small, medium, large, massive) (Table 2). Interestingly, 1 component of the Patte classification system, which describes the tear in the frontal plane (amount of retraction with proximal tendon stump close to bony insertion, tendon stump at the level of the humeral head, or tendon stump at the level of the glenoid) (Figure 1), had moderately good interobserver agreement (0.70, $\kappa = 0.54$).

DISCUSSION

In an evidence-based medicine approach to the treatment of rotator cuff tears, meta-analyses and systematic reviews of level I studies represent the highest level of evidence. This requires the ability to combine the data from different studies of rotator cuff tears. These studies must be consistent and precise in how rotator cuff tears are described. Interobserver agreement is a way of assessing the precision of a classification system. This study was conducted to determine which of these classification systems had the best interobserver agreement to guide research in how best to describe and classify rotator cuff tears.

The results from this study demonstrate that experienced clinicians show very high agreement when deciding that a rotator cuff tear is a partial tear or a full-thickness tear. In addition, clinicians are adept at agreeing on the side involved (articular or bursal) when a partial-thickness tear is identified; however, clinicians could not agree on the depth of the partial-thickness tears.

With regard to full-thickness tears, our results suggest that none of the classification systems resulted in excellent agreement. It is clear that the more complex the system, the less likely the agreement will be; therefore, rotator cuff classification systems should be as simple as possible. With regard to the most commonly used description of rotator cuff tears as small, medium, large, or massive in size, agreement was fair (0.56, $\kappa = 0.32$). The best agreement seemed to be with regard to describing the tear in the frontal plane (0.70, $\kappa = 0.54$), which describes full-thickness rotator cuff tears by the location of the tendon as near the insertion, at the level of the humeral head, and at the level of the glenoid (Figure 1). Classification by size has other drawbacks as well. For example, unless measured, all descriptions are estimates. Also, because classification by size is not normalized to the patient's size, the value of the absolute size of the rotator cuff tear is in question, and Ciepiela and Burkhead[2] suggested normalizing the size of the tear to the size of the humeral head. The Patte frontal plane description is already normalized and may be more accurate in predicting functional deficits or reparability than using size as a descriptive criterion.

This study is not without limitations. First, our literature search to identify rotator cuff tear classifications was confined to the English language, which may have limited the scope of this study. Second, some classification systems employ features such as chronicity or normalizing the tear size to the diameter of the humeral head,[2] and we could not evaluate this information with this study design, which rated tears based on their appearance on arthroscopic video. Video inspection of the rotator cuff tears does not allow for tactile information or motion of the arm, which may be helpful in determining items such as the depth of the partial-thickness tear. It is conceivable that agreement would be improved if each reviewer could perform the arthroscopy and manipulate the probe on each patient. The reviewers of the videos were fellowship-trained, experienced shoulder surgeons. As such, our results apply only when this population is reviewing arthroscopic videos of rotator cuff tears. It is important to note that this study was not designed to identify the most accurate method of describing rotator cuff tears. This study design was only able to assess agreement between physicians. The best method would have high interobserver agreement and accuracy.

Despite these limitations, this study clearly demonstrates that experienced, fellowship-trained shoulder surgeons are able to distinguish between partial tears and full-thickness tears reliably, and they can determine the side on which the partial tear exists. In addition, experienced shoulder surgeons can agree reasonably well when describing full-thickness rotator cuff tears by the degree of retraction in the frontal plane as proposed by Patte.[8] Other classification systems or distinctions that are more complex, offering more resolution with regard to tear size, extent of the tear, or shape of the tear, have surprisingly poor reliability. This is the only study to date to evaluate the interobserver agreement of a variety of rotator cuff tear classification systems.

As the field of shoulder surgery moves toward accepting tenets of evidence-based medicine, it becomes imperative that researchers use methods of characterizing patient outcomes and classifying disease states that have undergone validation testing. On the basis of this study, rotator cuff tears should be classified as to whether they are partial-thickness or full-thickness. Partial-thickness tears can be described by the side involved. The depth of partial-thickness tears is difficult to estimate with agreement.

Categorizing the full-thickness tear size as (a) proximal tendon stump close to the bony insertion, (b) tendon stump at the level of the humeral head, or (c) tendon stump at the level of the glenoid produces adequate interobserver reliability. Other currently existing descriptive terms for rotator cuff tears, most of which were developed before the advent of shoulder arthroscopy, have little agreement and, therefore, may be of little value.

REFERENCES

1. Bayne O, Bateman J. Long-term results of surgical repair of full-thickness rotator cuff tears. In: *Surgery of the Shoulder*. Philadelphia, Pa: CV Mosby; 1984:167-171.
2. Ciepiela MD, Burkhead WZ. Classification of rotator cuff tears. In: *Rotator Cuff Disorders*. Philadelphia, Pa: Williams & Wilkins; 1996:100-102.
3. DeOrio JK, Cofield RH. Results of a second attempt at surgical repair of a failed initial rotator cuff repair. *J Bone Joint Surg Am.* 1984;66:563-567.
4. Ellman H. Rotator cuff disorders. In: *Arthroscopic Shoulder Surgery and Related Disorders*. Philadelphia, Pa: Lea and Febiger; 1993:98-119.
5. Habermeyer P, Lehmann L. Rotatorenmanschette, rotatorenintervall und lange bizepssehne. In: *Schulterchirurgie*. München, Jena: Urban & Fischer; 2002:333-374.
6. Harryman DT, Mack LA, Wang K, Jackins SE, Richardson ML, Matsen FA. Repairs of the rotator cuff. Correlation of functional results with integrity of the cuff. *J Bone Joint Surg Am*. 1991;73:982-989.
7. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
8. Patte D. Classification of rotator cuff lesions. *Clin Orthop Relat Res*. 1990;254:81-86.
9. Snyder SJ. Arthroscopic classification of rotator cuff lesions and surgical decision making. In: *Shoulder Arthroscopy*. Philadelphia, Pa: Lippincott Williams & Wilkins; 2003:201-207.
10. Wolfgang GL. Surgical repairs of tears of the rotator cuff of the shoulder. *J Bone Joint Surg Am.* 1974;56:14-26.

APPENDIX
Questionnaire Used by Surgeons Evaluating Each Rotator Cuff Video[a]

Shoulder Moon Group Interobserver Reliability of Rotator Cuff Tear Classification Systems

REVIEWER NAME____ DATE____________________________
VIDEO NUMBER_______
IS THIS TEAR:

_____FULL-THICKNESS? (SKIP THE REST OF THIS PAGE, GO TO PAGE 2)
OR
_____PARTIAL-THICKNESS? (SEE BELOW)

If the tear is partial-thickness, classify by the system below and stop.
_____Articular surface partial tear
_____Bursal surface partial tear
_____Grade 1 (<3 mm deep)
_____Grade 2 (3-6 mm deep, or approximately 50% of the thickness of tendon)
_____Grade 3 (>6 mm deep, or more than half of the thickness of the tendon)

STOP
FULL-THICKNESS TEARS ONLY

1) Classify the SIZE of the tear by the widest diameter:
______Small (<1 cm)
______Medium (1-3 cm)
______Large (3-5 cm)
______Massive (>5 cm)

2) Classify the SHAPE of the tear by:
_____Transverse
_____Triangular or crescentic
_____Massive

3) Classify the CONFIGURATION of the tear by:
______Transverse Linear (Supraspinatus at insertion, no retraction)
______Crescent (A transverse tear deformed by the pull of infraspinatus and subscapularis)
______L-shaped (Extension between infraspinatus and supraspinatus)
______Reverse L (Extension into rotator interval)
______Trapezoidal (Insertions of infraspinatus and supraspinatus)
______Massive (Extension into teres or anterior subscapularis)

4) Classify the tear by the NUMBER OF TENDONS INVOLVED:
______Stage 1A (Partial tear)
______Stage 1B (Full-thickness tears isolated to supraspinatus)
______Stage II (Supraspinatus and at least a portion of infraspinatus)
______Stage III (Supraspinatus, infraspinatus, and subscapularis)
______Stage IV (Rotator cuff arthropathy)

5) The next assessments are part of the Patte Classification (TOPOGRAPHY AND TROPHICITY):

a) Patte EXTENT of tear:
______Group I: Partial or full-thickness <1 cm sagittal, or deep partial, or superficial, or small full-thickness tear
______Group II: Full-thickness entire supraspinatus involved
______Group III: Full-thickness more than 1 tendon involved
______Group IV: Massive tear with osteoarthritis

b) Patte TOPOGRAPHY IN SAGITTAL PLANE as divided into 6 segments (See Figure 1):
Segment 1 (Subscapularis)
Segment 2 (Coracohumeral ligament)
Segment 3 (Supraspinatus only)
Segment 4 (Supraspinatus and upper half of infraspinatus)
Segment 5 (Supraspinatus and entire infraspinatus)
Segment 6 (Subscapularis, supraspinatus, and infraspinatus)

RATE THIS TEAR AS ONE OF THE FOLLOWING CATEGORIES:
_____Anterosuperior tears (Segments 1, 2, 3)
_____Superior tears (Segments 2, 3)
_____Posterosuperior tears (Segments 4, 5)
_____Total cuff tears (Segment 6)

c) Patte TOPOGRAPHY IN FRONTAL PLANE:
_____Stage 1 (Proximal tendon stump close to bony insertion)
_____Stage 2 (Tendon stump at level of humeral head)
_____Stage 3 (Tendon stump at level of glenoid)

d) Patte long head of biceps
_____Intact
_____Torn
_____Dislocated
_____For anterosuperior tear, no functional adversity
_____For Group III, may have superior migration of humeral head

[a]This questionnaire included a cover page with illustrations from the original manuscript to assist reviewers in making their assessments.