Interobserver Agreement in the Classification of Partial-Thickness Rotator Cuff Tears Using the Snyder Classification System

Christopher S. Lee,*[†] MD, MBA, Shane M. Davis,[†] BA, Brittany Doremus,[†] DO, Shalen Kouk,[†] MD, and William B. Stetson,[†] MD
Investigation performed at Stetson Powell Orthopedics and Sports Medicine, Burbank, California, USA

Background: At present, there is no widely accepted classification system for partial-thickness rotator cuff tears, and as a result, optimal treatment remains controversial.

Purpose: To examine the interobserver reliability and accuracy of classifying partial rotator cuff tears using the Snyder classification system. We hypothesized that the Snyder classification would be reproducible with high reliability and accuracy.

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

Methods: Twenty-seven orthopaedic surgeons reviewed 10 video-recorded shoulder arthroscopies. Each surgeon was provided with a description of the Snyder classification system for partial-thickness rotator cuff tears and was then instructed to use this system to describe each tear. Interrater kappa statistics and percentage agreement between observers were calculated to measure the level of agreement. Surgeon experience as well as fellowship training was evaluated to determine possible correlations.

Results: A kappa coefficient of 0.512 indicated moderate reliability between surgeons using the Snyder classification to describe partial-thickness rotator cuff tears. The mean correct score was 80%, which indicated "very good" agreement. There was no correlation between the number of shoulder arthroscopies performed per year and fellowship training and the number of correct scores.

Conclusion: The Snyder classification system is reproducible and can be used in future research studies in analyzing the treatment options of partial rotator cuff tears.

Keywords: Snyder classification; partial rotator cuff tears; articular surface; bursal surface; arthroscopy; supraspinatus

Partial-thickness rotator cuff tears may involve the articular surface, bursal surface, or both sides of the rotator cuff. They may be asymptomatic or a potential source of shoulder dysfunction. With the advent of magnetic resonance imaging (MRI) and shoulder arthroscopy, more tears are being diagnosed. Despite this increased awareness of partial tears, the optimal treatment of partial-thickness rotator cuff tears remains controversial.[11,18,19]

At present, there is no widely accepted classification system for partial-thickness rotator cuff tears.[8,9,17-19] A recent study of 12 fellowship-trained orthopaedic surgeons using the Ellman classification system of greater than or less than 50% tearing of partial-thickness rotator cuff tears showed low agreement when classifying the depth of these partial tears.[9] With no reliable classification system, it is difficult to communicate and determine which treatment options suit which type of tear.[11,19]

The Snyder classification system provides a comprehensive classification system for partial-thickness rotator cuff tears both on the articular and bursal side not provided by other systems.[16] Tears are graded on the degree of tearing on both the articular side and the bursal side. The degree of tearing is graded from 0 to IV, with 0 being normal and IV being a significant partial tear more than 3 cm in size (Table 1). The articular side of the tear is examined and then debrided of any loose collagenous fibers to determine whether it is a partial versus a complete tear. Once a partial

TABLE 1
Snyder Classification of Partial Rotator Cuff Tears[13,14]

Location of tear	
A	Articular surface
B	Bursal surface
Severity of tear	
0	Normal cuff with smooth coverings of synovium and bursa
I	Minimal superficial bursal or synovial irritation or slight capsular fraying in a small, localized area; usually <1 cm
II	Actually fraying and failure of some rotator cuff fibers in addition to synovial, bursal, or capsular injury; usually 1-2 cm
III	More severe rotator cuff injury, including fraying and fragmentation of tendon fibers, often involving the entire surface of a cuff tendon (most often the supraspinatus); usually 2-3 cm
IV	Very severe partial rotator tear that usually contains a sizable flap tear in addition to fraying and fragmentation of tendon tissue and often encompasses more than a single tendon; usually >4 cm

articular-sided tear is verified, it is graded using the Snyder system. A marker suture technique is then used, as described by Snyder, so that the bursal side of the tear can be easily located during the bursal examination.[15] The subacromial space is then entered into, bursal tissue is removed using a full radius shaver, and the marker suture is then localized. It is often necessary to rotate the arm both internally and externally to locate the marker suture. The bursal side of the cuff is then inspected and graded in a similar fashion. A normal rotator cuff is classified as an A-0, B-0, with no evidence of tearing or fraying involving either the articular or bursal side of the rotator cuff. Figures 1 and 2 show A-III, B-I and A-II, B-III partial tears, respectively.

Video-recorded surgeries are a useful tool for exploring interrater reliability and accuracy because they provide exact and reproducible information for a certain pathology or surgical procedure.[17] We present a study using the video-recorded arthroscopies of 10 patients and their intraoperative findings to assess the interobserver reliability and accuracy of the Snyder classification systems for partial rotator cuff tears. We hypothesized that the Snyder classification would be reproducible with high reliability and accuracy. We also believed that a widely used classification system should be reproducible despite surgeon experience and level of training. The purpose of this study was to determine whether the Snyder classification system was an effective tool for classifying partial rotator cuff tears and to determine its interobserver reliability.

METHODS

A prospective, nonrandomized study was conducted at the American Academy of Orthopaedic Surgeons Annual Meeting in Washington, DC, in February 2005 and the following year in Chicago, Illinois, in February 2006. Twenty-seven
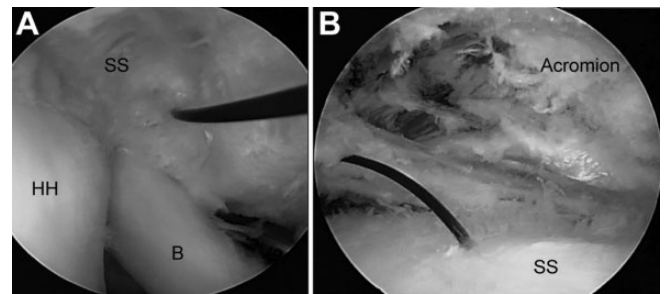

Figure 1. A-III, B-I partial-thickness rotator cuff tear. (A) The articular side of a partial rotator cuff tear viewing from the posterior portal with the marker suture through an A-III type tear. (B) The corresponding bursal side of the tear viewing from the posterior with minimal fraying (<1 cm) corresponding to a B-I tear. B, biceps; HH, humeral head; SS, supraspinatus.
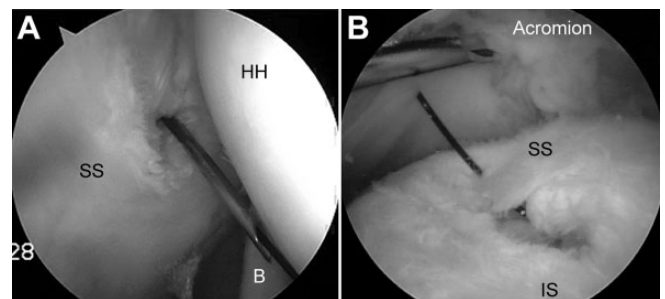

Figure 2. A-II, B-III partial-thickness rotator cuff tear. (A) The articular side of the supraspinatus tendon tear (A-II) viewed from the posterior portal with the marker suture being placed via a spinal needle. (B) In the subacromial space viewing from the posterior portal with the shaver lateral, the bursal side showing a significant partial tear (B-III). B, biceps; HH, humeral head; IS, infraspinatus; SS, supraspinatus.

orthopaedic surgeons volunteered for the study. The mean age of participants was 48 years (range, 36-68 years). The number of years in practice after residency or fellowship training averaged 16 years (range, 1-43 years). The number of shoulder arthroscopies performed per year averaged 124 (range, 0-350). Of the 27 participants, 16 were fellowship-trained in sports medicine and arthroscopy. Each surgeon reviewed a DVD that included 10 separate shoulder arthroscopies. Each individual shoulder arthroscopy video showed a view of the articular and bursal sides of the rotator cuff only. To provide consistent video reproduction for the reviewing surgeons, all arthroscopic procedures were performed by the senior author (W.B.S.), who is fellowship-trained in shoulder arthroscopy. Tear size was measured intraoperatively, and classifications were assigned by the senior author at the time of surgery.

Creation of Video Recordings

Videos were created during the arthroscopic evaluation of patients with partial-thickness rotator cuff tears. The

TABLE 2
Interpretive Scale for Interrater
Percentage Agreement Values

Value of Agreement	Strength of Agreement
0%	None
1%-20%	Very poor
21%-40%	Poor
41%-60%	Moderate
61%-80%	Good
81%-99%	Very good
100%	Perfect

TABLE 3
Kappa Coefficient Agreements[2]

Kappa Value	Interpretation
>0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost perfect

procedures were recorded without audio. Arthroscopies were performed in the lateral decubitus position with the operative arm in 10 lb of balanced suspension in the scapular plane. The glenohumeral examination showed only the articular side of the supraspinatus tendon viewing from the posterior and no other structures of the glenohumeral joint. If a partial articular-sided tear was noted, it was debrided and a marker suture was placed via a spinal needle to mark the articular side of the rotator cuff tear, which made for easy identification of the bursal side of the tear. After the glenohumeral examination, a bursoscopy examination was then performed viewing from the posterior portal. The bursal side of the rotator cuff was then video-recorded after removal of the bursal tissue, and if a marker suture had been placed, this was located in the subacromial space. Each video clip was approximately 15 to 20 seconds.

No clinical data were provided other than the intraoperative video recordings. The participant could rewind and replay and use the stop function as needed to review each video segment.

Collection of Classification Data

Each surgeon was provided a description of the Snyder classification of partial-thickness rotator cuff tears, which grades tears according to their location (A for articular, B for bursal) and size (Table 1).[10] Surgeons were able to review the cases as many times as necessary to provide a definitive description of the tear. Each participant was provided with a multiple choice questionnaire consisting of 4 possible choices for each case study. After each surgeon reviewed and classified each of the 10 cases, the data were recorded on an Excel spreadsheet (Microsoft Corp).

Statistical Analysis

Classifications were deemed to be correct if they matched those by the senior author determined intraoperatively. Interobserver reliability was calculated using interrater kappa statistics to measure agreement among the surgeons using the Snyder classification system for grading partial rotator cuff tears. We also used a percentage agreement between observers as another statistical representation of reliability, as described by Sasyniuk et al,[14] and used the associated interpretive scale (Table 2). No intraobserver reliability test was performed.

Pearson correlation coefficients were used to determine a correlation between age, years in practice, number of shoulder arthroscopies performed per year, and fellowship training with the number of correct scores. This correlation along with the percentage agreement and interrater agreement and all statistical analyses were performed by Minimax Consulting.

RESULTS

The mean correct score was 80% (range correct, 4-10). The mean number of correct responses for each case averaged 23 of 27 (85%; range, 19-25). Using the interpretive scale for interrater percentage agreement (Table 2), the strength of agreement was "very good."

The multirater kappa coefficient analysis showed a kappa of 0.512, with a standard error of 0.0104. This was calculated using a normal distribution ($Z = 49.5$) and an error probability ($P < .0001$). The resulting kappa coefficient of 0.512 is strongly statistically significant. Using the guidelines for kappa coefficient agreements,[9,13] the level of interobserver agreement among the 27 orthopaedic surgeon study participants is classified as "moderate" (Table 3).

Using Pearson correlation coefficients that were computed for all variables of age, score, years in practice, number of shoulder arthroscopies performed per year, and fellowship training, older age ($r = -0.419$) and those with more years in practice ($r = -0.428$) tended to have lower scores. The number of shoulder arthroscopies performed per year and fellowship training were not significantly correlated with an increased number of correct scores.

DISCUSSION

The purpose of this study was to determine whether the Snyder classification system was an effective tool for classifying partial rotator cuff tears and to determine its interobserver reliability. This prospective, nonrandomized study among a group of orthopaedic surgeons with a wide range of experience was able to validate its usefulness with interrater reliability. The level of agreement among the 27 orthopaedic surgeons participating in the study was classified as "moderate" based on the kappa coefficient. The kappa statistic, though widely used for the purpose of reliability, is felt by some authors and statisticians not to be an accurate tool as

the primary method of quantifying agreement, particularly with multiple raters and categories.[1,6] We therefore also used a percentage agreement, which was 85%, between observers as another statistical representation of reliability, resulting in a "very good" strength of agreement.[14]

The most widely used technique for comparing the ratings of any panel of observers all providing opinions on a series of events is the Cohen kappa coefficient.[3] There are 2 measures that contribute to the kappa statistic: expected agreement and observed agreement. Expected agreement is the probability that 2 surgeons provide the same response to a question for any given patient (chance agreement). Observed agreement is the probability that 2 surgeons provide the same response to a question for a specific patient. The kappa statistic is the amount of observed agreement that is beyond the agreement expected due to chance alone. A kappa coefficient of 1.00 indicates that there is perfect agreement, whereas a kappa coefficient of 0.00 implies no more agreement than would be expected by chance alone. The kappa coefficients were interpreted according to the guidelines described by Landis and Koch[10]: values of <0.00 indicate poor agreement; 0.00 to 0.20, slight agreement; 0.21 to 0.40, fair agreement; 0.41 to 0.60, moderate agreement; 0.61 to 0.80, substantial agreement; and >0.80, almost perfect agreement (Table 3).

We implemented the multirater version of the kappa coefficient using the algorithm developed in detail by Fleiss.[7] The development of this multirater kappa was verified by randomly selecting just 2 of the orthopaedic surgeons and performing the kappa analysis in both a standard and customized version. The resulting kappa coefficient of 0.512 and all associated error terms agreed to the third significant digit with an error of probability of $P < .0001$ and a standard of error 0.0104, making the kappa coefficient strongly statistically significant

Classifying and treating partial rotator cuff tears is still a matter of controversy as there is still no accepted system for classifying these tears and hence no accepted treatment algorithm. Treating these tears is controversial, as some surgeons recommend debriding partial tears while others recommend completing the tear and treating it like a full-thickness tear. Other surgeons recommend an in situ–type repair as described by Snyder.[15] What tears to debride versus what tears to be more aggressive with depends on the extent of the tear and whether it involves the articular, bursal, or both sides of the rotator cuff. Other factors also need to be taken into consideration, including the age of the patient, their occupation, history, physical examination findings, and also the etiology or pathogenesis of these tears.

With no accepted classification system, it is difficult to compare studies of partial rotator cuff tears and to come to some sort of conclusion as to the best way to treat them. Neer[12] first described 3 stages of rotator cuff disease looking at histological specimens of the rotator cuff. Stage I is characterized by hemorrhage and cuff edema, stage II by cuff fibrosis, and stage III by cuff tear.

TABLE 4
Participant Data

Doctor	Score	Age, y	Years Practicing, n	Scopes per Year, n	Fellowship-Trained?
1	10	50	20	120	Yes
2	4	68	43	0	No
3	10	49	15	350	Yes
4	9	40	3	150	Yes
5	10	52	24	175	No
6	6	56	25	150	Yes
7	10	47	15	55	Yes
8	6	38	7	150	Yes
9	8	50	19	35	Yes
10	9	44	13	300	Yes
11	8	51	20	200	No
12	10	55	10	100	Yes
13	5	52	19	100	Yes
14	10	58	35	0	No
15	7	52	21	12	No
16	10	34	5	300	Yes
17	8	43	10	200	Yes
18	9	60	25	40	No
19	9	34	0	0	No
20	9	48	16	80	No
21	10	43	10	200	Yes
22	4	61	30	0	No
23	8	42	4	120	Yes
24	10	40	12	180	No
25	10	60	26	30	No
26	10	36	5	150	Yes
27	10	37	3.5	150	Yes

However, this system has significant limitations clinically and intraoperatively and does not address partial-thickness tears.

Ellman[5] recognized the difficulty in using the Neer classification system and proposed a classification for partial-thickness rotator cuff tears. For partial-thickness tears, grade I is less than 3 mm deep, grade II is 3 to 6 mm deep, and grade III tears are more than 6 mm deep. He also recognized that partial tears could occur on the articular side, bursal side, or be interstitial. He felt that grade III tears involving more than 50% of the tendon should be repaired (assuming mean cuff thickness of 9-12 mm in size).[5] However, the Ellman classification system has never been validated and does not account for the varying sizes of the rotator cuff footprint. In a 2007 multicenter study (Multi-Center Orthopaedic Outcome Network [MOON] Shoulder Group) that consisted of 12 fellowship-trained orthopaedic surgeons who each perform at least 30 rotator cuff repairs per year, there was poor interobserver agreement using the Ellman classification system.[9] They could not agree when classifying the depth of the partial-thickness tear, with an observed agreement of only 0.49 and a kappa coefficient of only 0.19 (slight agreement).[9]

Curtis et al[4] and Ruotolo et al[13] have recommended a similar system, grading tears based on whether the tear size exceeded 50% of the rotator cuff thickness. If one assumes that the average rotator cuff is approximately 12 mm in size, it is possible to grade the percentage of tearing.

Using the supraspinatus footprint as a guide, if more than 6 mm of the footprint is exposed, a greater than 50% tear of the supraspinatus insertion has occurred.[4,13] However, studies have shown that the thickness of the rotator cuff tendon, in particular, the supraspinatus tendon, can vary widely in size from 9 to 21 mm in thickness.[4] This classification system also has never been validated in a study. Again, the greater than or less than 50% rule is arbitrary and does not address the difference between partial articular-sided and partial bursal-sided tears like the Snyder classification does.

One of the drawbacks of our study was that data collection was performed in a multiple choice fashion instead of allowing the surgeons to free-text their descriptions. This method certainly does not reproduce the clinical scenario; however, we felt that a multiple choice format would reduce transcription errors and would not prevent us from assessing the reproducibility and interobserver reliability of the Snyder classification. In addition, the Snyder classification is somewhat complicated for the first-time user so we did not want to overwhelm the participants by using a free-text study design. Interestingly, the study by the MOON Shoulder Group in 2007 also used a multiple choice format when evaluating the Ellman classification system and found very poor interobserver reliability even using this multiple choice format.[9]

This study also lacks an estimation of intrarater reliability, as each surgeon was involved in only 1 trial. This is an aspect of the study that must be addressed in the future. The variability in training and experience could have also been considered a limitation, but it was found to have little effect on the individual scores (Table 4). One of the other weaknesses of our study is that video inspection of partial rotator cuff tears does not allow for palpation of the tears with a probe or shaver, which may be helpful in determining the depth and size of a partial-thickness tear.

CONCLUSION

This study demonstrates that the Snyder classification system can be used to categorize partial-thickness rotator cuff tears in a moderately reliable fashion by both experienced and inexperienced surgeons regardless of their fellowship training. More information is needed and a more comprehensive study would further evaluate the effectiveness of the Snyder classification system; however, we believe this is the first step in an evidence-based approach to managing partial-thickness rotator cuff tears, for it is imperative that we have a method of characterizing these tears as we propose and trial treatment strategies for them.

REFERENCES

1. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol*. 1990;43:551-558.
2. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37-46.
3. Cohen J. This week's Citation Classic. A coefficient of agreement for nominal scales. *Curr Contents*. 1986;3:18.
4. Curtis AS, Burbank KM, Tierney JJ, Scheller AD, Curran AR. The insertional footprint of the rotator cuff: an anatomic study. *Arthroscopy*. 2006;22:609.e1.
5. Ellman H. Diagnosis and treatment of incomplete rotator cuff tears. *Clin Orthop Relat Res*. 1990;254:64-74.
6. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol*. 1990;43:543-549.
7. Fleiss JL. The measurement of inter-rater agreement. In: Fleiss JL, Levin B, Paik MC, eds. *Statistical Methods for Rates and Proportions*. 3rd ed. Hoboken, NJ: Wiley; 2003:598-626.
8. Habermeyer P, Krieter C, Tang KL, Lichtenberg S, Magosch P. A new arthroscopic classification of articular-sided supraspinatus footprint lesions: a prospective comparison with Snyder's and Ellman's classification. *J Shoulder Elbow Surg*. 2008;17:909-913.
9. Kuhn JE, Dunn WR, Ma B, et al. Interobserver agreement in the classification of rotator cuff tears. *Am J Sports Med*. 2007;35:437-441.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
11. McConville OR, Iannotti JP. Partial-thickness tears of the rotator cuff: evaluation and management. *J Am Acad Orthop Surg*. 1999;7:32-43.
12. Neer CS 2nd. Impingement lesions. *Clin Orthop Relat Res*. 1983;173:70-77.
13. Ruotolo C, Fow JE, Nottage WM. The supraspinatus footprint: an anatomic study of the supraspinatus insertion. *Arthroscopy*. 2004;20:246-249.
14. Sasyniuk TM, Mohtadi NG, Hollinshead RM, Russell ML, Fick GH. The inter-rater reliability of shoulder arthroscopy. *Arthroscopy*. 2007;23:971-977.
15. Snyder SJ. *Shoulder Arthroscopy*. Philadelphia, PA: Lippincott William & Wilkins; 2003.
16. Snyder SJ, Pachelli AF, Del Pizzo W, Friedman MJ, Ferkel RD, Pattee G. Partial thickness rotator cuff tears: results of arthroscopic treatment. *Arthroscopy*. 1991;7:1-7.
17. Spencer EE Jr, Dunn WR, Wright RW, et al. Interobserver agreement in the classification of rotator cuff tears using magnetic resonance imaging. *Am J Sports Med*. 2008;36:99-103.
18. Stetson WB, Ryu RKN. Evaluation and arthroscopic treatment of partial rotator cuff tears. *Sports Med Arthrosc Rev*. 2004;12:114-126.
19. Stetson WB, Ryu RKN, Bittar ES. Arthroscopic treatment of partial rotator cuff tears. *Oper Techn Sport Med*. 2004;12:135-148.