Reproducibility and Reliability of the Snyder Classification of Superior Labral Anterior Posterior Lesions Among Shoulder Surgeons

Xiaofeng Jia,* MD, PhD, Atsushi Yokota,† MD, PhD, Eric C. McCarty,‡ MD, Gregory P. Nicholson,§ MD, Stephen C. Weber,‖ MD, Patrick J. McMahon,¶ MD, Warren Reid Dunn,# MD, MPH, and Edward G. McFarland,**†† MD
Investigation performed at The Johns Hopkins University, Baltimore, Maryland

Background: Previous studies of the reliability of classifications for superior labral anterior and posterior (SLAP) lesions suggest that intraobserver/interobserver reliability is poor.

Purpose: The goals were to (1) evaluate intraobserver/interobserver reliability of the Snyder classification of labrum tears among experienced surgeons, (2) determine the effect of simplifying that classification into normal versus abnormal labrums, (3) determine the reliability of subdividing type II SLAP lesions, and (4) evaluate the effect of videotape quality on diagnostic confidence.

Study Design: Cohort study (diagnosis); Level of evidence, 3.

Methods: At 2 separate times, 5 experienced shoulder surgeons (all >10 years of practice) evaluated 90 videotapes of shoulder arthroscopy performed for SLAP lesions to identify lesion type, to grade video quality, and to determine degree of confidence in diagnosis. The results were compared for intraobserver reliability (κ analysis) and interobserver reliability (intraclass correlation coefficients) and effect of video quality on diagnostic confidence. Significance was set at $P < .05$.

Results: When the choices were normal labrum or the 4 types of SLAP lesions, the intraobserver agreements (mean κ, 0.670) and interobserver agreements (mean correlation, 0.804) were substantial. When the labrums were divided into normal (normal and type I) and abnormal (types II-IV), the intraobserver agreements (mean κ, 0.792) and interobserver agreements (mean correlation, 0.648) were substantial. When the 3 type II SLAP lesion subvariants were evaluated, the intraobserver agreement (mean κ, 0.598) was moderate, and the interobserver agreement (mean correlation, 0.804) was substantial. The confidence of the diagnosis was highly correlated with the perceived video quality (average Pearson correlation, 0.718; $P < .01$). Repeated measures analysis showed a significant relationship between confidence and quality (parameter estimate, 0.732; standard error, 0.021; $P < .01$) adjusting for rater and review.

Conclusion: For experienced shoulder surgeons, the Snyder classification is a reliable system for identifying SLAP lesions.

Keywords: arthroscopy; shoulder; labrum; superior labral anterior posterior lesions; reliability

Lesions of the superior labrum in the shoulder joint were first appreciated with the development of the arthroscope. In 1985, Andrews et al[1] viewed superior labrum lesions arthroscopically and noted the variability of the superior labrum. In 1990, Snyder et al[23] classified superior labral anterior and posterior (SLAP) lesion variants into 4 types: type I, fraying of the labrum without detachment; type II, detachment of the biceps and labrum from the superior glenoid; type III, detachment with a bucket-handle tear; and type IV, detachment with a bucket-handle tear and a tear of the biceps tendon. This classification system has been the one most commonly used by clinicians who evaluate and treat these lesions and by biomechanists who have sought to determine the cause of SLAP lesions.[7,17,23,25]

Subsequently, 2 modifications of that classification system have been reported.[15,19] Maffet et al[15] added 3 types of lesions: type V, detachment of the labrum from the biceps anchor across the front of the joint to include a Bankart lesion; type VI, a flap tear of the superior labrum; and type VII, a tear of the middle glenohumeral ligament. Then Morgan et al[19] divided type II lesions into 3 subtypes: type IIA, detachment of the biceps and anterior labrum from the superior glenoid; type IIB, detachment of the biceps and posterior labrum from the superior rim; and type IIC, detachment of the biceps and the anterior and posterior labrum from the glenoid rim.

Despite the widespread use of the Snyder classification system by clinicians, to our knowledge, only 1 study has analyzed the reliability of this classification scheme.[7]

The American Journal of Sports Medicine, Vol. 39, No. 5
DOI: 10.1177/0363546510392332

Although this extensive study by Gobezie et al[7] included 73 surgeons of varying experience who evaluated 22 videotapes of normal and abnormal superior labrums, it did not evaluate the effect of the quality of the videotapes on the results nor the reliability of the schema of Morgan et al.[19]

The goals of our study were to (1) evaluate the intraobserver and interobserver reliability of the Snyder classification of labrum tears among experienced shoulder surgeons, (2) determine the effect of simplifying the Snyder classification into normal versus abnormal labrums, (3) determine the reliability of subdividing type II SLAP lesions into 3 varieties as described by Morgan et al,[19] and (4) evaluate the effect of the quality of the videotapes on the diagnosis and on the confidence of the evaluator about the diagnosis.

On the basis of our clinical experience, we hypothesized that: (1) the Snyder classification system would produce excellent intraobserver and interobserver reliability for distinguishing normal from abnormal labrum tears but that the agreement for types II, III, and IV would be poor; (2) dividing type II SLAP lesions into 3 subgroups would produce lower intraobserver and interobserver reliability; and (3) the quality of the videotape would affect the confidence of the examiners when making the diagnosis.

MATERIALS AND METHODS

This study was approved by our Institutional Review Board, and all patient information was used according to current privacy regulations.

Videotape Selection

From the database records of the senior author's (E.G.M.) practice, we retrieved 90 archived videotapes of shoulder arthroscopies for patients with diagnosed SLAP lesions. The videotapes were archived from the patient charts of a shoulder database maintained by the senior author.[8,11,16] An independent observer who was a sports medicine research fellow at our institution and who had finished an orthopaedic surgery residency reviewed the videotapes and was instructed to select as equal a distribution as possible (between 15 and 20) of each of the 4 main types of SLAP lesions. The independent observer was instructed to also select from the database 15 to 20 videotapes of normal superior labrums (no detachment or fraying). Videotapes of the more complex SLAP lesions as reported by Maffet et al[15] were excluded—specifically, type V to type VIII lesions.

The independent observer edited the videotapes into video clips that showed only the examination of the superior labrum. Each clip included probing of (1) the superior labrum from the equator in the front of the glenoid to the equator in the posterior edge of the glenoid and (2) the biceps attachment to the superior glenoid tubercle. The probing of the labrum was done through an anterior portal with a nerve hook in all cases. In each videotape, the patient position was lateral decubitus with the arthroscope placed in a posterior portal, which is the usual clinical practice; the arthroscope was not placed through the anterior portal to visualize the posterior labrum.

The clips did not include any ancillary information that might influence the examiner's interpretation (eg, the presence of a Bankart lesion), and no other information about the patients (eg, history, examination, or radiologic studies) was available to the examiners.

Because we were interested in reliability of the classification of the labrum tears and not the treatment, there was no independent gold standard applied for the accuracy of the SLAP lesion diagnosis. The type of SLAP lesion chosen by the independent observer was based on the final diagnosis in the patient medical record and not by any other objective method. From the diagnoses in the patient records, the video clips selected by the independent observer included 20 normal labrum, 20 type I, 30 type II, 10 type III, and 10 type IV SLAP lesions. No attempt was made to obtain equal numbers of subtypes IIA, IIB, or IIC lesions.

Examiner Selection

For examiners, we chose 5 clinicians based on their extensive experience with shoulder arthroscopy; there were no other specific selection criteria. The senior author was not included among the 5 examiners. The examiners had been fellowship trained in sports medicine (n, 3) or shoulder surgery (n, 2), and the average time in practice after fellowship was 16.9 years (range, 11-23 years). On average, the examiners performed 380 shoulder operations annually (range, 250-500), of which 305 (range, 200-400) were shoulder arthroscopies. The examiners estimated that, on average,

**Address correspondence to Edward G. McFarland, MD, c/o Elaine P. Henze, BJ, ELS, Medical Editor and Director, Editorial Services, Department of Orthopaedic Surgery, Johns Hopkins Bayview Medical Center, 4940 Eastern Avenue, #A665, Baltimore, MD 21224-2780 (e-mail: ehenze1@jhmi.edu).
*Departments of Biomedical Engineering, Physical Medicine and Rehabilitation, The Johns Hopkins University, Baltimore, Maryland.
†Department of Orthopedic Surgery, Osaka Medical College, Takatsuki, Japan.
‡CU Sports Medicine, University of Colorado, Boulder, Colorado.
§Midwest Orthopaedics at Rush, Rush University Medical Center, Chicago, Illinois.
‖Sacramento Knee and Sports Medicine Center, Sacramento, California.
¶University of Pittsburgh and McMahon Orthopedics and Rehabilitation, Pittsburgh, Pennsylvania.
#Vanderbilt Orthopaedic Institute, Vanderbilt University Medical Center, Nashville, Tennessee.
††Division of Shoulder Surgery, Department of Orthopaedic Surgery, The Johns Hopkins University, Baltimore, Maryland.
Presented at the annual meeting of the American Orthopaedic Society for Sports Medicine, Providence, Rhode Island, July 2010.
One or more authors has declared a potential conflict of interest or source of funding: Dr Dunn was supported by grant No. 5K23 AR052392-04 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases, as well as the AOSSM-MTF Career Development Award Supplement; the Johns Hopkins University Division of Shoulder Surgery has also received financial support from DePuy Mitek for its shoulder fellowship program.

80% (range, 60%-90%) of their current surgical practice constituted shoulder procedures. The examiners were blinded to the distribution of the types of SLAP lesions.

Each examiner then viewed the videotapes and (1) classified the labrum as *normal* or as 1 of the 4 Snyder types (I, II, III, IV); (2) further classified any type II lesion as 1 of the subtypes described by Morgan et al[19] (anterior only, posterior only, or combined lesion); (3) graded the quality of each videotape via an analog scale (extremely poor to perfect visibility), which was 100 mm in length; and (4) graded his or her confidence about each diagnosis via another analog scale (no confidence to full confidence). Although there was the opportunity to grade any labrum as a type II and to grade it as anterior, posterior, or combined for each video, there was no indication on the grading sheet that a labrum displayed a type II tear. At a minimum of 2 months later, each examiner reviewed the same videotapes and repeated those 4 evaluations.

Statistical Analysis

To answer the objectives of the current study, all labrum tears were classified according to the schemes described above: original Snyder classification, where the variables were normal, I, II, III, or IV; simplified Snyder classification, where the variables were normal (normal and type I) versus abnormal (types II through IV); and the Snyder classification (normal, I, III, IV), including type II subtypes (anterior, posterior, combined) as suggested by Morgan et al.[19] For each classification scheme, we calculated interobserver and intraobserver agreement using standard statistical techniques. To determine the reliability of dividing the labrum tears into normal or abnormal categories, we placed no lesion and type I lesions into the normal group based on literature suggesting that type I lesions are associated with degenerative joint disease and age and not with distinctive symptoms.[11] We placed the remaining type II to type IV lesions into the abnormal group. To determine the reliability of dividing type II lesions into 3 subtypes, the analysis was performed with the options of normal, type I, type II anterior, type II posterior, type II combined, type III, and type IV.

For each analysis, intraobserver agreement was defined as the agreement of a single rater assessing the same entity across 2 sessions and was estimated using Cohen κ coefficient.[5] Interobserver agreement was defined as the agreement among multiple raters assessing the same entity across all ratings and entities, and it was estimated using the intraclass correlation coefficient.[18] These statistics are thought to be more robust than the simple percentage agreement calculation because Cohen κ statistic and the intraclass correlation coefficient take into account the agreement occurring by chance. The strength of agreement of κ and intraclass correlation coefficient was assessed using criteria developed by Landis and Koch.[13,14] Agreement among examiners was defined as poor (<0.0), slight (0.0-0.2), fair (0.21-0.4), moderate (0.41-0.6), substantial (0.61-0.8), or almost perfect (0.81-1.0).[14] The absolute agreement among observers was defined as the percentage of the observations of the surgeons where all 5 observers agreed on the diagnosis of the type of SLAP lesion.

To evaluate the relationship of the confidence of the diagnosis to the quality of the videotapes, Pearson correlation coefficients were calculated for each rater and review. A repeated measures linear regression model was used to estimate the relationship between confidence of the diagnosis and the quality of the videotape, adjusting for the interdependence of multiple ratings of the same videotape.

All data were entered into a spreadsheet and statistically analyzed with a standard computerized statistical package (SPSS 16.0, SPSS Inc, Chicago, Illinois).

RESULTS

Table 1 displays the distribution of the Snyder classifications for SLAP lesions by observer made during initial review. Tables 2 and 3 present the reliability statistics. Based on the original Snyder classification of labrum tears (normal and type I-IV), intraobserver agreement ranged between 0.555 and 0.787: 4 examiners showed substantial agreement and 1 showed moderate agreement. Interobserver agreement was 0.808 (initial review) and 0.800 (follow-up review), representing substantial agreement.

Based on the simplified Snyder classification (normal and type I versus types II-IV), intraobserver agreement ranged from 0.666 through 0.867. Two raters showed almost perfect agreement and 3 showed substantial agreement. Interobserver agreement was 0.641 (initial review) and 0.654 (follow-up review), representing substantial agreement. The absolute agreement (58.4%) when based on a classification of the labrum as either normal or abnormal was significantly higher ($P < .001$) than the absolute agreement when the traditional Snyder classification (normal and types I-IV) was used to classify the labrum tears (27.2%).

When based on the modified Snyder classification suggested by Morgan et al[19] (including the subclassifications for type II lesions), intraobserver agreement ranged from 0.404 through 0.719. Three raters showed substantial agreement, 1 showed moderate agreement, and 1 fair agreement. Interobserver agreement was 0.813 (initial review, almost perfect agreement) and 0.795 (follow-up review, substantial agreement).

The examiner's confidence about the diagnosis correlated significantly with the quality of the video (Pearson correlation range, 0.567-0.881; all $P < .01$) (Table 4). Repeated measures analysis showed a significant relationship between confidence and quality (parameter estimate, 0.732; standard error, 0.021; $P < .01$), adjusting for rater and review.

DISCUSSION

Our results show that experienced shoulder surgeons who evaluate the superior labrum and classify the findings using the Snyder classification have substantial intraobserver and interobserver reliability; that simplifying the labrum classification into normal or abnormal increased the absolute agreement and intraobserver reliability but

TABLE 1
Classification of Superior Labral Anterior Posterior Lesions by Examiner: Initial Review

	Videotapes Identified by Each Examiner as Specific SLAP Lesions, n				
Type of Lesion	1	2	3	4	5
Normal	58	20	39	18	19
I	4	24	20	19	26
IIA	4	0	3	12	2
IIB	2	7	10	3	6
IIC	10	23	2	25	24
III	3	5	9	7	6
IV	9	11	7	6	7

TABLE 2
Intraobserver Agreement for Classifications of Labrum Tears

	κ Coefficient by Examiner					
Lesion Classification	1	2	3	4	5	Average κ Coefficient
Original Snyder classification	.555	.692	.787	.623	.693	.670
Simplified Snyder classification (normal and type I vs type II-IV)	.666	.867	.855	.771	.800	.792
Modified Snyder classification as suggested by Morgan et al[19]	.404	.622	.719	.586	.658	.598

TABLE 3
Interobserver Agreement for Classifications of Labrum Tears: Superior Labral Anterior Posterior Lesion Classification

Rating	Absolute Agreement Among All Surgeons (%)	Intraclass Correlation Coefficient[a]
Original Snyder classification		
First	30.0	0.808 (0.751, 0.858)
Second	24.4	0.800 (0.742, 0.852)
Average	27.2	0.804
Simplified Snyder classification[b]		
First	58.9	0.641 (0.556, 0.723)
Second	57.8	0.654 (0.570, 0.734)
Average	58.4	0.648
Modified Snyder classification[c]		
First	21.1	0.813 (0.757, 0.861)
Second	16.7	0.795 (0.735, 0.847)
Average	18.9	0.804

[a]Type C single-measures intraclass correlation coefficients using agreement consistency definition. Two-way random effects model where people effects and measure effects are random. Confidence interval (95%) in parentheses.
[b]Normal and type I vs type II-IV.
[c]As suggested by Morgan et al.[19]

did not increase the interobserver reliability; that giving examiners the choice of the 3 subtypes of type II SLAP lesions did not affect the reliability of the responses; and that videotape quality significantly correlated with the examiner's confidence about the diagnosis and is therefore a factor that should be considered when using videotapes for reliability studies.

The major limitation of this study is that it is a reliability study designed to measure only the agreement between multiple observations by each examiner and by multiple examiners. As a result, a gold standard was not used to determine the validity of the Snyder classification; in other words, the examiners' responses were not judged to be right or wrong but were rather compared with the subsequent responses and to the responses of the other examiners. Although the independent examiner was asked to provide videotapes with a wide distribution of superior labrum anatomy and abnormalities, there were no other validity criteria used to determine the types of lesions used in this study. Similarly, the effects of the interpretation of the lesion on treatment or the results of treatment were not studied here; therefore, treatment decisions cannot be based on this study.

To our knowledge, only 2 previous studies[7,22] have addressed the intraobserver and interobserver reliability of the Snyder classification. Sasyniuk et al[22] reported on the interobserver agreement of orthopaedic surgeons who viewed the videotapes of 20 patients with anterior shoulder instability. They did not perform a rigorous statistical

TABLE 4
Correlation of Examiner Confidence About the Diagnosis With the Quality of the Video[a]

	Pearson Correlation by Examiner					
Review Sequence	1	2	3	4	5	Average Pearson Correlation
Initial review	.703	.645	.831	.741	.568	.698
Follow-up review	.863	.676	.881	.707	.567	.739

[a]All values, $P < .01$.

analysis, did not specify the experience of the surgeons involved, and found that the overall agreement for all the structures visualized in the shoulder was poor (<40%). The surgeons' agreement for superior labrum lesions was 60% if the surgeon who performed the arthroscopy was not included in the analysis; however, the agreement increased to 70% if the surgeon who performed the arthroscopy was included in the analysis. Also, in their study, the superior labrum was only determined to be normal or abnormal, and the utility of the Snyder classification was not evaluated.

The most extensive study of the reliability of the Snyder classification was by Gobezie et al,[7] who sent a survey to the members of an orthopaedic organization who might be expected to perform shoulder arthroscopy. Of the 73 surgeons who responded, 68% were fellowship trained in arthroscopy and 51% had more than 10 years in practice. Of the respondents, 64% said that they did not have arthroscopic shoulder training in their fellowships and 64% said that they had little exposure to shoulder arthroscopy during their residency training. This group of 73 surgeons interpreted 22 video vignettes to calculate the interobserver reliability of interpreting the type of SLAP lesion. Gobezie et al[7] found that intraobserver reliability was moderate (κ, 0.54) among the 17 surgeons who responded a second time and indicated that interobserver disagreement for normal through type IV lesions was "considerable."[7]

Our study differs from those 2 studies[7,22] in several important respects. First, we used only experienced fellowship-trained shoulder surgeons who perform shoulder arthroscopy frequently (on average, >305 per year), which may have accounted for our higher intraobserver and interobserver results. Because we did not analyze how much experience was needed (in terms of cases) or how many years of practice were needed to obtain higher intraobserver or interobserver agreements, the learning curve or the amount of experience needed was not quantified in our study. Although our study verifies that the Snyder classification has substantial reliability, we did not attempt to measure treatment decisions based on examiner choices. Last, we used a large number of videotapes, whereas the study by Gobezie et al[7] involved the interpretation of only 22 videotapes, which might have affected the statistical analysis.

In our study, simplifying the Snyder classification into normal or abnormal increased the absolute agreement among observers to more than 50%. Dividing the labrum into these 2 groups also increased intraobserver agreement to nearly perfect. It is unclear why the interobserver agreement was essentially unaffected by this attempt to simplify the classification. Note that we did not ask the observers to classify the labrum as normal or abnormal, nor did we ask them to simply classify the labrum as attached and detached. Asking the observers to use these descriptive terms might have produced different results. Similarly, although type I labrum tears may take many forms, such as fraying, splitting, or cracking, we did not ask the observers to categorize these tears as 1 of these subtypes. Additional clinical study would be necessary to determine if this helped surgeons determine one method of treatment over another based on this classification of SLAP lesions.

To our knowledge, this study is the first to evaluate the effect of dividing type II SLAP lesions into 3 subtypes based on the reliability of the interpretations by the examiners of the labrum lesions. However, it should be noted that in this study we did not ask the observers to specifically classify only lesions believed to be type II lesions; if the examiner believed the labrum to be a type II lesion, we asked the examiner to subclassify it as 1 of the 3 subtypes. We did not give the examiners a group of type II lesions and request that they classify them. As a result, our data reflect only the effect of dividing the possible type II SLAP lesions into 3 types based on the overall results for the reliability of the examiners for all the SLAP lesions. It is possible that if the examiners were given a cohort of only type II SLAP lesions, the results would be different.[2,3,19]

Another limitation of this study is that including internal impingement maneuvers in the arthroscopic analysis might affect the detection and classification of these subtypes of labrum lesions.[10,16,19] In our study, there were no videotapes of the labrum "peeling off," which might have helped the surgeons make a more reliable interpretation of the posterior-superior variant of type II SLAP lesions.

Other variables might have influenced our results. First, the surgeons were not able to physically probe the labrum, which might have increased their ability to make a diagnosis. Second, the surgeons had no clinical information about the patients and did not perform an examination of the patients. Although this information might have increased the chance of a more reliable agreement, to our knowledge, there is no evidence in the literature that history or physical examination increases the ability to make the diagnosis of SLAP lesions. The literature suggests that the physical examination for SLAP lesions is inexact and that arthroscopy is the gold standard for making the diagnosis of SLAP lesions.[9,16,17,20,25]

Third, this study might have been influenced by the possible ambiguity in the definition of the types of SLAP lesions. Our examiners were not given strict criteria for making the diagnosis of the different types of SLAP lesions and were given only a copy of the figures from the original

article by Snyder et al[23] as a guide. Davidson and Rivenburgh[6] suggested that elevation of the superior labrum more than 5 mm implies that there is an abnormal lesion of the posterior and superior labrum. It is possible that if more strict criteria had been used and if the examiners had been trained using the criteria by Davidson and Rivenburgh,[6] agreement might have been better.

Fourth, variability of the anatomy of the superior labrum, especially in the anterior-superior quadrant of the glenoid, has been well documented.[3,4,6,12,17,21,24] Although our examiners were given the option of grading the labrum as normal, they were not given the choice of choosing a sublabral hole or other normal variations. It is unknown if this distinction would have influenced the results. Sasyniuk et al[22] found that structures and abnormalities with less variability increased the diagnostic agreement.

Last, to our knowledge, our study is the first to find that the quality of the videotapes does influence the ability of the examiners to determine similar diagnoses. This finding is consistent with the observation that good visibility during shoulder arthroscopy is needed when performing the surgery. However, there was no independent control of quality of the videotapes, so the observers might have been biased in their assessments. To our knowledge, there are no known measures of videotape quality, so the assessment by the observers cannot be measured against any gold standard. Because our conclusion is based on subjective interpretation of the observers, additional study is needed on how to perform these types of studies with elimination of low-quality videotapes or confirmation that the abnormality intended to be observed has a proper amount of clarity.

In summary, this study substantiates that when experienced shoulder surgeons interpret superior labrum tears using the Snyder classification, the reliability is substantial. When the labrum tears were interpreted as normal or abnormal, the intraobserver reliability and the absolute agreement were significantly increased, but the interobserver reliability was not increased. There was substantial agreement among the same observers for the subtypes of type II SLAP lesions, which suggests that the schema by Morgan et al[19] can be reliably used by experienced surgeons. In addition, this study shows that the quality of the videotapes does influence the confidence of the surgeons viewing the tapes. This variable should be taken into account in future reliability studies involving the use of videotapes.

REFERENCES

1. Andrews JR, Broussard TS, Carson WG. Arthroscopy of the shoulder in the management of partial tears of the rotator cuff: a preliminary report. *Arthroscopy*. 1985;1(2):117-122.
2. Andrews JR, Carson WG Jr, McLeod WD. Glenoid labrum tears related to the long head of the biceps. *Am J Sports Med*. 1985;13(5):337-341.
3. Burkhart SS, Morgan CD, Kibler WB. The disabled throwing shoulder: spectrum of pathology. Part I: pathoanatomy and biomechanics. *Arthroscopy*. 2003;19(4):404-420.
4. Chang D, Mohana-Borges A, Borso M, Chung CB. SLAP lesions: anatomy, clinical presentation, MR imaging diagnosis and characterization. *Eur J Radiol*. 2008;68(1):72-87.
5. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37-46.
6. Davidson PA, Rivenburgh DW. Mobile superior glenoid labrum: a normal variant or pathologic condition? *Am J Sports Med*. 2004;32(4):962-966.
7. Gobezie R, Zurakowski D, Lavery K, Millett PJ, Cole BJ, Warner JJ. Analysis of interobserver and intraobserver variability in the diagnosis and treatment of SLAP tears using the Snyder classification. *Am J Sports Med*. 2008;36(7):1373-1379.
8. Jia X, Ji JH, Petersen SA, Freehill MT, McFarland EG. An analysis of shoulder laxity in patients undergoing shoulder surgery. *J Bone Joint Surg Am*. 2009;91(9):2144-2150.
9. Jones GL, Galluch DB. Clinical assessment of superior glenoid labral lesions: a systematic review. *Clin Orthop Relat Res*. 2007;455:45-51.
10. Kibler WB. Specificity and sensitivity of the anterior slide test in throwing athletes with superior glenoid labral tears. *Arthroscopy*. 1995;11(3):296-300.
11. Kim TK, Queale WS, Cosgarea AJ, McFarland EG. Clinical features of the different types of SLAP lesions: an analysis of one hundred and thirty-nine cases. *J Bone Joint Surg Am*. 2003;85(1):66-71.
12. Kreitner KF, Botchen K, Rude J, Bittinger F, Krummenauer F, Thelen M. Superior labrum and labral-bicipital complex: MR imaging with pathologic-anatomic and histologic correlation. *AJR Am J Roentgenol*. 1998;170(3):599-605.
13. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. 1977;33(2):363-374.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
15. Maffet MW, Gartsman GM, Moseley B. Superior labrum-biceps tendon complex lesions of the shoulder. *Am J Sports Med*. 1995;23(1):93-98.
16. McFarland EG, Kim TK, Savino RM. Clinical assessment of three common tests for superior labral anterior-posterior lesions. *Am J Sports Med*. 2002;30(6):810-815.
17. McFarland EG, Tanaka MJ, Garzon-Muvdi J, Jia X, Petersen SA. Clinical and imaging assessment for superior labrum anterior and posterior lesions. *Curr Sports Med Rep*. 2009;8(5):234-239.
18. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients [published correction appears in Psychol Methods. 1996;31(34):390]. *Psychol Methods*. 1996;1(1):30-46.
19. Morgan CD, Burkhart SS, Palmeri M, Gillespie M. Type II SLAP lesions: three subtypes and their relationships to superior instability and rotator cuff tears. *Arthroscopy*. 1998;14(6):553-565.
20. Parentis MA, Glousman RE, Mohr KS, Yocum LA. An evaluation of the provocative tests for superior labral anterior posterior lesions. *Am J Sports Med*. 2006;34(2):265-268.
21. Rao AG, Kim TK, Chronopoulos E, McFarland EG. Anatomical variants in the anterosuperior aspect of the glenoid labrum: a statistical analysis of seventy-three cases. *J Bone Joint Surg Am*. 2003;85(4):653-659.
22. Sasyniuk TM, Mohtadi NGH, Hollinshead RM, Russell ML, Fick GH. The inter-rater reliability of shoulder arthroscopy. *Arthroscopy*. 2007;23(9):971-977.
23. Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ. SLAP lesions of the shoulder. *Arthroscopy*. 1990;6(4):274-279.
24. Waldt S, Metz S, Burkart A, et al. Variants of the superior labrum and labro-bicipital complex: a comparative study of shoulder specimens using MR arthrography, multi-slice CT arthrography and anatomical dissection. *Eur Radiol*. 2006;16(2):451-458.
25. Walton DM, Sadi J. Identifying SLAP lesions: a meta-analysis of clinical tests and exercise in clinical reasoning. *Phys Ther Sport*. 2008;9(4):167-176.