Journal of ISAKOS 7 (2022) 56–61

Contents lists available at ScienceDirect

Journal of ISAKOS

journal homepage: www.elsevier.com/locate/jisakos

Original Research

Reliable interobserver and intraobserver agreement of the International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine (ISAKOS) classification system of rotator cuff tears

Emilio Calvo [a,*], Carlos Rebollón [b], Eiji Itoi [c], Andreas Imhoff [d], Felix H. Savoie [e], Guillermo Arce [f]

[a] *Department of Orthopedic Surgery and Traumatology, IIS-Fundación Jiménez Díaz, Universidad Autónoma. Madrid, Spain*
[b] *Centro Ortopédico Panama Clinic, The Panama Clinic, Panama City, Panama*
[c] *Department of Orthopaedic Surgery, Tohoku University School of Medicine, Sendai, Japan*
[d] *Department of Orthopaedic Sports Medicine, Klinikum Rechts der Isar, Technische Universitat Munchen, School of Medicine, Munich, Germany*
[e] *Tulane Institute of Sports Medicine, Department of Orthopaedics, Tulane University School of Medicine, New Orleans, Louisiana, USA*
[f] *Instituto Argentino de Diagnóstico y Tratamiento, Buenos Aires, Argentina*

ARTICLE INFO

Keywords:
Shoulder
Rotator cuff tear
Classification
ISAKOS
Arthroscopy
Validity
Reliability
Agreement

ABSTRACT

Objective: The ISAKOS Shoulder Committee developed a new comprehensive classification system aimed to describe all rotator cuff tears. The five characteristics of the tears included are pattern (P), extension (E), fatty atrophy (A), retraction (R), and location (L), conforming to the acronym "PEARL." The objective of this study is to assess if the ISAKOS Rotator Cuff Tear Classification System is reliable by measuring the intraobserver and interobserver multirater reliability.

Methods: Arthroscopic videos of 36 rotator cuff tears, including tears of varying sizes and configurations, were evaluated twice by four surgeons from different continents blinded to tear type. Intraobserver and interobserver reliability was measured using multirater and intraobserver Kappa coefficients.

Results: Intraobserver reliability: An almost perfect agreement for "location" ($\kappa = 0.98$), substantial agreement for "extension" ($\kappa = 0.73$) were obtained for mean intrarater kappa, lower κ agreement for "pattern" ($\kappa = 0.58$), and relatively high agreement of 0.79 for "retraction." Every characteristic of "location" had an almost perfect agreement among the surgeons ($\kappa = 0.91$). Intrarrater reliability: In the partial-thickness posterosuperior tears "location," there was good agreement in tears involving less than 50% of the tendon ($\kappa = 0.74$) and moderate in those deeper than 50% of the tendon thickness ($\kappa = 0.58$). "Extension" in full-thickness posterosuperior RCT achieved moderate agreement. Within the anterior subscapularis tears, we have a substantial agreement in Lafosse's Type 1 ($\kappa = 0.73$), moderate in Types 2 and 3 ($\kappa = 0.45$ and $\kappa = 0.46$) and slight agreement in Type 4 ($\kappa = 0.06$). The overall kappa summarizing all categories indicated moderate agreement ($\kappa = 0.52$). The articular pattern in the partial-thickness posterosuperior RCT "location" had a perfect kappa of 1, while the bursal pattern showed an almost complete agreement ($\kappa = 0.87$). In the full-thickness posterosuperior RCTs, we observed fair agreement in C, U, and reverse L configurations and slight agreement in L pattern ($\kappa = 0.18$). The overall kappa is 0.44 with a confidence interval of 0.41–0.47. There was substantial agreement in every category and the overall kappa for "retraction" ($k = 0.70$).

Conclusion: ISAKOS rotator cuff tear classification system provides sufficient interobserver reliability for communicating among surgeons and for pooling of data from clinical studies.

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

Abbreviations: RCT, Rotator cuff tear.

* Corresponding author. Department of Orthopedic Surgery and Traumatology, IIS-Fundación Jiménez Díaz, Universidad Autónoma, Avda Reyes Católicos, 2, Madrid, 28040, Spain. Tel.: +34619145936.

E-mail address: ecalvo@fjd.es (E. Calvo).

https://doi.org/10.1016/j.jisako.2021.12.004
Received 8 December 2021; Accepted 12 December 2021
Available online 27 December 2021

What are the new findings?

- The shoulder committee of ISAKOS has developed a comprehensive system for rotator cuff tears, including a combination of previously existing systems.
- Pattern (P), extension (E), fatty atrophy (A), retraction (R), and location (L) were the key characteristics used to describe rotator cuff tears.
- The interobserver and intraobserver reliability of the ISAKOS Classification System of Rotator Cuff Tears is described.

Introduction

Rotator cuff tendon pathology is the most common cause of shoulder pain [2,17]. Rotator cuff tears (RCT) span a broad spectrum of injuries that vary greatly in the indications for treatment and surgical repair technique (2). The outcome and prognosis are not only determined by the tear size or the degree of retraction but also by other characteristics such as the specific tendons involved, the shape of the tear, or the severity of muscular atrophy (5,6,3,17). The existing controversy on the treatment of RCTs is partly due to the uncertainty of surgeons about the type and characteristics of RCTs reported in the different investigations (2,3). At present, there is no widely accepted classification system for RCTs, although some systems have been used to describe restricted types of RCTs [3,4,8,13,18]. Although addressing the details of the specific problem, they are not comprehensive and lack critical information to establish a correct therapeutic approach or prognosis. Consequently, none have been validated nor are used universally.

After reviewing the currently available systems for RCTs, the ISAKOS Shoulder Committee has developed a new comprehensive classification system for RCTs based in part on previously published classifications. The new system aims to describe all rotator cuff tears in a comprehensive and straightforward manner, which meets three criteria: allows the surgeon to establish indications, predicts difficulties during the procedure, and prognosticates. The five essential characteristics of the rotator cuff tears included in this system are pattern (P), extension (E), fatty atrophy (A), retraction (R), and location (L), conforming to the acronym "PEARL" [2].

We hypothesize that the ISAKOS Rotator Cuff Tear Classification System, as proposed by the ISAKOS Shoulder Committee, is reliable and reproducible. The purpose of this study is to assess if the ISAKOS Rotator Cuff Tear Classification System is reliable and reproducible by measuring the intraobserver and interobserver multirater agreement using surgical arthroscopic videos.

Materials and methods

The ISAKOS Shoulder Committee agreed that the new classification should allow: (1) validity, (2) reliability, (3) to guide an adequate therapeutic approach, (4) and to establish an accurate prognosis, ensuring realistic patient expectations. Based on previously reported classifications, a comprehensive system was developed, including among a combination of previously existing systems trying to classify all types of RCTs. The first step was to agree on the key characteristics to be defined in any RCT to outline the correct therapeutic approach in terms of indications, surgical technique, and realistic prognosis. RCT pattern (P), extension (E), muscular fatty atrophy (A), tendon retraction (R), and tear location (L) "PEARL" were the primary assets agreed by consensus to be included in the ISAKOS RCT classification system (Fig. 1) [2].

The etiology, clinical symptoms, natural history, and therapeutic approach of posterior RCTs involving the supraspinatus, infraspinatus, and teres minor compared to anterior subscapularis tears are different [15]. Therefore, RCT location, i.e. posterosuperior or anterior, is the first condition to be defined.

RCT extension determines treatment and prognosis, and the committee agreed to follow the tear size system suggested by de Orio and Cofield for full-thickness posterosuperior tears (4). This system provides information not only on the size but also on the number of tendons involved. The full-thickness tears are classified as C1 (small complete tear, pinhole-sized), C2 (moderate tear less than 2 cm of only one tendon without retraction), C3 (large complete tear with an entire tendon with minimal retraction, usually 3–4 cm), or C4 (massive rotator cuff tear involving two or more rotator cuff tendon with associated retraction and scarring of the remaining tendon). The Snyder et al. classification system was followed for posterosuperior partial-thickness RCTs that were classified as articular-sided and bursal-sided or interstitial tears [9]. Partial-thickness RCTs involving more than half of the tendon thickness are more prone to progress, and that they outperform better if treated surgically compared to more superficial tears [14,18]. Consequently, partial-thickness RCTs involving over or fewer than 50 % of tendon thickness were differentiated. For subscapularis tears, the Lafosse et al. classification was suggested since it is the most used [8]. According to Lafosse et al., type 1 subscapularis lesions are simple fibrillation or erosions of the superior third. Type 2 consists of detachment confined to the superior subscapularis third. Type 3 involves the entire height of the tendon insertion, but the muscular portion is preserved, limiting tendon retraction. Type 4 is complete subscapularis detachment from the lesser tuberosity, but the humeral head remains well centered. Type 5 also represents a complete rupture, but with the anterosuperior migration of the humeral head with associated fatty infiltration.

Recognizing the geometric pattern of RCTs was considered crucial for preoperative planning to perform the repair. The committee agreed to use the classification of RCT described by Davidson and Burkhart because it contains the four types of geometric RCT patterns: crescent-shaped tears, U-shaped tears, L-shaped tears, and reverse L-shaped tears [3].

The classification most used to describe tendon retraction is that suggested by Patte [13]. This classification uses the distance between the retracted tendon and its original insertion on the greater tuberosity in the coronal plane and defines three stages. Stage 1 is a tear with minimal

LOCATION (L)	EXTENSION (E)	PATTERN (P)	FATTY ATROPHY (A)			RETRACTION (R)
Partial-thickness posterosuperior	>50% thickness <50% thickness	A (articular) B (bursal) I (interstitial)				
Full-thickness posterosuperior	C1(small) C2 (moderate) C3 (large) C4 (massive)	C (crescent shape) U (U shape) L (L shape) rL (reverse L shape)	SS0 SS1 SS2 SS3 SS4	IS0 IS1 IS2 IS3 IS4	TM0 TM1 TM2 TM3 TM4	1 2 3
Full-thickness anterior	Type 1 Type 2 Type 3 Type 4 Type 5		SC0 SC1 SC2 SC3 SC4			

Fig. 1. ISAKOS Rotator cuff tear classification system. SS, supraspinatus; IS, infraspinatus; TM, teres minor; SC, subscapularis. Anterior subscapularis tears are classified following Lafosse et al[8]. Degree of fatty infiltration is scored according to Fuchs et al[5] and Goutallier et al[6]. Retraction is graded following Patte et al[13].

retraction; in stage 2, the tendon is retracted medially to a level between the greater tuberosity footprint and the glenoid, and stage 3 represents a tear retracted to the level of the glenoid.

Fatty infiltration and muscle atrophy of rotator cuff muscles are significant prognostic factors in RCTs and should be included in any classification system. The severity of fatty infiltration and muscle atrophy is classified according to the computed tomography classification by Goutallier et al. [6], which was later adapted to magnetic resonance imaging by Fuchs et al. [5]. This system for evaluating fatty infiltration has already been validated, and it relies on imaging studies and not on arthroscopy [16]. Therefore, the committee decided to exclude it from the validity and reliability analysis.

Collection and editing of classification data

Arthroscopic videos of 36 RCTs from 36 patients operated by the senior author were collected. Videos included a well-adjusted number of all patterns of varying sizes and configurations covered by the ISAKOS classification system. Exclusion criteria for the videos to be eligible for the study include previous surgery, previous proximal humerus fracture, and the presence of severe glenohumeral osteoarthritis.

Procedures were performed under combined brachial plexus block and general anesthesia in the beach chair position. The glenohumeral joint was first examined from a conventional posterior portal using a standard 30° scope. If a partial articular-sided tear was noted, it was debrided from an anterior portal to better judge thickness and extension. Partial articular-sided ruptures involving the supraspinatus and infraspinatus tendons were also assessed, visualizing from the anterior portal.

Once the glenohumeral examination was completed, the subacromial space was inspected from the posterior portal, and a bursectomy performed from a lateral portal located 3 cm lateral to the lateral edge of the acromion in continuity with the posterior side of the acromioclavicular joint. Bursectomy and tendon debridement were completed to provide a clear image of the tear characteristics. If deemed indicated, biceps tenotomy or tenodesis was carried out before the tendon evaluation. The scope was then moved to the lateral portal to evaluate the RCT. All rotator cuff tendon insertion was also debrided, clearly visualized, and palpated. The tendon mobilized in different directions using grasping forceps to check mobility and to help in defining the geographic pattern of the tear. A calibrated probe with 1-mm laser-etched lines was introduced through anterior or posterior portals to measure retraction, as well as the anterior-posterior and mediolateral dimensions of the tear.

Two shoulder surgeons not participating in the reliability study and blinded to the cases collected the videos ascertaining that they were recorded following the previously described criteria. Large subscapularis tears were less prevalent and therefore required a longer recruiting period. All videos showed a complete view of the rotator cuff tendons, and measurements with the calibrated probe, as well as RCT mobilization and reduction of the tear were included. The arthroscopies were recorded without audio and lasted less than 60 s. According to the ISAKOS RCT classification system, there were 25 posterosuperior RCTs, including five supraspinatus partial-thickness tears, two bursal and three articular-sided, and 21 full-thickness posterosuperior tears, including a balanced number of sizes and geographical configurations. Five cases had isolated subscapularis tendon involvement.

Intra and interobserver reliability study

International intraobserver and interobserver reliability was tested among four orthopedic surgeons experienced in shoulder surgery who practiced in four countries from three continents (Argentina, Germany, Japan, and Spain) and that were members of four continental sports medicine and/or shoulder surgery societies: AANA, ESSKA, ESSSE-SECEC, JSS, JOSKAS, and SLARD. Videos were anonymized and randomized on four USB memory sticks that were distributed to the evaluating surgeons. For each case, a questionnaire a video tutorial with instructions was provided on the memory stick. The questionnaire included identification and classification of partial-thickness tears concerning side; full-thickness tears were classified by pattern, extension, retraction, and location of tendons involved according to the ISAKOS RCT classification system, as described in the instructions. Neither clinical data nor information on the final reconstruction technique was given for the cases corresponding to the provided videos. For the intraobserver reliability study, all participants evaluated videos twice, separated by a minimum of seven days between evaluations. Otherwise, observers were allowed to review the cases and use the rewind, replay, and pause functions as many times as needed at each assessment. Data from observations were transferred to an Excel spreadsheet (Microsoft, Redmond, Washington, USA) for descriptive and statistical analysis.

Statistical analysis

Multirater Kappa statistics were used to measure agreement among the surgeons with the ISAKOS RCT Classification System [15]. Overall intraobserver Kappa coefficients were calculated for every surgeon and category (11). A deeper look at every subcategory was done exclusively in interobserver analysis, as it considers all observations at once and provides a broader of view agreement. A κ of 0.00 represents agreement equivalent with random chance alone, whereas a κ of 1.00 represents the perfect agreement. The jackknife resampling technique was used to generate 95% confidence intervals for the interobserver Kappa coefficients, whereas, for the intraobserver 95% confidence intervals it was used according to the formula established by McHugh ML based on the standard error, since jackknife is not recommended for samples with only two observations [11]. The interpretation criteria for the κ statistic by Landis and Koch was followed: a κ of 0–0.2 represents a slight agreement, 0.21 to 0.40 is a fair agreement, 0.41 to 0.60 is a moderate agreement, 0.61 to 0.80 is a substantial agreement, and 0.81 to 1.0 is an almost perfect agreement [7].

Results

Videos were collected and edited in three months. Large subscapularis tears were less prevalent and therefore required a longer recruiting period. A total of 288 observations were given by the four surgeons (36 videos × 4 surgeons × 2 times = 288 evaluations). The intraobserver analysis is shown in Table 1, where averages have been added for clarification. Within the "location" criteria, an almost perfect agreement for every surgeon was obtained, with an average interval range at 95% of 0.03. Regarding "extension," every surgeon achieved substantial agreement, although the 95% intervals were wide. "Pattern" coefficients have a lower Kappa average of 0.58 across surgeons. Finally, although "retraction" has a relatively high overall Kappa average of 0.79, we can see that intraobserver agreement quite variable, with Kappas ranging from 0.56 to 1 (Table 1).

The overall Kappa summarizing all subcategories indicated moderate agreement (k = o.52; 95% CI 0.48, 0.56) Interobserver agreement for "location" and every subcategory achieved an almost perfect agreement among the surgeons, with the overall kappa 0.91 and the 95% confidence interval 0.88–0.94 (Table 2).

As categorization of "extension" depends on the "location" of the tear, values reflect the agreement for these combinations (Table 3). Considering the category "partial-thickness posterosuperior RCT," there was substantial consensus in tears involving less than 50% of the tendon ($\kappa = 0.74$) and moderate agreement in partial tears deeper than 50% of the tendon thickness ($\kappa = 0.58$). For the "full-thickness posterosuperior" RCT category, ranging from 0.44 to 0.53 there was a moderate agreement in every extension subcategory. Finally, within the anterior subscapularis tears, a substantial agreement was observed in Lafosse's Type 1 ($\kappa = 0.73$), moderate agreement for Types 2 and 3 ($\kappa = 0.45$ and $\kappa = 0.46$, respectively) and minimal agreement for Type 4 ($\kappa = 0.06$). There were no cases considered for type 5.

Table 1
Overall intraobserver agreement of the ISAKOS RCT classification system.

Surgeon	Location		Extension		Pattern		Retraction	
	Overall Kappa	Interval Range at 95% (±)	Overall Kappa	Interval Range at 95% (±)	Overall Kappa	Interval Range at 95% (±)	Overall Kappa	Interval range at 95% (±)
1	1.00	0.00	0.64	0.17	0.56	0.19	1.00	0.00
2	0.90	0.13	0.74	0.16	0.59	0.21	0.56	0.37
3	1.00	0.00	0.77	0.15	0.56	0.22	0.70	0.29
4	1.00	0.00	0.77	0.14	0.63	0.19	0.90	0.18
Mean	0.98	0.03	0.73	0.16	0.58	0.20	0.79	0.21

Table 2
LOCATION: Interobserver agreement of the ISAKOS RCT classification system.

Location	Number of observations	Expected agreement	Observed agreement	Kappa	Interval range at 95% (±)
Partial-thickness posterosuperior	54	0.63	0.97	0.91	0.01
Full-thickness posterosuperior	164	0.42	0.94	0.89	0.04
Anterior (subscapularis)	70	0.57	0.97	0.93	0.04
Overall	288	0.42	0.95	0.91	0.03

Table 3
EXTENSION: Interobserver agreement of the ISAKOS RCT classification system.

Location	Extension	Number of observations	Expected agreement	Observed Agreement	Kappa	Interval range at 95% (±)
Partial-thickness posterosuperior	<50% thickness	29	0.82	0.95	0.74	0.04
	>50% thickness	25	0.84	0.93	0.58	0.06
Full-thickness posterosuperior	1	18	0.88	0.94	0.51	0.12
	2	65	0.65	0.83	0.53	0.07
	3	52	0.70	0.84	0.44	0.08
	4	29	0.82	0.90	0.45	0.08
Anterior (subscapularis)	Type 1	15	0.90	0.97	0.73	0.04
	Type 2	30	0.81	0.90	0.45	0.06
	Type 3	21	0.86	0.93	0.46	0.10
	Type 4	4	0.97	0.97	0.06	0.11
	Type 5	0	–	–	–	–
Overall		288	0.13	0.59	0.52	0.04

The overall interrater kappa is 0.44 with a 95% confidence interval of 0.41-0.47 for a pattern (Table 4). Like "extension" criteria, "pattern" agreement must be considered in combination with "location." The articular pattern in the partial-thickness posterosuperior RCT location had a perfect kappa of 1, while the bursal pattern showed substantial agreement ($\kappa = 0.87$). In the full-thickness posterosuperior RCTs, we observe fair agreement in C, U, and reverse L configurations (0.34, 0.22, and 0.36, respectively) and slight agreement in the L pattern ($\kappa = 0.18$). There was substantial agreement in every category of "retraction," from 0.69 to 0.74, with similar overall interrater agreement ($\kappa = 0.70$; 95% CI 0.66–0.74; Table 5).

Discussion

Classifying RCTs is controversial since there is no uniformly acknowledged international system. A globally accepted system is necessary to precisely define the characteristics of the tear, for communicating among colleagues making researching investigations comparable, to guide in an adequate therapeutic approach, and to establish an accurate prognosis, ensuring realistic patient expectations. The ISAKOS RCT classification system has several innovative features that improve upon other systems. First, it provides a comprehensive classification system encompassing all types of RCTs. Second, it involves separate grading of the subscapularis and posterosuperior tears. Third, it integrates existing and commonly used classifications, familiar to surgeons but were not applicable to all RCTs and had not been validated. Finally, the system includes all key factors required for an adequate therapeutic approach and for establishing a prognosis.

The edited, video-recorded surgeries have proven to be a useful method for exploring inter and interrater reliability and accuracy of classifications systems because they provide exact and reproducible information [1,9,10]. Atoun et al. [1] and Lee [9] et al. respectively published, respectively, validation studies on subacromial impingement and partial-thickness RCTs classification and evaluated the rotator cuff from a

Table 4
PATTERN: Interobserver agreement of the ISAKOS RCT classification system.

Location	Pattern	Number of observations	Expected agreement	Observed Agreement	Kappa	Interval range at 95% (±)
Partial-thickness posterosuperior	Articular	16	0.94	1.00	1.00	0.00
	Bursal	38	0.71	0.96	0.87	0.02
Full-thickness posterosuperior	C	70	0.56	0.71	0.34	0.06
	U	30	0.76	0.81	0.22	0.02
	L	42	0.69	0.74	0.18	0.04
	Reverse L	20	0.83	0.89	0.36	0.10
Overall		216	0.21	0.56	0.44	0.03

Table 5
RETRACTION: Interobserver agreement of the ISAKOS RCT classification system.

Retraction	Number of observations	Expected agreement	Observed Agreement	Kappa	Interval range at 95% (±)
1	23	0.74	0.93	0.74	0.04
2	93	0.53	0.85	0.69	0.04
3	36	0.64	0.89	0.70	0.07
Overall	152	0.45	0.84	0.70	0.04

conventional posterior portal. We believe that RCT should be evaluated from a lateral portal because it provides a wide visualization of the entirety of the rotator cuff tendons -from subscapularis to teres minor- and allows a more precise estimation of retraction. Visualization from the lateral portal also provides a clear-cut assessment of the tear pattern furnishing the different approaches to reduce the tendon and restore anatomy. Furthermore, arthroscopic posterior portals only show the upper part of the subscapularis tendon if visualized from the glenohumeral joint, while the supraspinatus and coracohumeral ligament blocks visualization of the tendon and muscle if examined from the subacromial space.

This study demonstrates that many aspects of the ISAKOS RCT Classification System are repeatable and reliable between surgeons, even when they did not have the advantages of tactile sensation provided by probing to evaluate the tears personally. All the surgeons involved in the validation process were experienced in shoulder arthroscopy. Further, no discussion occurred between the surgeons about the cases or their particular method for grading injuries so as to replicate a first approach to the classification system without any guidance.

The intraobserver analysis yields an almost perfect agreement in the category "location," considering the average of all four surgeons ($\kappa = 0.98$). The interobserver agreement is also very high ($\kappa = 0.91$) considering all eight observations. Indeed, this high grade of the agreement was expected due to the nature of the features, as posterosuperior and anterior RCTs are easily differentiated. "Retraction" is more difficult to discern than "location" but has high agreement both in intraobserver ($\kappa = 0.79$) and interobserver analysis ($\kappa = 0.70$) overall. This substantial agreement in "retraction" is also expected since the anatomic landmarks used to define the degree of retraction in the Patte classification can be clearly recognized [13]. Therefore, we can conclude that both "location" and "retraction" are clearly acceptable within the classification system.

Concerning "extension," both the intraobserver and interobserver overall Kappa coefficients were slightly lower (average of $\kappa = 0.73$ Table 1, and $\kappa = 0.52$ Table 2) than that for location and retraction. This is expected due to the difficulty of correctly identifying these features on video clips. Also, the correct assignment of extension relies on location. The surgeons evaluated the "extension" over the anteroposterior and mediolateral diameter using a calibrated probe and should choose the longest diameter; however, the 95% confidence intervals were large. As we only have two observations in every intraobserver Kappa value, instead of 8, intraobserver interval ranges will be generally broader than interobserver's. In interobserver analysis, all subcategories in thickness posterosuperior RCT location reached at least a moderate agreement, with 164 observations in total. This makes it the "location" division with more cases in the study. In fact, posterosuperior RCTs are the most prevalent in clinical practice. Partial-thickness RCT obtained higher Kappa values concerning extension ($\kappa = 0.74$ and $\kappa = 0.58$) since only two categories were defined depending on the amount of tendon tissue involved in the tear (higher or lower 50% of tendon thickness). On the other hand, the level of agreement of subscapularis tears was notably lower (ranging from $\kappa = 0.06$ to $\kappa = 0.73$). Type 4 in subscapularis tear showed a level of agreement close to random choice ($\kappa = 0.06$). This low Kappa value can be explained by the fact that only one video with Type 4 subscapularis tear was included in the series because this specific category is much less prevalent than the other types of subscapularis tears. Finally, Type 5 subscapularis tears were not included, consequently, no reliability could be determined. However, this subcategory is rarely indicated for arthroscopic treatment. For all the reasons mentioned above, "extension" has an acceptable grade of agreement among the surgeons evaluated.

Finally, we find the lowest overall Kappa values in "pattern," both in intraobserver and interobserver analysis (average of $\kappa = 0.58$, and $\kappa = 0.44$, respectively, indicating moderate agreement. As "extension," "pattern" relies on "location," subcategories values in the interobserver analysis allow us to identify those with better or worse results. While there was an almost perfect agreement in allocating the pattern in posterosuperior partial-thickness RCT to any of the two groups (i.e. bursal or articular sided), the agreement in describing the pattern in full-thickness posterosuperior RCT was only fair to moderate (with the exception of L-shape pattern, with slight agreement). This can be explained because most operated patients whose surgeries were recorded for the study had chronic RCTs. Although different ways to reduce the tear using soft tissue forceps were shown in the videos, the initial pattern in chronic RCTs tends to be obscured by progressive retraction and the appearance of scar tissue. Consequently, RCT "pattern" can be accepted as described within the classification system.

This study has strengths and weaknesses. The strengths include a system created using existing systems that, although frequently used, only described a limited number of RCT types and had not been validated. It was ascertained that a balanced number of all patterns categories covered by the classification system were included. Video only was used, which allowed measurements, as well as retraction and reducibility testing. The study performed both intrareliability and interrater reliability. Intraobserver agreement was higher than interobserver in every characteristic of the system. Finally, the results should be internationally generalizable as the four surgeons who participated in the assessment live and practice medicine in four different countries from three different continents. Also, only experienced surgeons validated the ISAKOS RCT system; no trainee orthopedic surgeons participated. A limitation is that no sample size calculation was performed, but the number of videos considered is in line with similar reviewed articles [1,9,10,16]. Furthermore, recent types of RCTs described, such as musculotendinous tears, are not included [12]. It is reasonable to assume that the continuous updating of the system will be necessary as the knowledge of RCTs evolves. Also, experienced surgeons validated the ISAKOS RCT system, and trainee orthopedic surgeons did not participate in the evaluation process. Although this might question the reliability because of inexperience, it is expected that similar results could be obtained after learning the system.

In conclusion, results demonstrate that the ISAKOS RCT Classification System provides sufficient intraobserver and interobserver reliability for consistent communication amongst surgeons and uniform definitions for research applications of RCTs.

Ethics

The authors declare no competing interests related to this publication. The study obtained ethics approval from the institutional review board "Comite de Etica de la Investigacion de la Fundacion Jimenez Diaz" under act nr 19/20, and all participants gave informed consent before taking part.

Conflict of interest

There is no conflict of interest with regard to this study.

Acknowledgments

This work was supported by a grant from the International Society of Arthroscopy Knee Surgery and Orthopedics (ISAKOS).

The authors wish to acknowledge the support provided by David Alvarez and Daniel Mateos from Growth from Knowledge Spain for their technical assistance in the statistical analysis.

References

[1] Atoun E, Gilat R, van Tongel A, Pradhan R, Cohen O, Rath E, et al. Intraobserver and interobserver reliability of the Copeland-Levy classification for arthroscopic evaluation of subacromial impingement. J Shoulder Elbow Surg 2017;26:2167–72. https://doi.org/10.1016/j.jse.2017.07.018.

[2] Calvo E, Rebollón C, Morcillo D, Arce G. Diagnosis and classification of rotator cuff tears. In: Imhoff AB, Savoie III FH, editors. Rotator cuff across the life span. ISAKOS consensus book. Berlin, Germany: Springer-Verlag GmbH; 2019. p. 3–10.

[3] Davidson J, Burkhart SS. The geometric classification of rotator cuff tears: a system linking tear pattern to treatment and prognosis. Arthroscopy 2010;26(3):417–24. https://doi.org/10.1016/j.arthro.2009.07.009.

[4] De Orio 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(4):563–7. pmid: 6707035.

[5] Fuchs B, Weishaupt D, Zanetti M, Hodler J, Gerber C. Fatty degeneration of the muscles of the rotator cuff: assessment by computed tomography versus magnetic resonance imaging. J Shoulder Elbow Surg 1999;8:599–605. https://doi.org/10.1016/s1058-2746(99)90097-6.

[6] Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. Clin Orthop Relat Res 1994;304:78–83. pmid: 8020238.

[7] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74. https://doi.org/10.2307/2529310.

[8] Lafosse L, Jost B, Reiland Y, Audebert S, Toussaint B, Gobezie R. Structural integrity and clinical outcomes after arthroscopic repair of isolated subscapularis tears. J Bone Joint Surg Am 2007;89:1184–93. https://doi.org/10.2106/JBJS.F.00007.

[9] Lee CS, Davis SM, Doremus BAB, Kuok S, Stetson WB. Interobserver agreement in the classification of partial-thickness rotator cuff tears using the Snyder classification system. Orthop J Sports Med 2016;4. https://doi.org/10.1177/2325967116667058.

[10] Marx RG, Connor J, Lyman S, Amendola A, Andrish JT, Kaeding C, et al. Multirater agreement of arthroscopic grading of knee articular cartilage. Am J Sports Med 2005;33:1654–7. https://doi.org/10.1177/0363546505275129.

[11] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012; 22(3):276–82. https://doi.org/10.11613/BM.2012.031.

[12] Millett PJ, Hussain ZB, Fritz EM, Warth RJ, Katthagen JC, Pogorzelski J. Rotator cuff tears at the musculotendinous junction: classification and surgical options for repair and reconstruction. Arthrosc Tech 2017;6:e1075–85. https://doi.org/10.1016/j.eats.2017.03.023. eCollection 2017 Aug.

[13] Patte D. Classification of rotator cuff lesions. Clin Orthop Relat Res 1990;254:81–6. pmid: 2323151.

[14] Sano H, Wakabayashi I, Itoi E. Stress distribution in the supraspinatus tendon with partial-thickness tears: an analysis using two-dimensional finite element model. J Shoulder Elbow Surg 2006;15:100–5. https://doi.org/10.1016/j.jse.2005.04.003.

[15] Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. Psychol Bull 1979;86(2):420–8. https://doi.org/10.1037/0033-2909.86.2.420.

[16] Slabaugh MA, Friel NA, Karas V, Romeo AA, Verma NN, Cole BJ. Interobserver and intraobserver reliability of the Goutallier classification using magnetic resonance imaging: proposal of a simplified classification system to increase reliability. Am J Sports Med 2012;40:1728–34. https://doi.org/10.1177/0363546512452714.

[17] Teratani T. Comparison of epidemiology and outcomes of arthroscopic rotator cuff repair for anterosuperior and posterosuperior rotator cuff tears. J Orthop 2017; 14(4):430–3. https://doi.org/10.1016/j.jor.2017.07.003.

[18] Weber SC. Arthroscopic debridement and acromioplasty versus mini-open repair in the treatment of significant partial-thickness rotator cuff tears. Arthroscopy 1999; 15:126–31. https://doi.org/10.1053/ar.1999.v15.0150121.