Skeletal Radiology (2019) 48:1723–1733
https://doi.org/10.1007/s00256-019-03214-z

SCIENTIFIC ARTICLE

Diagnostic accuracy of MRI for detection of tears and instability of proximal long head of biceps tendon: an evaluation of 100 shoulders compared with arthroscopy

Eduardo Baptista[1] · **Eduardo A. Malavolta**[2] · **Mauro E. C. Gracitelli**[2] · **Daniel Alvarenga**[1] · **Marcelo Bordalo-Rodrigues**[1] · **Arnaldo A. Ferreira Neto**[2] · **Nestor de Barros**[1]

Received: 4 December 2018 / Revised: 15 March 2019 / Accepted: 25 March 2019 / Published online: 2 April 2019

Abstract

Objective To evaluate the diagnostic accuracy of magnetic resonance imaging (MRI) for detection of instability and tears of the proximal long head of biceps tendon (LHBT). To assess intraobserver and interobserver agreement.

Materials and methods We performed a retrospective analysis of prospectively collected data of 100 consecutive shoulders who underwent non-contrast 1.5-T MRI prior to arthroscopic surgery due to rotator cuff injury. Images were independently analyzed by two musculoskeletal radiologists. LHBT was evaluated for presence of tearing (intact, longitudinal split, partial-thickness, or full-thickness) and position (normal, subluxated, and dislocated). Anterosuperior rotator cuff tears were also assessed. The reference standard was arthroscopic surgery. The ramp test was performed in order to evaluate LHBT stability. Diagnostic performance measures were determined and Kappa coefficients assessed agreement.

Results Concerning the detection of overall tears, sensitivity ranged from 71 to 73% and specificity was 73%. The specificity for full-thickness tears ranged from 75 to 96%. Overall displacement showed sensitivity ranging from 51 to 58% and specificity ranging from 70 to 86%. The specificity of overall displacement combined with anterosuperior rotator cuff tears ranged from 73 to 91%. Interobserver Kappa values were between 0.59 and 0.69. Intraobserver Kappa values were between 0.74 and 0.82.

Conclusions MRI has moderate accuracy and good agreement for detection of LHBT tears and instability. There is a tendency for increased specificity for full-thickness tears and for instability in the coexistence of anterosuperior rotator cuff tears.

Keywords MRI · Long head of biceps tendon · Tendon tears · Instability · Diagnostic accuracy

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00256-019-03214-z) contains supplementary material, which is available to authorized users.

✉ Eduardo Baptista
baptista100@hotmail.com

Eduardo A. Malavolta
eduardomalavolta@gmail.com

Mauro E. C. Gracitelli
mgracitelli@gmail.com

Daniel Alvarenga
alvarenga.dan@gmail.com

Marcelo Bordalo-Rodrigues
bordalo.m@gmail.com

Arnaldo A. Ferreira Neto
aafneto1@gmail.com

Nestor de Barros
nestor.barros.102@gmail.com

1 Department of Radiology and Oncology, School of Medicine, University of São Paulo (USP), 333 Dr. Ovídio Pires de Campos street, ground floor, Cerqueira César, São Paulo, SP 05403-010, Brazil

2 Department of Orthopedics and Traumatology, School of Medicine, University of São Paulo (USP), 333 Dr. Ovídio Pires de Campos street 3rd floor, Cerqueira César, São Paulo, SP 05403-010, Brazil

Abbreviations

MRI Magnetic resonance imaging
LHBT Long head of biceps tendon
MRA Magnetic resonance arthrography

Introduction

Long head of biceps tendon (LHBT) injuries are a recognized cause of functional impairment and shoulder pain [1, 2]. Proximal LHBT tears and instability within the rotator interval are well-established indications for surgical treatment, isolated or in combination with rotator cuff tears. They are commonly associated with subacromial impingement and rotator cuff disease, especially with supraspinatus tendon tears [3]. Some authors argue that an injured LHBT not addressed during a rotator cuff repair may be a cause of treatment failure and poor outcome [4, 5]. Consequently, the coexistence of tenodesis or tenotomies with rotator cuff repair has increased significantly, reaching 60 to 70% of cases [6].

Clinical examination lacks specificity to detect LHBT abnormalities, particularly with concomitant rotator cuff tears and labral lesions [7]; however, predicting LHBT injuries is desirable because it may be helpful to plan the pain management and contribute to the decision of surgical treatment when conservative measures have failed [8]. This would be especially beneficial for isolated LHBT disease, since these patients would not undergo surgery and, therefore, might not be diagnosed otherwise. Magnetic resonance imaging (MRI) is the method of choice for assessment of rotator cuff tendon injuries and is systematically employed to evaluate LHBT [9]; however, practice shows that recognizing LHBT injuries may be challenging, particularly for partial tears and subtle subluxations [10]. Current literature lacks solid evidence concerning the diagnostic accuracy of MRI. Most published studies are small series, retrospective, and showed heterogeneous methods [3, 11–21]. Some utilize non-contrast MRI and MRA (magnetic resonance arthrography) in the same analysis [3, 12, 21], while others use low magnetic field units [19, 21]. To the best of our knowledge, there is no prospectively collected study to date that focuses on non-contrast 1.5-Tesla (T) unit MRI.

We performed a cross-sectional retrospective analysis of prospectively collected data, comparing MRI findings with arthroscopic shoulder surgery data in consecutive patients who underwent rotator cuff repair, here considered the reference standard. Our purpose was to investigate the accuracy of non-contrast 1.5-T MRI for the detection of LHBT instability and tears. We also sought to determine the diagnostic value of MRI with concomitant medial LHBT displacement and anterosuperior rotator cuff tears to assess instability. Interobserver, intraobserver, and intermethods agreement were evaluated as well. Lastly, we assessed predictor factors that might affect MRI performance in detecting these abnormalities.

Methods

Study design

This cross-sectional diagnostic study was approved by the local institutional ethics review board. Patients were selected consecutively between April 2013 and March 2017 (47 months). All agreed to participate and provided written informed consent. Subjects were included in this study if (1) there was an indication for arthroscopic surgery for rotator cuff repair and (2) had undergone non-contrast MRI before arthroscopic surgery at our institution. Exclusion criteria were: time elapsed between MRI and surgery surpassing 1 year; previous surgical manipulation in the affected shoulder; and insufficient imaging quality (motion artifacts, improper positioning). Patients' age, gender and preoperative functional scores by UCLA [22] and ASES scales [23] were also collected.

Magnetic resonance imaging

Magnetic resonance imaging was performed in our institution. Patients were at supine position with their arms at their side in neutral rotation. A 1.5-T MRI system (GE HDxt, General Electric Medical System, Walchesha, WI) was used with a dedicated shoulder coil. We performed intermediate-weighted and fat-saturated images in the three standard planes and T1 weighted-images in coronal and sagittal planes. Further MRI parameters are detailed in Table 1. Two musculoskeletal radiologists (E.B. and D.A.) with 4 and 8 years of experience, respectively, independently analyzed the images and were blinded of surgical findings and previous imaging reports. A training session on the basis of five cases (not included in the study) was performed in order to standardize image interpretation. Images were analyzed on an image archiving and communication system workstation (isite PACS System, Philips Healthcare, Best, The Netherlands). A second analysis was made by one of the radiologists 4 months after the first assessment.

The evaluation of LHBT integrity was performed in the three standard planes, preferentially on fluid-sensitive sequences [24]. All identified tendons were evaluated for the presence of tendinopathy, regardless of the presence of a tear. Tendinopathy was defined as tendon thickening and/or increase of tendon signal intensity in fluid-sensitive sequences [25]. Tendons were classified as intact, partial-thickness tear, or full-thickness tear, according to Yamaguchi and Bindra [26]. Intact LHBT was defined as a tendon presenting regular

Table 1 MRI parameters

	FSE axial intermediate-weighted	FSE coronal T1-weighted	FSE sagittal T1-weighted	FSE sagittal intermediate-weighted	FSE coronal intermediate-weighted
Fat suppression	Yes	No	No	Yes	Yes
Field of view	15 cm	15 cm	14 cm	15 cm	15 cm
Section thickness	3.0 mm	3.0 mm	4.0 mm	4.0 mm	3.5 mm
Gap	1.0 cm	0.8 cm	0.8 cm	0.8 cm	0.6 cm
Echo time (TE)	42 ms	Minimum	Minimum	45 ms	46 ms
Repetition time (TR)	3,784 ms	517 ms	350 ms	2667 ms	2717 ms
ETL	12	3	4	14	12
Matrix size	288 × 192	288 × 192	288 × 192	288 × 192	256 × 192
Receiver bandwidth	25	15.63	15.63	25	25
Nex	4	2	2	4	4

MRI magnetic resonance imaging, *FSE* fast spin-echo, *ETL* echo train length, *NEX* number of excitations

contours, and signal intensity lower than fluid. Based on clinical practice and interdisciplinary exchange, we proposed to modify Yamaguchi and Bindra classification [26], distributing partial thickness-tears into longitudinal split and partial-thickness tears. A longitudinal split was defined as a tendon delamination with signal intensity similar to liquid and without loss of transverse area. Criteria for partial-thickness tears included focal discontinuity of fibers with the presence of fluid or scar tissue, resulting in a reduction of tendon caliber and surface irregularity. These tears were graded subjectively, as reaching more or less than 50% of LHBT thickness. Full-thickness tears were characterized as lack of identification or complete discontinuation of LHBT fibers. Once a longitudinal split or a partial-thickness tear was recognized, observers indicated one or more sites of involvement as follows—origin, articular course, deflexional course, bicipital groove. The position of the LHBT was evaluated as well, except when a full-thickness tear was present. The upper slice encompassing the lesser tuberosity was the reference in this assessment. Partial contact loss between LHBT and bicipital groove and/or the presence of the perched sign was defined as subluxation [27, 28]. Complete contact loss between LHBT and bicipital groove was defined as dislocation [27]. Anterosuperior rotator cuff tears were also accessed and classified as intact or torn [29]. The subscapularis tendon was assessed exclusively in the upper two slices in the axial plane. The supraspinatus tendon was assessed exclusively in the most anterior slice in the coronal plane. Criteria for supraspinatus and subscapularis tendons tears included discontinuity of fibers with hyperintense signal within the tendon corresponding to fluid.

Reference standard

Arthroscopic surgery is the orthopedic surgeon's first choice to approaching rotator cuff repair and was considered the reference standard. Patients were operated in beach chair positioning under brachial plexus block and general anesthesia. Posterior, anterior, and lateral portals were performed. Diagnostic inspection is routine of the surgical procedure; therefore, it did not result in any additional morbidity. Initially, a static inspection was performed through the posterior portal with a 30° scope view. A dynamic inspection followed in which LHBT was pulled through a probe inserted in an anterior portal, providing a more comprehensive view [30]; then, the ramp test was performed in order to evaluate LHBT stability [31]. The ramp test is a dynamic arthroscopic technique performed routinely in our institution to test the integrity of soft tissue restraint. In this test, the tendon in pulled down repeatedly by a probe inserted in anterior arthroscopic portal. The LHBT remains in the bicipital groove and assumes a V-shape morphology when restraints of soft tissue pulley are intact. When soft tissue restraints are damaged—that is, LHBT is unstable—tendons may undergo medial displacement and/or assume a U-shaped morphology.

The same three orthopedic surgeons with 10, 11, and 12 years of experience performed all arthroscopies, all having specialized training in shoulder surgery. They were aware of routine radiology reports but blinded for the study analysis. LHBT was classified as intact or torn (longitudinal split; partial-thickness tear under or above 50%; full-thickness) and also as stable, unstable or dislocated, except when a full-thickness tear was found.

Sample size estimation

We performed the sample size estimation using the likelihood ratio test, based on data obtained by Zanetti et al. [19], targeting a statistical power of 80% with a significance level of 5%. A sample of at least 86 shoulders was determined. With the consideration for sample loss, we decided to expand the sample to 100 shoulders.

Statistical analysis

Continuous variables were presented with means and standard deviations, while categorical variables were presented with absolute and percentage values. MRI performance for detection of overall tears, full-thickness tears, and instability was calculated by means of sensitivity, specificity, predictive values, and accuracy, with 95% confidence intervals. Instability accuracy was calculated in the following settings: dislocations or subluxations alone, overall medial displacement (dislocation or subluxation), anterosuperior rotator cuff tendon tears depicted on MRI and combined settings (medial displacement and anterosuperior rotator cuff tendon tears simultaneously).

Interobserver and intraobserver agreement were determined by Kappa statistics (simple Kappa for dichotomous variables and weighted Kappa for multiclass variables with a weight of one per differential class). The agreement was calculated in five circumstances: overall tears, simplified tear classification (intact, overall partial-thickness tears, full-thickness tears), complete tear classification (modified Yamaguchi and Bindra classification), position (centered, subluxated/unstable, dislocated) and instability (stable versus unstable). We also calculated Kappa coefficients between imaging readers and intraoperative findings. The strength of agreement was classified as follows: 0–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost perfect [32].

Several variables were tested as predictors of MRI accuracy for tears and instability detection. For this calculation, the shoulders were divided into two groups: correctly evaluated (true positives and true negatives) and incorrectly evaluated (false positives and false negatives). Bivariate logistic regression was used in this assessment, complemented by chi-square, Mann-Whitney, and likelihood ratio tests. Age, functional score, tear site, surgical instability, and anterosuperior rotator cuff tears were tested regarding LHBT tears. Age, functional score, LHBT partial-thickness tears, and anterosuperior rotator cuff tears were tested regarding LHBT instability. Lastly, chi-square statistic was used to test the association between LHBT partial-thickness tears and instability. IBM-SPSS Statistics for Windows (version 20.0, Armonk, NY) software was used to perform the analysis. All tests were performed with a significance level of 5%.

Results

The study group comprised 100 shoulders of 98 patients from 38 males (38%) and 62 females (62%). Mean age of subjects was 56.4 ± 8.5 years. There were 65 right-sided shoulders (65%) and 35 left-sided shoulders (35%). Regarding functional scores, mean ASES was 38.7 ± 19.1, while mean UCLA was 13.7 ± 4.6. Mean time interval between MRI and surgery was 172 days ±94.

Observer 1 found 52 intact and 48 torn LHBT. The LHBT position was evaluated in 90 cases, of which 49 were considered centered in the bicipital groove, while 33 were subluxated and 8 were dislocated. Observer 2 found 53 intact and 47 torn LHBT. The LHBT position was evaluated in 86 cases, of which 57 were centered in the bicipital groove, while 21 were subluxated and 8 were dislocated. In relation to LHBT tears, observers 1 and 2 designated 32 and 33 true positives (Fig. 1); 40 and 40 true negatives; 15 and 15 false positives; 13 and 12 false negatives (Fig. 2), respectively. In relation to LHBT instability, observers 1 and 2 designated 26 and 22 true positives (Fig. 3); 31 and 38 true negatives; 13 and 7 false positives; 18 and 20 false negatives (Fig. 4), respectively.

Of the 100 shoulders, there were 55 intact and 45 torn LHBT at arthroscopy. Among these tears, 3 were described as longitudinal split; 19 as partial-thickness tears affecting less than 50% of thickness; 14 as partial-thickness tears affecting

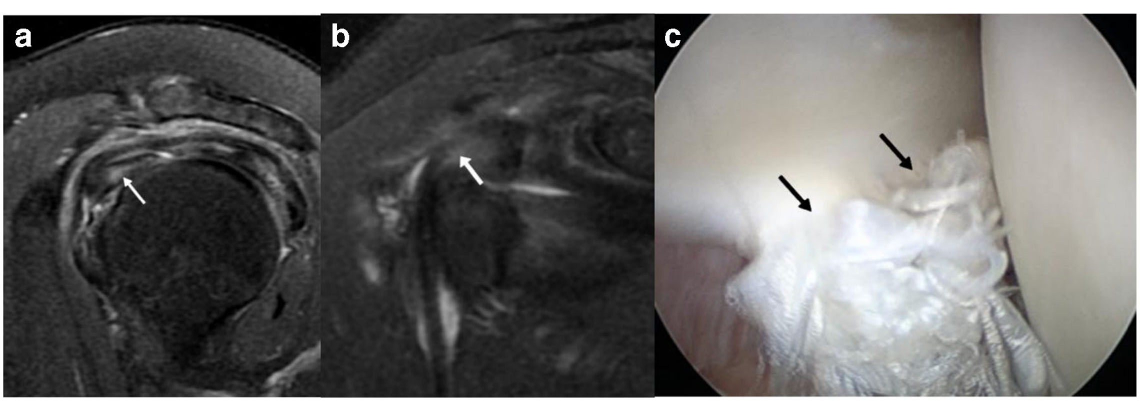

Fig. 1 57-year-old woman presenting with true positive MRI findings for both readers. Sagittal (**a**) and coronal (**b**) fast-spin-echo intermediate-weighted and fat-suppressed images show irregular tendon surfaces and increased signal intensity, equivalent to fluid (white arrows), consistent with a partial-thickness LHBT tear, judged by readers as involving less than 50% of tendon thickness. Posterior viewing arthroscopic portal evaluation (**c**) in agreement with the radiologists (black arrows)

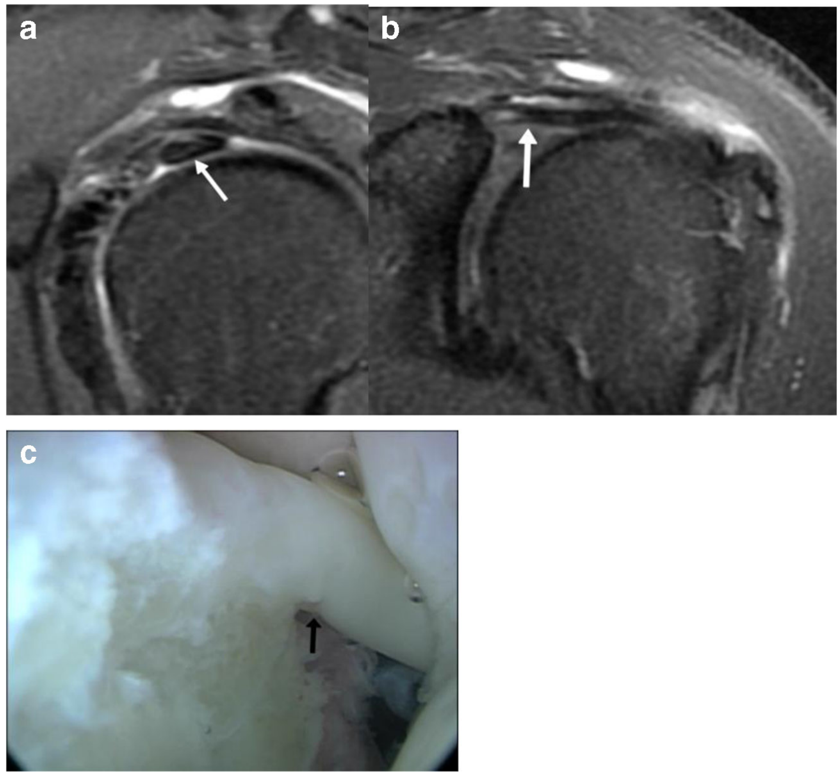

Fig. 2 False negative by both readers in the right shoulder of a 58-year-old man with partial-thickness LHBT tear. Sagittal and coronal fast-spin echo intermediate-weighted and fat-suppressed images (**a** and **b**) demonstrating slight signal changes, consistent with tendinopathy (white arrows). No tear was identified by either of the MRI readers. Arthroscopic image through the posterior view (**c**) revealed a small partial-thickness tear, reaching less than 50% of tendon thickness (black arrow)

more than 50% of thickness; and 9 were full-thickness tears. All the full-thickness LHBT tears were retracted. Further details regarding the distribution of LHBT tears for each observer and their correlation with surgery data are presented in Table 2. No labral lesion extended to the LHBT anchor. One patient presented with a bifid tendon that was considered an anatomic variant during surgery and was correctly diagnosed as intact by both MRI observers. Among the 91 LHBT in which instability was tested, 45 (50.5%) were stable and 46 (49.5%) unstable, of which 16 (17.6%) were dislocated and 30 (33%) demonstrated instability during the ramp test. Among the 46 unstable LHBT, 46 presented with medial bicipital pulley lesion and 37 with lateral bicipital pulley lesion. Thus, of the 100 shoulders, 25 were treated with tenodesis, 34 treated with tenotomies, and 41 were not treated. The most frequent indications for LHBT tenodesis or tenotomy were tearing and instability; however, five tenotomies were indicated by tendinopathy.

Regarding depiction of overall LHBT tears, sensitivity ranged from 71 to 73%, specificity was 73% for both observers, and accuracy ranged from 72 to 73%. When evaluating full-thickness tears, sensitivity ranged from 56 to 67%, specificity ranged from 75 to 96%, and accuracy from 68 to 88%. Further details are given in Table 3.

Long head of biceps tendon dislocation alone showed sensitivity ranging from 30 to 31%, specificity of 100%, and

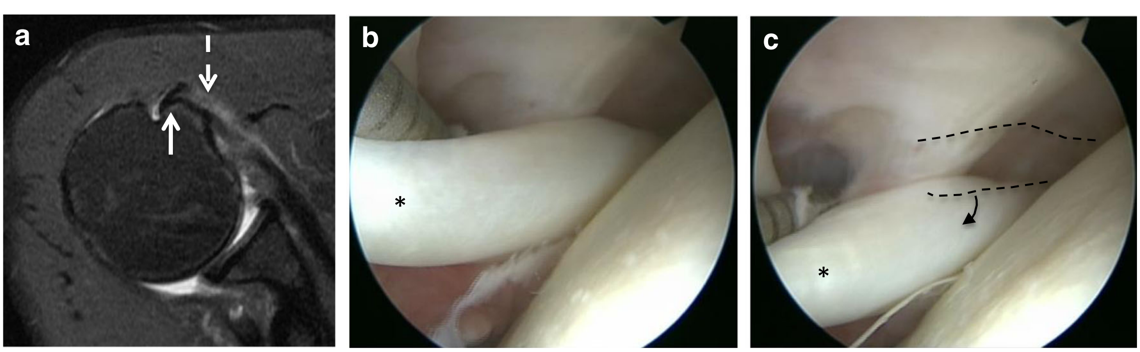

Fig. 3 61-year-old woman with true positive MRI findings for LHBT instability. Intermediate-weighed axial sequence (**a**) showing medial subluxation of LHBT (white arrow) and partial-thickness tear of the SScT (dotted white arrow). Photographs of arthroscopy, before (**b**) and after (**c**) the ramp test depict the subluxated tendon (*). The position of LHBT before the maneuver is represented by the dotted lines and its displacement by the curved arrow

Fig. 4 68-year-old man with false negative findings at MRI. Axial intermediate-weighed image (**a**) showing the LHBT (white arrow) centered in the bicipital groove. Photographs of arthroscopy, before (**b**) and after (**c**) the ramp test reveal the tendon (*) underwent medial subluxation after the ramp test. The position before the maneuver is represented by white-dotted lines and the direction of displacement by the curved arrow. After the maneuver, the tendon presented "U" appearance (black-dotted line)

accuracy ranging from 68 to 70% for detection of LHBT instability. Isolated subluxation showed sensitivity ranging from 39 to 49%, specificity ranging from 70 to 86%, and accuracy ranging from 61 to 66%. Overall medial displacement obtained sensitivity between 51 to 58%, specificity between 70 to 86%, and accuracy between 64 to 69%. The association of overall LHBT displacement with supraspinatus and subscapularis tendons tears led to sensitivities ranging from 17 to 48% and specificity ranging from 73 to 91%. Further details are given in Table 4.

Interobserver reliability for LHBT tears detection was substantial in all settings. Kappa coefficients achieved were 0.62 for overall LHBT tears, 0.65 for simplified LHBT tears classification, and 0.65 when considered complete LHBT tears classification. Intraobserver correlation showed to be almost perfect for detection of overall LHBT tears (Kappa coefficient = 0.82) and for simplified LHBT tears classification (weighted Kappa coefficient = 0.82), while it was substantial when considering complete LHBT tears classification (weighted Kappa coefficient = 0.74). Agreement between methods (arthroscopy and MRI) for LHBT tears was reasonable to moderate (weighted Kappa coefficients ranging from 0.39 to 0.48). Regarding the presence of LHBT instability, the Kappa coefficient was 0.59 for interobserver agreement and 0.76 for intraobserver agreement, whereas for LHBT position, Kappa coefficient obtained was 0.69 for interobserver correlation and 0.8 for intraobserver agreement. The correlation between methods ranged between 0.35 and 0.43 (further details are provided in Table 5).

No predictors of MRI accuracy were identified in the assessment of LHBT tears or instability (Tables 6 and 7). There was significant correlation between LHBT partial-thickness tears and instability ($p < 0.05$). Of the 36 tendons with surgically proven partial-thickness tears, 27 (75%) were unstable; likewise, of the 46 unstable LHBT, 27 (59%) exhibited concomitant partial-thickness tears.

Table 2 Correlation between MRI observers and arthroscopic surgery for intact and torn LHBT

Observer 1	Surgery					
	Intact	Longitudinal split	Partial-thickness tear <50%	Partial-thickness tear >50%	Full-thickness tear	Total
Intact	40 (40)	1 (1)	7 (7)	5 (5)	0 (0)	53 (53)
Longitudinal split	3 (3)	0 (0)	5 (5)	3 (3)	0 (0)	11 (11)
Partial-thickness tear <50%	3 (3)	2 (2)	3 (3)	2 (2)	1 (1)	11 (11)
Partial-thickness tear >50%	6 (6)	0 (0)	3 (3)	4 (4)	2 (2)	15 (15)
Full-thickness tear	3 (3)	0 (0)	1 (1)	0 (0)	6 (6)	10 (10)
Observer 2						
Intact	40 (40)	1 (1)	7 (7)	4 (4)	0 (0)	52 (52)
Longitudinal split	7 (7)	1 (1)	3 (3)	4 (4)	0 (0)	15 (15)
Partial-thickness tear <50%	1 (1)	1 (1)	1 (1)	2 (2)	1 (1)	6 (6)
Partial-thickness tear >50%	4 (4)	0 (0)	4 (4)	2 (2)	3 (3)	13 (13)
Full-thickness tear	3 (3)	0 (0)	4 (4)	2 (2)	5 (5)	14 (14)
Total	55 (55)	3 (3)	19 (19)	14 (14)	9 (9)	100 (100)

Data are absolute numbers with percentages *in parenthesis*

Table 3 Diagnostic performance of MRI for detection of LHBT tears

	Accuracy	Sensitivity	Specificity	PPV	NPV
Overall tears					
Reader 1	72% (63; 81)	71% (56; 84)	73% (59; 84)	68% (53; 81)	75% (62; 86)
Reader 2	73% (64; 82)	73% (58; 85)	73% (59; 84)	69% (54; 81)	77% (63; 87)
Full-thickness tears					
Reader 1	88% (76; 99)	67% (30; 92)	96% (78; 100)	86% (42; 100)	88% (69; 97)
Reader 2	70% (54; 85)	56% (21; 86)	75% (53; 90)	45% (17; 77)	82% (60; 95)

Data are percentages with 95% confidence intervals in parenthesis. *MRI* magnetic resonance imaging, *LHBT* long head of biceps tendon, *PPV* positive predictive value, *NPV* negative predictive value

Discussion

Long head of biceps tendon instability and tears are causes of shoulder pain and functional impairment. An accurate imaging diagnosis is desirable because it can change therapeutic strategy and contribute to surgical treatment decisions; however, the role of MRI remains unclear. To the best of our knowledge, this is the first prospective study focused on the LHBT in which MRI showed to have moderate accuracy for the detection of overall tears with balanced sensitivity and specificity. The specificity and negative predictive value for full-thickness LHBT tears were substantially higher for one of the observers. Although dislocation has shown to be the most reliable signal for LHBT instability, with specificity and negative predictive value of 100%, its sensitivity was poor. Overall displacement achieved somewhat better sensitivities, but at the expense of reduced specificity. The combined evaluation of LHBT displacement with anterosuperior rotator cuff tendons tears led to higher specificities (above 90% in some settings), which may enhance the radiologist's confidence in this diagnosis. Intraobserver agreement was substantial or almost perfect. Interobserver agreement was substantial in most of the measures, consistent with data previously reported [11, 12, 17].

Several issues may explain the moderate accuracy of MRI for LHBT tears. First, the conventional shoulder MRI protocol is designed to evaluate rotator cuff disease, particularly supraspinatus tendon. Consequently, LHBT is obliquely viewed at each slice, requiring a multiplanar viewing and effort-consuming evaluation. Second, limited spatial resolution may induce partial volume artifacts and obscure small partial-thickness tears. Third, signal intensity may lead to

Table 4 Diagnostic performance of MRI for detection of LHBT instability

	Accuracy	Sensibility	Specificity	PPV	NPV
Observer 1					
Dislocation	68% (56; 81)	31% (14; 52)	100% (89; 100)	100% (63, 100)	63% (48, 77)
Subluxation	61% (50; 72)	49% (31; 66)	71% (55; 83)	57% (37; 75)	63% (48; 77)
Overall LHBT displacement	64% (54; 74)	58% (42; 73)	70% (55; 83)	66% (49; 80)	63% (48; 77)
SScT tear	54% (44; 64)	74% (59; 86)	33% (20; 49)	53% (40; 66)	56% (35; 75)
SST tear	49% (39; 60)	80% (66; 91)	18% (8; 32)	50% (38; 62)	47% (23; 72)
LHBT displacement + SScT tear	62% (52; 72)	46% (31; 62)	77% (62; 89)	67% (47; 83)	60% (46; 72)
LHBT displacement + SST tear	60% (50; 71)	48% (33;2 63)	73% (57; 85)	64% (45;80)	58% (44; 71)
LHBT displacement + SScT and SST tears	58% (48; 68)	36% (22; 52)	80% (65; 90)	64% (43; 82)	56% (43; 68)
Observer 2					
Dislocation	70% (59; 82)	30% (14; 50)	100% (91; 100)	100% (63; 100)	66% (52; 78)
Subluxation	66% (55; 77)	39% (22 58)	86% (72; 95)	67% (41; 87)	66% (52; 78)
Overall LHBT displacement	70% (60; 80)	51% (35; 68)	86% (72; 95)	77% (56; 91)	66% (52; 78)
SScT tear	54% (44; 64)	63% (48; 77)	44% (30; 60)	54% (40; 67)	54% (37; 71)
SST tear	49% (39; 60)	74% (59; 86)	24% (13; 40)	50% (38; 62)	48% (27; 69)
LHBT displacement + SScT tear	61% (50; 72)	28% (15; 45)	91% (78; 97)	73% (45; 92)	58% (46; 70)
LHBT displacement + SST tear	60% (50; 71)	33% (19; 49)	86% (72; 95)	68% (43; 87)	58% (45; 70)
LHBT displacement + SScT and SST tears	55% (48; 66)	18% (7; 33)	91% (78; 97)	64% (31; 89)	54% (42; 66)

Data are percentages with 95% confidence intervals in parenthesis. *MRI* magnetic resonance imaging, *LHBT* long head of biceps tendon, *PPV* positive predictive value, *NPV* negative predictive value, *SScT* subscapularis tendon, *SST* supraspinatus tendon

Table 5 Interobserver, intraobserver, and inter-methods agreement

	Interobserver	Intraobserver	Intermethods observer 1	Intermethods observer 2
Overall LHBT tears	0.62 (0.46; 0.77)	0.82 (0.71; 0.93)	–	–
Simplified LHBT tear	0.65 (0.52; 0.78)	0.82 (0.72; 0.91)	0.48 (0.31; 0.64)	0.43 (0.72; 0.91)
Full LHBT tear classification	0.65 (0.53; 0.77)	0.74 (0.64; 0.85)	0.39 (0.23; 0.54)	0.40 (0.28; 0.58)
LHBT position	0.69 (0.54; 0.84)	0.80 (0.70; 0.91)	0.35 (0.17; 0.52)	0.43 (0.26; 0.60)
LHBT instability	0.59 (0.42; 0.77)	0.76 (0.63; 0.89)	–	–

Data are Kappa or weighted Kappa coefficients, with 95% confidence intervals in parenthesis. *LHBT* long head of biceps tendon

doubt, since a mildly increased signal on fluid-sensitive sequences may be judged as tendinopathy or partial-thickness tear. The limited sensitivity of the MRI for detection of LHBT instability seems reasonable to us, since a dynamic condition is being evaluated by a static method. As already demonstrated in the literature [33], an unstable LHBT may be centered in the bicipital groove, limiting sensibility of MRI.

Literature is contradictory in regard to MRI's sensitivity concerning LHBT tears, varying widely from 7 to 100% [3, 12, 13, 15, 17, 19, 21], making comparison troublesome. Among noncontrast MRI; however, our sensitivity was similar to those reported by Malavolta et al. [20] and superior to Mohtadi et al. [16] and Razjmou et al. [14], which ranged from 0 to 57%. Our specificity values are consistent with those reported by Zanetti et al. [19] and Tadrus et al. [17], although modest compared to other authors [2, 17–19, 22, 24–26, 35]. Razmjou et al. [14] and De Maeseneer et al. [13], for example, reported high specificities, between 86 and 96% for partial-thickness LHBT tears; however, corresponding sensitivities were low, between 25 and 50% [13, 14]. This observation might be related to heterogeneity of diagnostic criteria and it may suggest an intrinsic limitation of MRI accuracy, regardless of the criteria adopted. Few studies have evaluated MRI accuracy for the detection of LHBT instability. Our specificity was consistent with the literature, which ranges from 72 to 97%. Although our sensitivity is in accordance with the findings of Baumann et al. [34] and Malavolta et al. [20], some authors found higher values, ranging from 73 to 100% [14, 35, 36].

We do believe that the discrepancy of diagnostic performance of this study compared to some series may be justified due to methodological issues. Retrospective articles predominate in the literature, based on radiological and/or arthroscopic reports [11–13, 15, 17–21, 24, 37, 38]. An unsure radiologist or surgeon may not express his doubts and be elusive in the report. It leads to the exclusion of most challenging cases, resulting in bias and overestimating diagnostic measures. Besides, once the orthopedic surgeon identifies a partial LHBT tear, it is followed by tenotomy or tenodesis, thereby skipping instability tests. Consequently, previous series may have underestimated intraoperative LHBT instability prevalence and miscalculate diagnostic performance.

Table 6 Predictor factors for MRI performance in detecting LHBT tears

Predictors	OR (CI) observer 1	*p*	OR (CI) observer 2	*p*
SScT tear	0.75 (0.29; 1.95)	0.557	1.02 (0.4; 2.61)	0.966
SSP tear	1.26 (0.5; 3.17)	0.620	0.67 (0.25; 1.80)	0.428
LHBT tear site		0.245[b]		0.186[b]
Origin[c]			1.00	
Articular course	1.00		0.41 (0.11; 1.62)	
Deflexional course	0.50 (0.12; 2.15)		0.13 (0.1; 1.76)	
Bicipital groove[c]	0.63 (0.09; 4.53)			
LHBT instability	0.7 (0.28; 1.76)	0.446	2.22 (0.79; 6.28)	0.127
Age	0.99 (0.94; 1.04)	0.727[a]	0.99 (0.94; 1.05)	0.822[a]
ASES	0.99 (0.97; 1.02)	0.899[a]	0.99 (0.97; 1.02)	0.879[a]
UCLA	1.0 (0.9; 1.11)	0.701[a]	0.99 (0.87; 1.09)	0.814[a]

OR odds ratio, *CI* confidence intervals, *LHBT* long head of biceps tendon, *SScT* subscapularis tendon, *SST* supraspinatus tendon. Otherwise indicated, *p*-value was obtained with chi-squared test

[a] Mann-Whitney test

[b] Likelihood ratio test

[c] Impossible to calculate

Table 7 Predictor factors for MRI performance in detecting LHBT instability

Predictors	OR (CI) observer 1	*p*	OR (CI) observer 2	*p*
SScT tear	0.6 (0.23; 1.57)	0.293	0.78 (0.29; 2.09)	0.623
SSP tear	0.54 (0.2; 1.47)	0.224	0.65 (0.24; 1.79)	0.402
LHBT tear	0.47 (0.19; 1.15)	0.094	0.53 (0.2; 1.4)	0.307[b]
ASES	1.0 (0.98; 1.03)	0.882[a]	0.99 (0.97; 1.02)	0.782
UCLA	1.06 (0.95; 1.18)	0.451[a]	0.98 (0.89; 1.09)	0.939

CI confidence intervals, *OR* odds ratio, *LHBT* long head of biceps tendon, *SScT* subscapularis tendon, *SST* supraspinatus tendon. Otherwise indicated *p*-value obtained with chi-squared test

[a] Mann-Whitney test

[b] Likelihood ratio test

Moreover, the concept of LHBT instability has evolved over the past two decades and is no longer limited to static subluxations and dislocations described by Walch et al. [27], but rather it is becoming a dynamic condition. Walch et al. [27] and Bennet [39] stressed that this entity may be overlooked during surgery. Motley et al. [31] developed the ramp test, a dynamic arthroscopic maneuver that refines the detection of LHBT instability at arthroscopy. Previous articles such as published by Spritzer et al. [28], in 2001, had not yet assimilated these concepts. Even more recent studies like those of Kang et al. [35] and Razmjou et al. [14], did not mention any specific maneuvers to explore LHBT instability. We believe these authors did not depict mild LHBT instabilities and probably overestimated MRI sensitivity.

Additionally, technical issues may also be responsible for the observed differences. Many authors employed MRA [12, 13, 15, 19] and 3-T units, which have proven to increase detection of rotator cuff tears [40, 41]. It might have the same effect on the depiction of LHBT tears, although a recent study has shown no difference between MRI and MRA for detection of LHBT tendinopathy and tears [17]. Prospective research comparing 1.5-T and 3.0-T magnets for LHBT tears are needed to clarify this issue. Finally, there is considerable disparity about echo time (TE) utilized on fluid-sensitive sequences. Although intermediate-weighted sequences predominate, some authors perform at least one T2-weighted image with TE above 80 ms [13–15, 17]. An elevated TE may increase the conspicuity between tendon and fluid and reduce the magic angle artifact, while conversely, lower TE values improve the signal-to-noise ratio.

We did not observe predictors of MRI accuracy to detect LHBT injuries, unlike Razmjou et al. [14] who reported lower accuracies when massive rotator cuff tears were seen. This can be explained by the fact that we limited this evaluation to anterosuperior tears, while Razmjou et al. [14] found this result for rotator cuff tears surpassing 3 cm of extension. Besides, we did not observe the effect of age on MRI accuracy, as opposed to Borrero et al. [12] who reported lower accuracy for the middle-aged group; however, in agreement with Borrero et al. [12], we noted a significant correlation between partial LHBT tears and instability. Although this association does not establish a causal relationship, it reinforces the hypothesis of a pathophysiological link between these abnormalities. It is possible that the abrasion secondary to instability may have led to partial LHBT tears, while partial LHBT tears likewise may have injured the primary stabilizers, resulting in instability.

Limitations and conclusions

This study has several limitations. Firstly, all patients in our sample had a surgical indication for rotator cuff repair, which may create a selection bias. However, our results are more representative of clinical practice, since LHBT degeneration occurs more often in this population. Secondly, orthopedic surgeons were aware of routine MRI reports at the time of the surgeries, while thirdly, radiologists were aware of the study's aim and that patients had indication of surgical treatment for rotator cuff repair. Fourthly, there are limitations inherent to arthroscopic surgery, such as moderate interobserver agreement and limitations in detecting intratendinous or extra-articular tears [42, 43]. However, since this study focused on proximal LHBT tears, limitations to evaluate distal extra-articular LHBT tears were not a concern. Fifthly, time elapsed between MRI and arthroscopy may have contributed to underestimating the detection of tears, although the 1-year limit is considered acceptable in the literature and has been adopted by a systematic review of rotator cuff tears [44].

In conclusion, we have shown that MRI has moderate accuracy and good agreement for detection of LHBT tears and instability. LHBT dislocation is the most reliable sign of instability. Radiologists may be more confident to diagnose full-thickness tears and instability in the coexistence of subluxation and anterosuperior rotator cuff tears; therefore, preoperative imaging may be useful to guide treatment strategy, as it gives the opportunity of conservative treatment with cortisone injections and it alerts the surgeon about the possibility of LHBT abnormalities. Besides, preoperative diagnosis may contribute to the surgical decision in case of failed pain management. It also may be valuable to surgical planning, since a choice must be made among tenotomy and several different techniques of tenodesis; however, careful surgical inspection remains mandatory, since false positives and negatives may occur. Further work should focus on improving the sensitivity to detect LHBT instability, in which dynamic imaging may play a role.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Watson ST, Robbins CB, Bedi A, Carpenter JE, Gagnier JJ, Miller BS. Comparison of outcomes 1 year after rotator cuff repair with and without concomitant biceps surgery. Arthroscopy. 2017;33:1928–36.
2. Alpantaki K. Sympathetic and sensory neural elements in the tendon of the long head of the biceps. J Bone Joint Surg. 2005;87:1580–3.
3. Beall DP, Williamson EE, Ly JQ, Adkins MC, Emery RL, Jones TP, et al. Association of biceps tendon tears with rotator cuff abnormalities: degree of correlation with tears of the anterior and superior portions of the rotator cuff. AJR Am J Roentgenol. 2003;180:633–9.
4. Williams GR. Painful shoulder after rotator cuff surgery. J Am Acad Orthop Surg. 1997;5:97–108.
5. Ahrens PM, Boileau P. The long head of biceps and associated tendinopathy. J Bone Joint Surg (Br). 2007;89:1001–9.
6. Gilmer BB, DeMers AM, Guerrero D, Reid JB, Lubowitz JH, Guttmann D. Arthroscopic versus open comparison of long head of biceps tendon visualization and pathology in patients requiring tenodesis. Arthroscopy. 2015;31:29–34.
7. Holtby R, Razmjou H. Accuracy of the Speed's and Yergason's tests in detecting biceps pathology and SLAP lesions: comparison with arthroscopic findings. Arthroscopy. 2004;20:231–6.
8. Creech MJ, Yeung M, Denkers M, Simunovic N, Athwal GS, Ayeni OR. Surgical indications for long head biceps tenodesis: a systematic review. Knee Surg Sports Traumatol Arthrosc. 2016;24:2156–66.
9. Morag Y, Bedi A, Jamadar DA. The rotator interval and long head biceps tendon: anatomy, function, pathology, and magnetic resonance imaging. Magn Reson Imaging Clin N Am. 2012;20:229–59.
10. Tuckman GA. Abnormalities of the long head of the biceps tendon of the shoulder: MR imaging findings. AJR Am J Roentgenol. 1994;163:1183–8.
11. Lee RW, Choi SJ, Lee MH, Ahn JH, Shin DR, Kang CH, et al. Diagnostic accuracy of 3T conventional shoulder MRI in the detection of the long head of the biceps tendon tears associated with rotator cuff tendon tears. Skelet Radiol. 2016;45:1705–15.
12. Borrero CG, Costello J, Bertolet M, Vyas D. Effect of patient age on accuracy of primary MRI signs of long head of biceps tearing and instability in the shoulder: an MRI-arthroscopy correlation study. Skelet Radiol. 2018;47:203–14.
13. De Maeseneer M, Boulet C, Pouliart N, Kichouh M, Buls N, Verhelle F, et al. Assessment of the long head of the biceps tendon of the shoulder with 3T magnetic resonance arthrography and CT arthrography. Eur J Radiol. 2012;81:934–9.
14. Razmjou H, Fournier-Gosselin S, Christakis M, Pennings A, ElMaraghy A, Holtby R. Accuracy of magnetic resonance imaging in detecting biceps pathology in patients with rotator cuff disorders: comparison with arthroscopy. J Shoulder Elb Surg. 2016;25:38–44.
15. Omoumi P, Bafort A-C, Dubuc J-E, Malghem J, Vande Berg BC, Lecouvet FE. Evaluation of rotator cuff tendon tears: comparison of multidetector CT arthrography and 1.5-T MR arthrography. Radiology. 2012;264:812–22.
16. Mohtadi NG, Vellet AD, Clark ML, Hollinshead RM, Sasyniuk TM, Fick GH, et al. A prospective, double-blind comparison of magnetic resonance imaging and arthroscopy in the evaluation of patients presenting with shoulder pain. J Shoulder Elb Surg. 2004;13:258–65.
17. Tadros AS, Huang BK, Wymore L, Hoenecke H, Fronek J, Chang EY. Long head of the biceps brachii tendon: unenhanced MRI versus direct MR arthrography. Skelet Radiol. 2015;44:1263–72.
18. Dubrow SA, Streit JJ, Shishani Y, Robbin MR, Gobezie R. Diagnostic accuracy in detecting tears in the proximal biceps tendon using standard nonenhancing shoulder MRI. Open Access J Sports Med. 2014;5:81–7.
19. Zanetti M, Weishaupt D, Gerber C, Hodler J. Tendinopathy and rupture of the tendon of the long head of the biceps brachii muscle: evaluation with MR arthrography. AJR Am J Roentgenol. 1998;170:1557–61.
20. Malavolta EA, Assunção JH, Guglielmetti CLB, de Souza FF, Gracitelli MEC, Ferreira Neto AA. Accuracy of preoperative MRI in the diagnosis of disorders of the long head of the biceps tendon. Eur J Radiol. 2015;84:2250–4.
21. Houtz CG, Schwartzberg RS, Barry JA, Reuss BL, Papa L. Shoulder MRI accuracy in the community setting. J Shoulder Elb Surg. 2011;20:537–42.
22. Oku EC, Andrade AP, Stadiniky SP, Carrera EF, Tellini GG. Tradução e adaptação cultural do modified-University of California at Los Angeles Shoulder Rating Scale para a língua Portuguesa. Rev Bras Reumatol. 2006;46:246–52.
23. Knaut L, Moser AD, Melo SDA, Richards RR. Tradução e adaptação cultural à língua portuguesa do American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form (ASES) para avaliação da função do ombro. Rev Bras Psiquiatr. 2010;50:176–83.
24. Gaskin CM, Anderson MW, Choudhri A, Diduch DR. Focal partial tears of the long head of the biceps brachii tendon at the entrance to the bicipital groove: MR imaging findings, surgical correlation, and clinical significance. Skelet Radiol. 2009;38:959–65.
25. Buck FM, Grehn H, Hilbe M, Pfirrmann CWA, Manzanell S, Hodler J. Degeneration of the long biceps tendon: comparison of MRI with gross anatomy and histology. AJR Am J Roentgenol. 2009;193:1367–75.
26. Yamaguchi K, Bindra R. Disorders of the biceps tendon. In: Iannott JP, Willians G, Miniaci A, Zuckerman J, editors. Disorders of the shoulder: diagnosis and management. 1st ed. Philadelphia: Lippincott Willians and Wilkins; 1999.
27. Walch G, Nove-Josserand L, Levigne C, Renaud E. Tears of the supraspinatus tendon associated with "hidden" lesions of the rotator interval. J Shoulder Elb Surg. 1994;3:353–60.
28. Spritzer CE, Collins AJ, Cooperman A, Speer KP. Assessment of instability of the long head of the biceps tendon by MRI. Skelet Radiol. 2001;30:199–207.
29. Habermeyer P. Rotatorenintervall und lange bizepssehne. Rotatorenmanschette. Munchen: Jena, Urban and Ficher; 2002. p. 333–74.
30. Festa A, Allert J, Issa K, Tasto JP, Myer JJ. Visualization of the extra-articular portion of the long head of the biceps tendon during intra-articular shoulder arthroscopy. Arthroscopy. 2014;30:1413–7.
31. Motley GS, Osbahr DC, Holovac TF, Speer KP. An arthroscopic technique for confirming intra-articular subluxation of the long head of the biceps tendon: the ramp test. Arthroscopy. 2002;18:1–5.
32. Landis JRL, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74.
33. Braun S, Horan MP, Elser F, Millett PJ. Lesions of the biceps pulley. Am J Sports Med. 2011;39:790–5.
34. Baumann B, Genning K, Böhm D, Rolf O, Gohlke F. Arthroscopic prevalence of pulley lesions in 1007 consecutive patients. J Shoulder Elb Surg. 2008;17:14–20.

35. Kang Y, Lee JW, Ahn JM, Lee E, Kang HS. Instability of the long head of the biceps tendon in patients with rotator cuff tear: evaluation on magnetic resonance arthrography of the shoulder with arthroscopic correlation. Skelet Radiol. 2017;46:1335.
36. Weishaupt D, Zanetti M, Tanner A, Gerber C, Hodler J. Lesions of the reflection pulley of the long biceps tendon. MR arthrographic findings. Investig Radiol. 1999;34:1–10.
37. Jung JY, Yoon YC, Yi SK, Yoo J, Choe BK. Comparison study of indirect MR arthrography and direct MR arthrography of the shoulder. Skelet Radiol. 2009;38:659–67.
38. Skendzel JG, Jacobson JA, Carpenter JE, Miller BS. Long head of biceps brachii tendon evaluation: accuracy of preoperative ultrasound. AJR Am J Roentgenol. 2011;197:942–8.
39. Bennett WF. Visualization of the anatomy of the rotator interval and bicipital sheath. Arthroscopy. 2001;17:107–11.
40. Smith TO, Daniell H, Geere J, Toms AP, Hing CB. The diagnostic accuracy of MRI for the detection of partial- and full-thickness rotator cuff tears in adults. Magn Reson Imaging. 2012;30:336–46.
41. De Jesus JO, Parker L, Frangos AJ, Nazarian LN. Accuracy of MRI, MR arthrography, and ultrasound in the diagnosis of rotator cuff tears: a meta-analysis. AJR Am J Roentgenol. 2009;192:1701–7.
42. Taylor SA, Khair MM, Gulotta LV, Pearle AD, Baret NJ, Newman AM, et al. Diagnostic glenohumeral arthroscopy fails to fully evaluate the biceps-labral complex. Arthroscopy. 2015;31:215–24.
43. Kuhn JE, Dunn WR, Ma B, Wright RW, Jones G, Spencer EE, et al. Interobserver agreement in the classification of rotator cuff tears. Am J Sports Med. 2007;35:437–41.
44. Lenza M, Buchbinder R, Takwoingi Y, Rv J, Nca H, Faloppa F. Magnetic resonance imaging, magnetic resonance arthrography and ultrasonography for assessing rotator cuff tears in people with shoulder pain for whom surgery is being considered (review). Cochrane Database Syst Rev. 2013;24(9):CD009020.