Title
Reliability of MR Quantification of Rotator Cuff Muscle Fatty Degeneration Using a 2-point Dixon Technique in Comparison with the Goutallier Classification: Validation Study by Multiple Readers.

Journal
Academic radiology, 24(11)

ISSN
1076-6332

Authors
Horiuchi, Saya
Nozaki, Taiki
Tasaki, Atsushi
et al.

Publication Date
2017-11-01

DOI
10.1016/j.acra.2017.03.026

Original Investigation

Reliability of MR Quantification of Rotator Cuff Muscle Fatty Degeneration Using a 2-point Dixon Technique in Comparison with the Goutallier Classification: Validation Study by Multiple Readers

Saya Horiuchi, MD, Taiki Nozaki, MD, PhD, Atsushi Tasaki, MD, PhD, Akira Yamakawa, MD, Yasuhito Kaneko, MD, PhD, Takeshi Hara, PhD, Hiroshi Yoshioka, MD, PhD

Rationale and Objectives: Presurgical assessment of fatty degeneration is important in the management of patients with rotator cuff tears. The Goutallier classification is widely accepted as a qualitative scoring system, although it is highly observer-dependent and has poor reproducibility. The objective of this study was to quantify fatty degeneration of the supraspinatus muscle using a 2-point Dixon technique in patients with rotator cuff tears by multiple readers, and to evaluate the reproducibility compared to Goutallier classification.

Materials and Methods: Two hundred patients with rotator cuff tears who underwent magnetic resonance imaging (MRI), including 2-point Dixon sequence at 3.0-T, were selected retrospectively. Qualitative and quantitative analyses of fatty degeneration were performed by two radiologists and three orthopedic surgeons independently. The fat quantification was performed by measuring signal intensity values of in phase (S(In)) and fat image (S(Fat)), and calculating fat fraction as S(Fat)/S(In). The reproducibility of MR quantification was analyzed by the intra- and interclass correlation coefficients and Bland-Altman plots.

Results: The interobserver agreement of the Goutallier classification among five readers was moderate ($k = 0.51$), whereas the interclass correlation coefficient regarding fat fraction value quantified in 2-point Dixon sequence was excellent (0.893). The mean differences in fat fraction values from the individual segmentation results were from −0.072 to 0.081. Proposed fat fraction grading and Goutallier grading showed similar frequency and distribution in severity of rotator cuff tears.

Conclusions: Fat quantification in the rotator cuff muscles using a 2-point Dixon technique at 3.0-T MRI is highly reproducible and clinically feasible in comparison to the qualitative evaluation using Goutallier classification.

Key Words: 2-point Dixon MR quantification; fatty degeneration; reproducibility; Goutallier classification; imaging biomarker.

INTRODUCTION

The assessment of rotator cuff muscle abnormalities after tear is one of the main factors in clinical decision-making in the management of patients with rotator cuff tears (1–3). Severity of presurgical fatty degeneration is especially relevant to the higher frequency of retear and worse functional outcomes after treatment (4–8). An accurate determination of fatty degeneration in the rotator cuff muscle is therefore a crucial element in evaluating surgical indications and postoperative prognosis.

For the qualitative assessment of fatty degeneration, computed tomography and magnetic resonance imaging (MRI)-based classifications have been established. The Goutallier classification has become widely accepted as standard 5-grade

Acad Radiol 2017; ■:■■–■■

From the Department of Radiology, St. Luke's International Hospital, 9-1 Akashi-cho, Chuo-ku, Tokyo 104-8560, Japan (S.H., T.N.); Department of Radiological Sciences, University of California, Irvine, California (T.N., Y.K., H.Y.); Department of Orthopaedic Surgery, St. Luke's International Hospital, Tokyo (A.T.); Division of Clinical Biotechnology, University of Tokyo Graduate School of Medicine, Tokyo, Japan (A.Y.); Department of Orthopaedic Surgery, Saitama Municipal Hospital, Saitama (Y.K.); Department of Intelligent Image Information, Division of Regeneration and Advanced Medical Sciences, Gifu University Graduate School of Medicine, Gifu (T.H.). Received February 16, 2017; accepted March 25, 2017. **Address correspondence to:** T.N. e-mail: nojyakki@gmail.com

http://dx.doi.org/10.1016/j.acra.2017.03.026

scoring systems for muscle fatty degeneration in current practice (9,10). The Goutallier classification was initially created based on axial computed tomography images; however, recently it was adapted for the evaluation of fatty degeneration on sagittal MR images. Although the Goutallier classification has been widely used, it is shown to be highly observer-dependent, and inter- and intraobserver reliability is not high (10–14). This fact has been a critical problem for the preoperative evaluation of fatty degeneration and assessment of operative indication.

Nozaki et al. have previously reported a method for quantitative assessment of fatty degeneration in the rotator cuff muscles based on a novel 2-point Dixon technique on 3.0-T MRI. Such MR quantification of fatty degeneration is useful pre- and postoperatively (3,15). To our knowledge, reproducibility of this quantitative assessment has not been assessed among multiple readers, including radiologists, orthopedic surgeons, board-certified physicians, and residents.

We hypothesized fat fraction values of the rotator cuff are more reliable and reproducible compared to the Goutallier classification. The purpose of this study was to determine the inter- and intraobserver reliability of the quantitative assessment of fatty degeneration in the rotator cuff muscles based on a 2-point Dixon technique, and to compare fat fraction values in MR quantification with qualitative assessment using Goutallier classification among multiple specialty-trained observers.

MATERIALS AND METHODS

Patients

The study was approved by our institutional review board, and the requirement for informed consent was waived because of its retrospective nature. Two hundred consecutive patients with 100 partial-thickness and 100 full-thickness rotator cuff tears (mean: 66.1 years, range: 34–91 years, 86 men and 114 women) who underwent MRI of the shoulder at 3.0-T between August 2010 and February 2014 were included in our study. Criteria for study inclusion included partial-thickness or full-thickness tears of the supraspinatus tendon as diagnosed on shoulder MRI at the time of exam. We excluded patients with history of shoulder surgery at the time of imaging.

MRI Acquisition

3.0-T MR examinations were performed (MAGNETOM Verio, Siemens AG, Erlangen, Germany) using a 4-channel flex coil with the following sequences: fast spin-echo (FSE) proton density-weighted images (PDWI) on oblique coronal (matrix size: 320 × 320; repetition time (TR): 2500 milliseconds; echo time (TE): 19 milliseconds; field of view (FOV): 180 mm; number of acquisitions: 1; slice thickness: 3 mm) and sagittal planes (matrix size: 320 × 320; TR: 2500 milliseconds; TE: 19 milliseconds; FOV: 180 mm; number of acquisitions: 1; slice thickness: 3 mm), fat-saturated FSE PDWI on oblique coronal (matrix size: 320 × 320; TR: 2500 milliseconds; TE: 37 milliseconds; FOV: 180 mm; number of acquisitions: 1; slice thickness: 3 mm) and sagittal planes (matrix size: 320 × 320; TR: 2500 milliseconds; TE: 37 milliseconds; FOV: 180 mm; number of acquisitions: 1; slice thickness: 3 mm), and a T2* sequence on the axial plane (matrix size: 320 × 320; TR: 900 milliseconds; TE: 13 milliseconds; FOV: 180 mm; number of acquisitions: 1; slice thickness: 3 mm). We also performed three-dimensional 2-point Dixon volumetric interpolated breath-hold examination (VIBE) sequence on the oblique-sagittal plane. The 2-point Dixon sequence used an acquisition matrix size of 128 × 128, TR of 6.5 milliseconds, TE of 1.225/2.4 milliseconds, flip angle of 10°, FOV of 196 mm, number of acquisitions 1, slice thickness of 2.5 mm, and acquisition time of 2 minutes 30 seconds. From the 2-point Dixon sequence, water-only, fat-only, in-phase (water and fat), and out-of-phase (water minus fat) images were produced.

Evaluation of the Rotator Cuff

The MR evaluation of the rotator cuff was performed by two radiologists (reader 1, a radiology resident with 4 years of experience; reader 2, a board-certified radiologist subspecialized in musculoskeletal radiology with 14 years of experience) in consensus. We categorized the type of tendon tears into two groups: partial-thickness tears and full-thickness tears. We also divided full-thickness tears into four subgroups—small tears (<1 cm), medium tears (1–3 cm), large tears (3–5 cm), and massive tears (>5 cm or more than two rotator cuff tears)—according to the Cofield grading system (16).

Qualitative Analysis of Fatty Degeneration

The supraspinatus muscles were assessed for fatty degeneration using the Goutallier classification in the most lateral oblique-sagittal section on FSE PDWI, which appears as "scapula Y-view." From these slices, the degree of fatty degeneration was graded according to a 5-point scale described by Goutallier et al. (9) and Fuchs et al. (10), shown in Figure 1: grade 0, normal; grade 1, some fatty streaks; grade 2, less than 50% fatty muscle; grade 3, as much fat as muscle; and grade 4, more fat than muscle. The qualitative evaluation of all subjects was performed using the Goutallier grading by two radiologists (readers 1 and 2) and three orthopedic surgeons (reader 3, a board-certified orthopedic surgeon subspecialized in spine with 15 years of experience; reader 4, an orthopedic surgery resident with 4 years of experience; reader 5, a board-certified orthopedic surgeon subspecialized in the shoulder with 19 years of experience), independently. Readers were blinded to the clinical information. Discrepancies were resolved by means of consensus, in accordance with the rules of the Goutallier grading (10).

Quantitative Analysis of Fatty Degeneration

Fat fraction value was calculated in a 2-point Dixon sequence after measurement of mean signal intensity value within the region

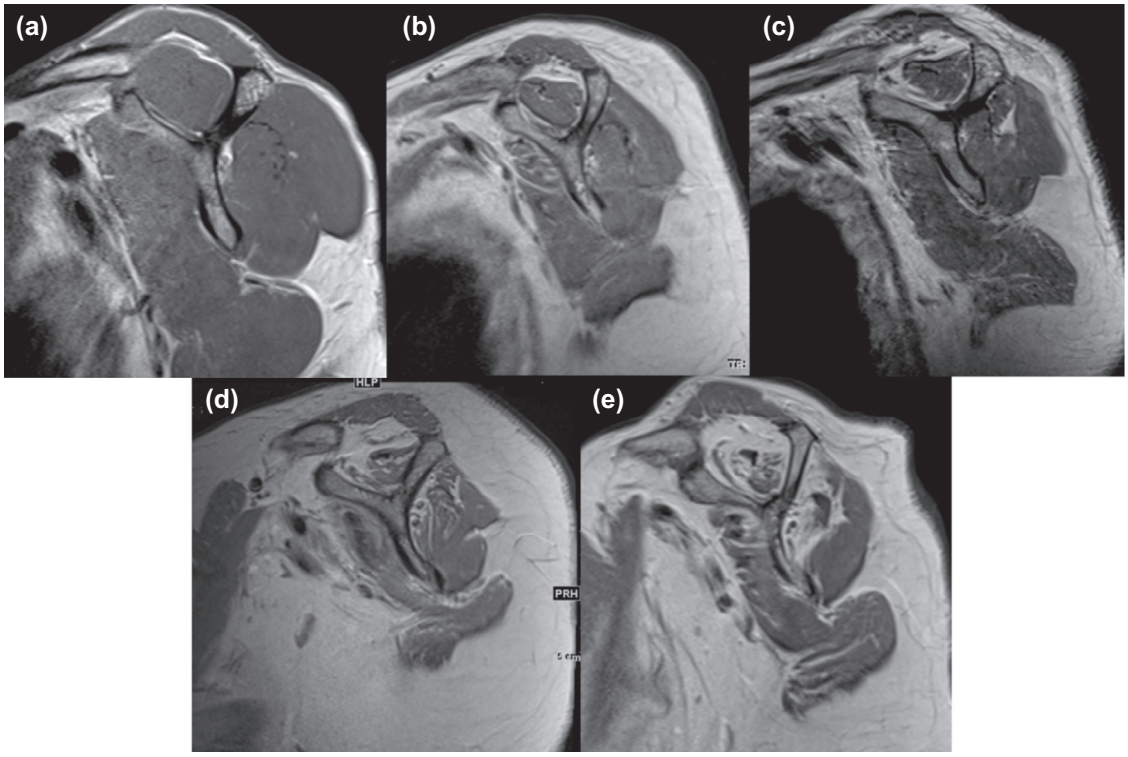

Figure 1. The Goutallier classification: oblique-sagittal proton density-weighted images show different degrees of fatty degeneration of the supraspinatus muscle: normal = grade 0 (**a**), some fat streaks = grade 1 (**b**), less fat than muscle = grade 2 (**c**), as much fat as muscle = grade 3 (**d**), and more fat than muscle = grade 4 (**e**).

of interest placed over supraspinatus muscle in the most lateral oblique-sagittal section, where the acromion, coracoid, and scapular body were all visible, and which appears as "scapula Y-view" (Fig 2), as previously reported (3,13). Fifty subjects were selected randomly from a pool of 200 subjects for the quantitative evaluation of the supraspinatus muscle. All five readers repeated segmentation twice independently with a 1-month interval in a randomized order to prevent recall bias. Next, two raters (readers 1 and 2) performed segmentation in all 200 subjects independently for comparison between the Goutallier grade and the fat fraction values in supraspinatus muscle. All raters performed segmentation on a picture archiving and communication system (Centricity, Ver.4.0, GE Healthcare, Barrington, IL). Signal intensity values on in-phase and fat images, which were produced using the three-dimensional 2-point VIBE sequence, were defined as S(In) and S(Fat). Fat fraction value in the supraspinatus muscle was calculated as the following:

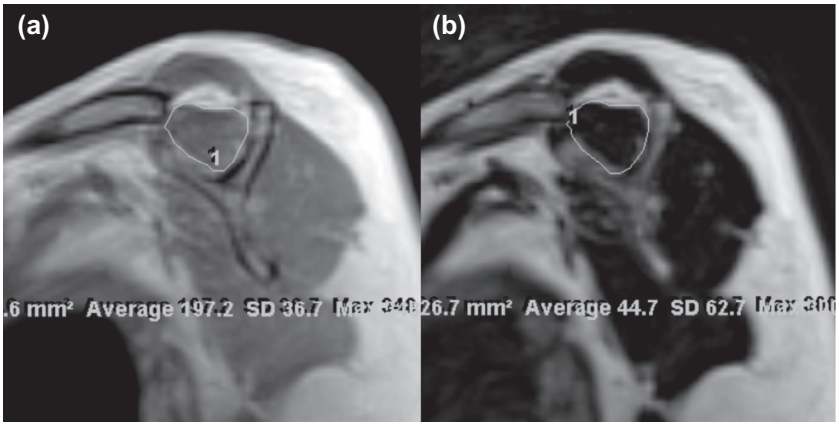

Figure 2. Measurements of signal intensity within the region of interest (*white outline*) over supraspinatus muscle were performed on oblique-sagittal in-phase image (**a**) and fat image (**b**) of a 2-point Dixon sequence. We calculated fat fraction value within supraspinatus muscle as signal intensity on fat image divided by signal intensity on in-phase image, equal to 0.227 in this case.

$$S(\text{Water}) + S(\text{Fat}) = S(\text{In})$$

$$\text{Fat fraction} = S(\text{Fat})/(S(\text{Water}) + S(\text{Fat})) = S(\text{Fat})/S(\text{In})$$

Statistical Analysis

According to a power analysis based on a chi-square goodness-of-fit test model from a prior study (17), the sample size ($n = 200$), mean, and variance yielded 100% statistical power at the significance level of 0.05 for a medium effect size.

The Fleiss' kappa analysis was performed to evaluate the interobserver reliability among five readers scoring using the Goutallier classification; the unweighted/weighted Cohen's kappa analyses were also performed for all possible combinations of reader pairs.

We used the intra- and interclass correlation coefficients (ICCs) to assess the intra- and interobserver reliability of the fat fraction values performed by five readers in the quantitative analysis of fatty degeneration. We also performed Bland-Altman plot analysis in the quantitative measurement of each five rater (18,19).

One-way analysis of variance and the Steel-Dwass post-hoc test were used to compare the relationship between fat fraction value and the Goutallier grade, as well as the relationship between fat fraction value and the degree of rotator cuff tear.

All statistical analyses were performed using statistical software R for MAC software (R Development Core Team, version 3.2.3; Vienna, Austria). The level of significance for all calculations was defined at $P < 0.05$.

RESULTS

Of the 200 subjects included in this study, 100 patients were included in the partial-thickness rotator cuff tear group (mean age: 63.3 ± 11.0, range: 34–91), nine patients in the full-thickness small tear group (mean age: 56.6 ± 9.2, range: 41–72), 39 patients in the full-thickness medium tear group (mean age: 68.6 ± 8.5, range: 47–84), 21 patients in the full-thickness large tear group (mean: 68.8 ± 8.2, range: 53–82), and 31 patients in the full-thickness massive tear group (mean: 72.9 ± 10.0, range: 53–91) (Table 1).

TABLE 2. Unweighted/Weighted Cohen's Kappa Value of the Goutallier Classification Between Reader Pairs

	Reader 1	Reader 2	Reader 3	Reader 4	Reader 5
Reader 1		0.47/0.73	0.44/0.80	0.65/0.86	0.53/0.78
Reader 2			0.16/0.64	0.54/0.78	0.50/0.76
Reader 3				0.36/0.76	0.34/0.68
Reader 4					0.52/0.79
Reader 5					

Interobserver agreement of the Goutallier classification among the five readers was moderate with a Fleiss' kappa value of 0.51. Interobserver agreement of the Goutallier classification between reader pairs was poor to good with a Cohen's kappa value from 0.16 to 0.65, and it was good to excellent with a weighted Cohen's kappa value from 0.64 to 0.86 (Table 2). Interobserver agreement for quantitative evaluation by fat fraction values among the five readers was excellent (ICC, 0.893; 95% confidence interval [CI], 0.845–0.931). The intraclass correlation coefficients for fat fraction values in each reader were also excellent (ICC, 0.956; 95% CI, 0.924–0.975, for reader 1; ICC, 0.953; 95% CI, 0.919–0.973, for reader 2; ICC, 0.899; 95% CI, 0.830–0.942, for reader 3; ICC, 0.931; 95% CI, 0.883–0.960, for reader 4; ICC, 0.893; 95% CI, 0.819–0.937, for reader 5). In Figure 3, Bland-Altman plots of the five readers' measurements are shown for the fat fraction values using a 2-point Dixon MR quantification. Tables 3 and 4 show the results of Bland-Altman inter- and intraobserver analysis. The interobserver correlation coefficients of the fat fraction value in each pair among the five readers were from 0.835 to 0.948. The mean differences in the fat fraction values from the individual segmentation results were from −0.072 to 0.081. Several fixed biases were observed: between reader 4 (orthopedic surgery resident) and all other readers, between reader 1 (radiology resident) and reader 3 (board-certified orthopedic surgeon subspecialized in spine), and between reader 1 and reader 5 (board-certified orthopedic surgeon subspecialized in shoulder). Proportional biases were also observed: between reader 4 and 3 of the four other readers (reader 1, reader 2, and reader 5) (Table 3). Regarding the intraobserver agreement, the correlation coefficients of the fat fraction value for each reader were from 0.904 to 0.966. The mean differences in the fat

TABLE 1. Patient's Demographic Data

	Group					
	Partial Tear	Small Tear	Medium Tear	Large Tear	Massive Tear	Combined
No. of patients	100	9	39	21	31	200
Percentage of patients	50	4.5	19.5	10.5	15.5	100
Age, y, mean ± standard deviation	63.3 ± 11	56.6 ± 9.2	68.6 ± 8.5	68.8 ± 8.2	7.29 ± 10	66.1 ± 10.8
Sex, *n* (%)						
Male	48 (48.0)	4 (44.4)	16 (41.0)	8 (38.0)	10 (32.3)	86 (43.0)
Female	52 (52.0)	5 (55.6)	23 (59.0)	13 (62.0)	21 (67.7)	114 (57.0)

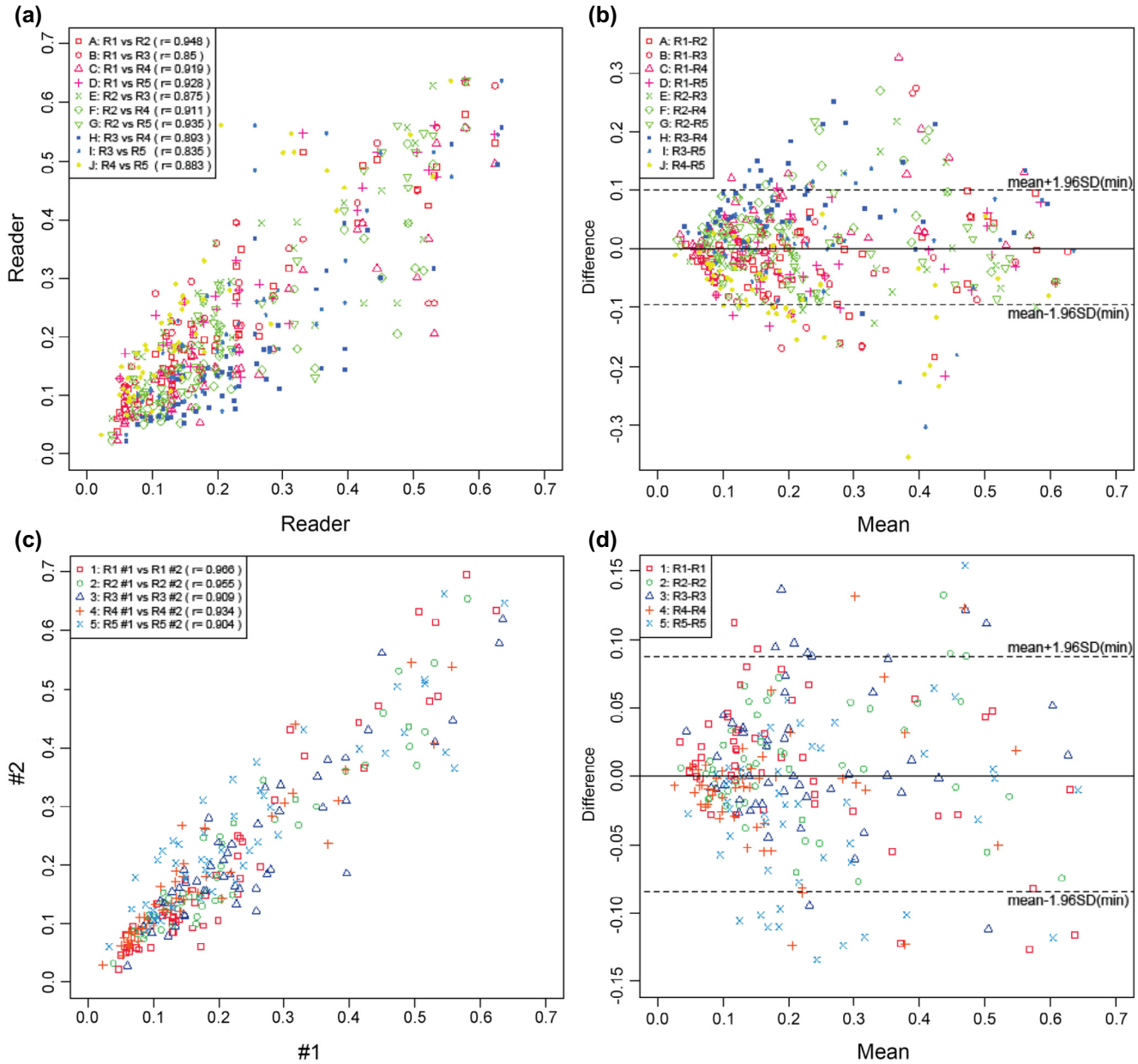

Figure 3. Bland-Altman plots of the five readers' measurements of fat fraction values using a 2-point Dixon MR quantification between inter- and intraobserver. Correlation plots of inter- and intraobserver agreements in fat fraction value within supraspinatus muscle (**a** and **c**). Difference plots of inter- and intraobserver agreements in fat fraction value within supraspinatus muscle. The *dotted line* indicates 95% limits of agreements (m ± 1.96 SD) using mean difference with minimum SD among the pairs (**b** and **d**). MR, magnetic resonance; SD, standard deviation.

fraction values from the individual segmentation results were from −0.022 to 0.021. Fixed biases in reader 3 (board-certified orthopedic surgeon subspecialized in spine) and reader 5 (board-certified orthopedic surgeon subspecialized in shoulder) and proportional bias in reader 1 (radiology resident) were observed (Table 4). However, on fixed biases in both intra- and interobserver agreement, the 95% limits of agreement as the mean difference and 1.96 standard deviation were presented in all pairs.

Figure 4 shows the results of rating based on Goutallier classification with fat fraction values in each reader. Figure 5 shows the box plot about the correlation between Goutallier grade by consensus of the five readers and fat fraction value. As the stage of Goutallier classification increases, the fat fraction value becomes larger with statistically significant difference in all pairs ($P < 0.001$).

Figure 6 shows the frequency and distribution of Goutallier grade in each rotator cuff tear group. The Cofield grade (small,

TABLE 3. Bland-Altman Analysis of Interobserver Agreement of Measurement of Fat Fraction Value of Supraspinatus Muscles

Proportional Bias			Fixed Bias			
Reader	PMC*	95% CI	*P* Value†	Mean Difference	SD	95% Limits of Agreement
R1-R2	0.211	(–0.103, 0.437)	0.353	–0.011	0.05	(–0.109, 0.087)
R1-R3	0.149	(–0.076, 0.459)	**0.048**	–0.032	0.083	(–0.195, 0.131)
R1-R4	**0.001**	(0.191, 0.644)	**<0.001**	0.049	0.064	(–0.076, 0.174)
R1-R5	0.863	(–0.256, 0.301)	**0.048**	–0.023	0.059	(–0.139, 0.093)
R2-R3	0.451	(–0.175, 0.376)	0.176	–0.021	0.072	(–0.162, 0.120)
R2-R4	**0.025**	(0.042, 0.547)	**<0.001**	0.06	0.061	(–0.060, 0.180)
R2-R5	0.345	(–0.400, 0.148)	0.353	–0.012	0.055	(–0.120, 0.096)
R3-R4	0.203	(–0.100, 0.439)	**<0.001**	0.081	0.063	(–0.042, 0.204)
R3-R5	0.205	(–0.439, 0.101)	0.444	0.009	0.086	(–0.160, 0.178)
R4-R5	**0.008**	(–0.588, –0.103)	**<0.001**	–0.072	0.073	(–0.215, 0.071)

CI, confidence interval; SD, standard deviation.
Bold: statistically significant difference at $P < 0.05$.
* Pearson's product-moment correlation.
† Obtained by multiple comparisons with Holm-Bonferroni correction.

TABLE 4. Bland-Altman Analysis of Intraobserver Agreement of Measurement of Fat Fraction Value of Supraspinatus Muscles

Proportional Bias			Fixed Bias			
Reader	PMC*	95% CI	*P* Value†	Mean Difference	SD	95% Limits of Agreement
R1(#1-#2)	**<0.001**	(–0.668, –0.231)	0.27	0.008	0.05	(–0.090, 0.106)
R2(#1-#2)	0.508	(–0.188, 0.364)	0.227	0.01	0.044	(–0.076, 0.096)
R3(#1-#2)	0.51	(–0.188, 0.364)	**0.017**	0.021	0.059	(–0.095, 0.137)
R4(#1-#2)	0.275	(–0.127, 0.417)	0.189	–0.009	0.046	(–0.099, 0.081)
R5(#1-#2)	0.149	(–0.075, 0.459)	**0.026**	–0.022	0.066	(–0.151, 0.107)

CI, confidence interval; SD, standard deviation.
Bold: statistically significant difference at $P < 0.05$.
* Pearson's product-moment correlation.
† Obtained by paired *t* test.

medium, large, or massive) of rotator cuff tear is correlated with the grade of Goutallier classification (Goutallier grades 0–4). No patients with Goutallier grade 0 were found in the massive tear group, whereas the patients with Goutallier grades 3 and 4 were not found in the partial tear group, the small tear group, and the medium tear group. Figure 7 shows the frequency and distribution of fat fraction grading (fat fraction value: <0.1 = grade 0; 0.1–0.2 = grade 1; 0.2–0.3 = grade 2; 0.3–0.4 = grade 3; >0.4 = grade 4) in each rotator cuff tear group. The frequency and distribution of fat fraction grade (Fig 7) resembled that of Goutallier grade (Fig 6).

DISCUSSION

Radiographic 5-grade scoring systems as proposed by Goutallier et al. (9) and Fuchs et al. (10) assess muscle fatty degeneration in a qualitative and subjective manner with poor reproducibility as previously reported (10–14). Our study also found just moderate interobserver agreement (k = 0.51) among the five readers for Goutallier classification grading of fatty degeneration, although it was good or excellent interobserver reliability between reader pairs using weighted Cohen's kappa. On the other hand, quantitative assessment of fatty degeneration based on a 2-point Dixon technique showed high consistency with the excellent ICC (0.893) among the five readers who have different subspecialties with different levels of experience.

We also assessed the reproducibility of quantitative measurement of fatty degeneration among multiple specialty-trained observers with Bland-Altman and correlation analysis from the viewpoints of data agreement and correlation. Bland-Altman plots for intraobserver agreement showed the proportional bias of reader 1 (a resident of radiology). There was a tendency to perform segmentation inconsistently in severe degeneration cases. The inter- and intraobserver reliability was slightly lower in reader 4 (a resident of orthopedic surgery) among the five readers. Due to less clinical experience, reader 4 tended to measure smaller areas than other observers, as seen in Table 3. He may have excluded peripheral fat around the supraspinatus muscle too precisely, resulting in a tendency for smaller fat fraction values. Bland-Altman plots for intraobserver agreement also showed fixed bias of readers 3 and 5

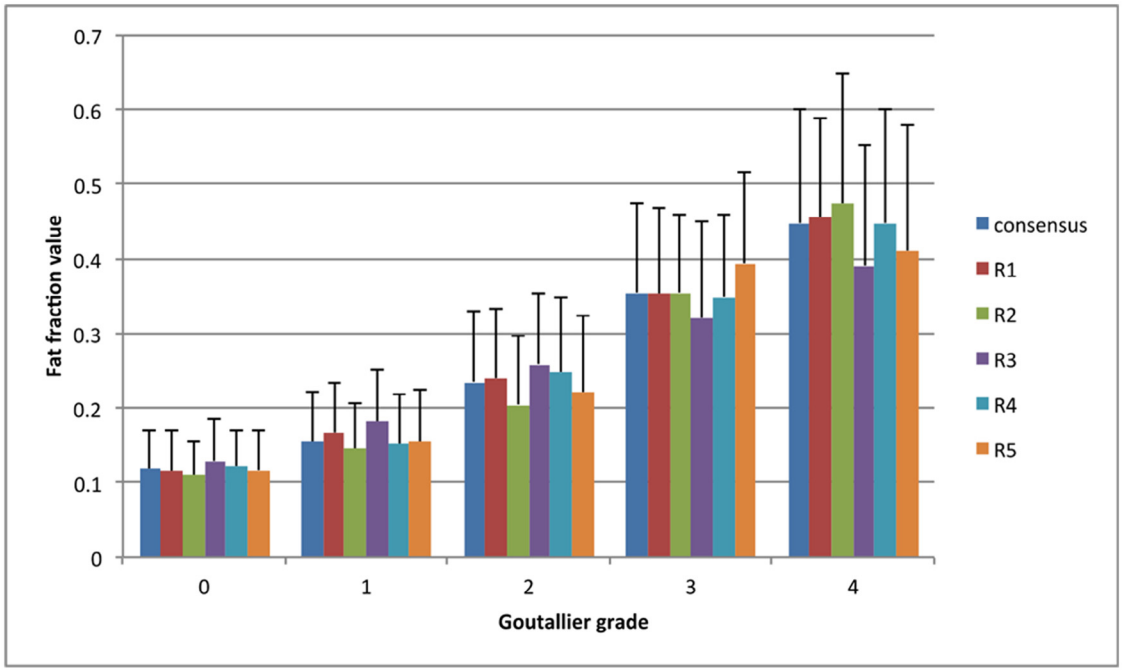

Figure 4. The correlation between fat fraction values and the rating based on Goutallier grade in each reader.

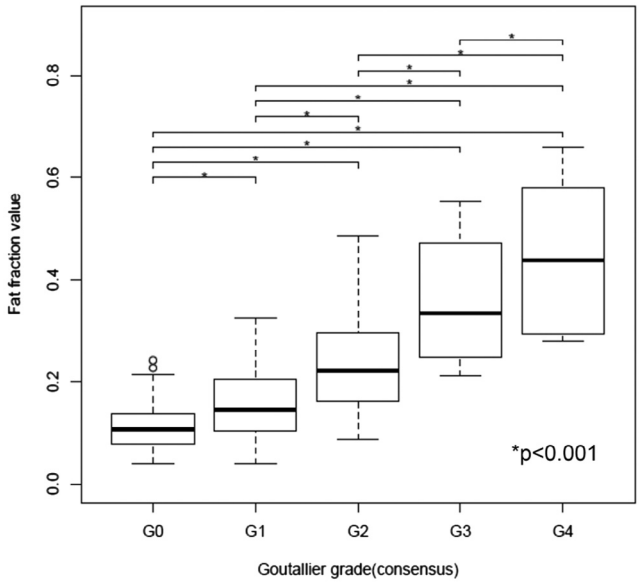

Figure 5. The box plot about the correlation between Goutallier grade and fat fraction value. As the grade of Goutallier classification elevates, the fat fraction value becomes larger with significant difference (*$P < 0.001$).

(board-certificated orthopedic surgeons). However, because the 95% limits of agreement as the mean difference and 1.96 standard deviation were presented in all pairs, the measurements between initial and second segmentation within intrareader as well as between two readers are judged as equivalent regarding fixed bias. Overall, although there are subtle variations of intra- and interrater reproducibility of the measurement of fatty degeneration by five readers, reproducibility looks higher than the Goutallier grading.

Another issue of the Goutallier classification is that fat fraction grading is inaccurate. The percentage of fat within supraspinatus muscle measured with a 2-point Dixon method in our study was different compared to the original definition of the Goutallier classification. The Goutallier grade 3 was defined as equal amounts of fat and muscle (=50% fat fraction), and grade 4 as more fat than muscle (=more than 50% fat fraction). However, the range of fat fraction value (S(fat)/S(in)) within the muscle was 21.1%–55.4% (average: 35.5%) for grade 3, and 27.9%–65.9% (average 44.8%) for grade 4 in our study; the average fat fraction values in both Goutallier grades 3 and 4 were below 50%. Nardo et al. also reported that the percentage of fat within rotator cuff muscles assessed with chemical shift-based multipoint water-fat separation MRI techniques was lower than the original Goutallier grade (13). They mentioned the possibility of overestimation of the fat content using qualitative Goutallier classification. Our result with a larger number of subjects supports their results.

In general, the degree of fatty degeneration of the supraspinatus muscle increases as rotator cuff tears progress, with the most severe degeneration occurring in massive tears as previously reported (2,3,6,20). A high correlation between 2-point Dixon-based fat quantification and Goutallier classification was found with statistical significance in our study. The group of patients with partial, small, or medium rotator cuff tears was composed of fatty degeneration less than or equal to Goutallier grade 2, whereas only patients with large or massive rotator cuff tears showed fatty degeneration greater than or equal to Goutallier grade 3. The presence of fatty degeneration is considered a negative prognostic indicator of surgical rotator cuff

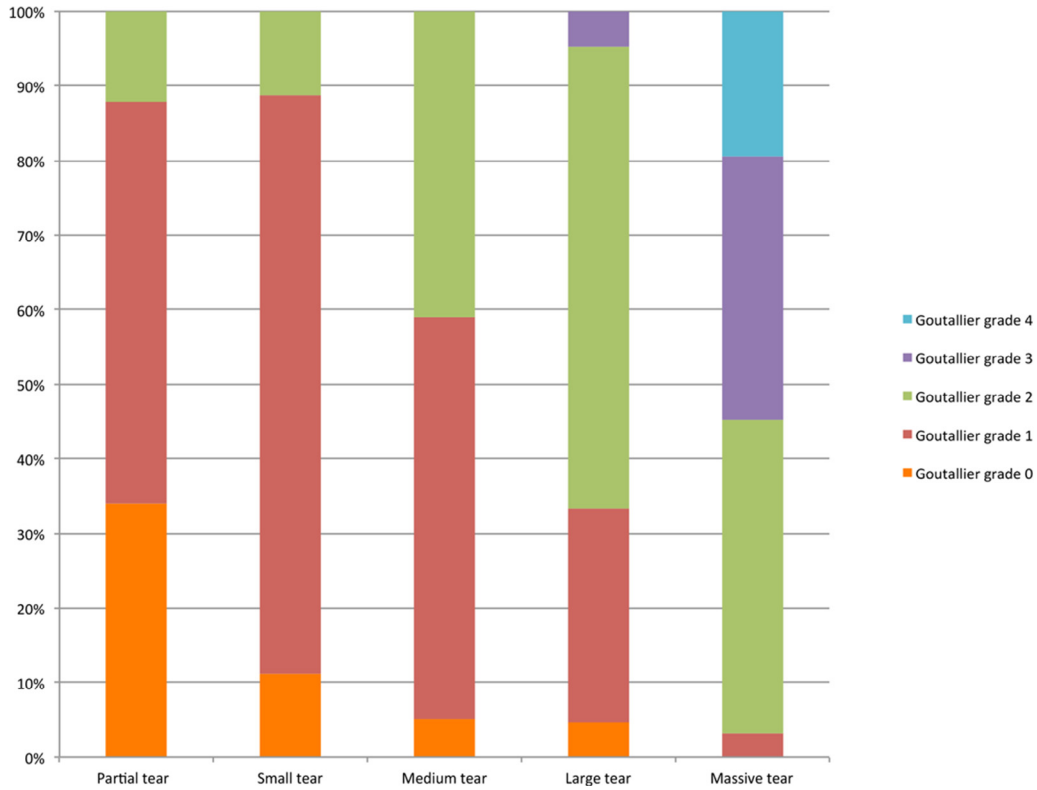

Figure 6. Frequency and distribution of Goutallier grade in each rotator cuff tear group. As the degree of rotator cuff tear progresses, the stage of Goutallier classification elevates.

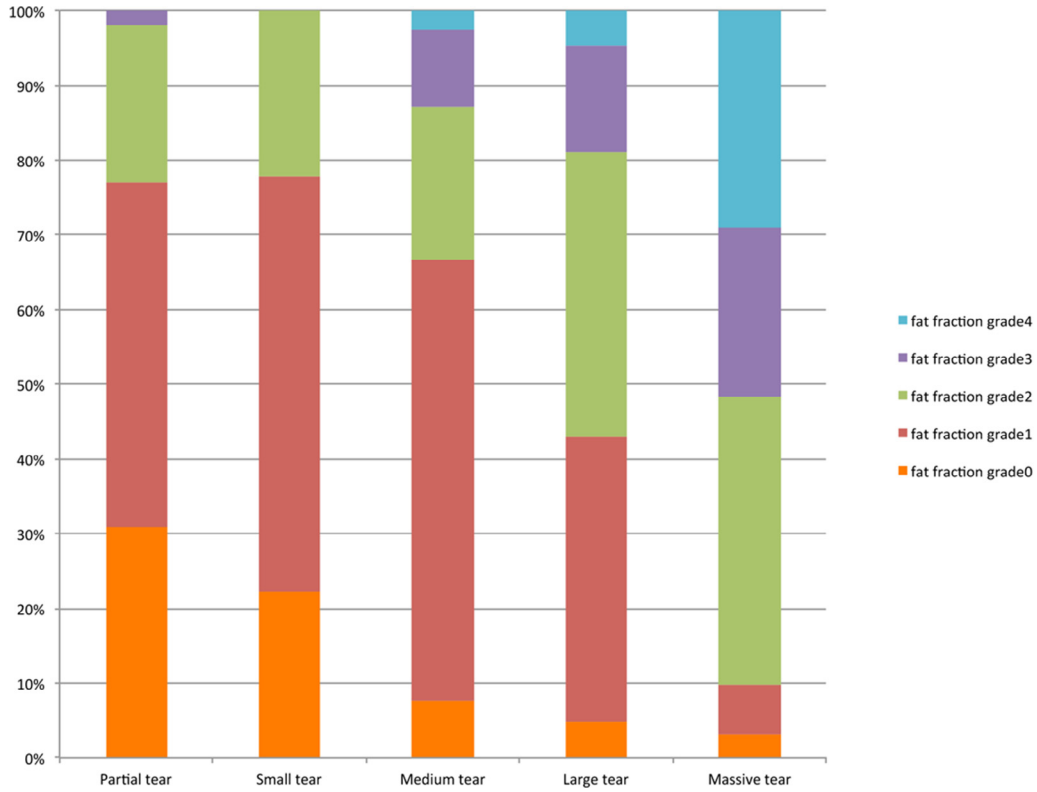

Figure 7. Frequency and distribution of fat fraction grade (fat fraction value <0.1 = grade 0; 0.1–0.2 = grade 1; 0.2–0.3 = grade 2; 0.3–0.4 = grade 3; >0.4 = grade 4) in each rotator cuff tear group. The frequency and distribution of fat fraction grade resembled that of Goutallier grade.

repair, and rotator cuff tears greater than or equal to Goutallier grade 3 are regarded as relative contraindications to rotator cuff repair (4,21). From the results of fat fraction values in the current study (Fig 7), the supraspinatus muscles in more than half of patients with massive rotator cuff tears had fat fraction grades 3 and 4. In other words, more than half of massive rotator cuff tears showed greater than 30% fat fraction value. Likewise, around 20% and 60% of large rotator cuff tears did greater than 30% and 20% of fat fraction value, respectively. This result suggests fat fraction value or grading is correlated with severity of rotator cuff tears similar to Goutallier grading, and shows the potential of its usefulness for presurgical management.

This study has several limitations. First, because evaluation of rotator cuff tears was performed only with MR images, we lack surgical proof of tendon tears with arthroscopy or surgery. A second practical limitation is the time-consuming nature of manual muscle segmentation in daily clinical practice. Further technical development of fully automated segmentation is needed to be useful in clinical settings. Third, we measured the fat fraction value in the most lateral oblique-sagittal section, which appears as "scapula Y-view." We do not know whether or not the fat fraction values in our study reflect those of the whole supraspinatus muscle. Further study will be needed to evaluate the fat fraction values of the whole muscle. Fourth, additional scan time in routine shoulder MRI protocol was needed to perform a 2-point Dixon sequence. However, its scan time is only about 2 and half minutes using our 3.0-T shoulder protocol. The 2-point Dixon technique as used in our study can also be performed on 1.5-T MRI units. Finally, the number of patients with full-thickness small rotator cuff tears was small in this study. A further study will be needed with a large cohort of rotator cuff tears to prove the usefulness of the quantitative assessment with fat faction values.

CONCLUSION

We demonstrate that quantitative assessment of fatty degeneration in the rotator cuff muscles based on a 2-point Dixon technique is feasible, yielding more objective and reproducible results with excellent inter- and intraobserver reliability in comparison to qualitative evaluation using the Goutallier classification. This improved and potentially widely available method of evaluating fatty degeneration is promising for presurgical and postsurgical evaluation of rotator cuff muscles, treatment selection, and prediction of treatment outcomes more objectively.

ACKNOWLEDGMENT

The authors would like to thank Dr. Jay Starkey for English editing and helpful comments.

REFERENCES

1. Gladstone JN, Bishop JY, Lo IK, et al. Fatty infiltration and atrophy of the rotator cuff do not improve after rotator cuff repair and correlate with poor functional outcome. Am J Sports Med 2007; 35:719–728.
2. Barry JJ, Lansdown DA, Cheung S, et al. The relationship between tear severity, fatty infiltration, and muscle atrophy in the supraspinatus. J Shoulder Elbow Surg 2013; 22:18–25.
3. Nozaki T, Tasaki A, Horiuchi S, et al. Quantification of fatty degeneration within the supraspinatus muscle by using a 2-point Dixon method on 3-T MRI. AJR Am J Roentgenol 2015; 205:116–122.
4. Goutallier D, Postel JM, Gleyze P, et al. Influence of cuff muscle fatty degeneration on anatomic and functional outcomes after simple suture of full-thickness tears. J Shoulder Elbow Surg 2003; 12:550–554.
5. Burkhart SS, Barth JR, Richards DP, et al. Arthroscopic repair of massive rotator cuff tears with stage 3 and 4 fatty degeneration. Arthroscopy 2007; 23:347–354.
6. Mellado JM, Calmet J, Olona M, et al. Surgically repaired massive rotator cuff tears: MRI of tendon integrity, muscle fatty degeneration, and muscle atrophy correlated with intraoperative and clinical findings. AJR Am J Roentgenol 2005; 184:1456–1463.
7. Sugihara T, Nakagawa T, Tsuchiya M, Ishizuki M. Prediction of primary reparability of massive tears of the rotator cuff on preoperative magnetic resonance imaging. J Shoulder Elbow Surg 2003; 12:222–225.
8. Liem D, Lichtenberg S, Magosch P, et al. Magnetic resonance imaging of arthroscopic supraspinatus tendon repair. J Bone Joint Surg Am 2007; 89:1770–1776.
9. Goutallier D, Postel JM, Bernageau J, et al. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. Clin Orthop Relat Res 1994; 304:78–83.
10. Fuchs B, Weishaupt D, Zanetti M, et al. 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.
11. Slabaugh MA, Friel NA, Karas V, et al. 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; 40:1728–1734.
12. Lippe J, Spang JT, Leger RR, et al. Inter-rater agreement of the Goutallier, Patte, and Warner classification scores using preoperative magnetic resonance imaging in patients with rotator cuff tears. Arthroscopy 2012; 28:154–159.
13. Nardo L, Karampinos DC, Lansdown DA, et al. Quantitative assessment of fat infiltration in the rotator cuff muscles using water-fat MRI. J Magn Reson Imaging 2014; 39:1178–1185.
14. Williams MD, Ladermann A, Melis B, et al. Fatty infiltration of the supraspinatus: a reliability study. J Shoulder Elbow Surg 2009; 18:581–587.
15. Nozaki T, Tasaki A, Horiuchi S, et al. Predicting retear after repair of full-thickness rotator cuff tear: two-point Dixon MR imaging quantification of fatty muscle degeneration-initial experience with 1-year follow-up. Radiology 2016; 280:500–509.
16. Cofield RH. Subscapular muscle transposition for repair of chronic rotator cuff tears. Surg Gynecol Obstet 1982; 154:667–672.
17. Cohen J. A power primer. Psychol Bull 1992; 112:155–159.
18. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307–310.
19. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999; 8:135–160.
20. Gokalp G, Yildirim N, Yazici Z, et al. Using chemical-shift MR imaging to quantify fatty degeneration within supraspinatus muscle due to supraspinatus tendon injuries. Skeletal Radiol 2010; 39:1211–1217.
21. Jost B, Pfirrmann CW, Gerber C, et al. Clinical outcome after structural failure of rotator cuff repairs. J Bone Joint Surg Am 2000; 82:304–314.