Severity of Glenohumeral Osteoarthritis Does Not Correlate With Patient-Reported Outcomes

Eitan M Kohan, MD[1] ⓘ, Jeffrey Ryan Hill, MD[1],
Joseph D Lamplot, MD[1], Alexander W Aleem, MD[1],
Jay D Keener, MD[1] ⓘ and Aaron M Chamberlain, MD[1]

Abstract

Background: Patient pain and clinical function are important factors in decision-making for patients with glenohumeral osteoarthritis (GHOA). The correlation between radiographic severity of arthritis and demographic factors with modern patient-reported outcome measures has not yet been well defined.

Methods: This cross-sectional study included 256 shoulders in 246 patients presenting with isolated GHOA. All patients obtained standard radiographs and completed the American Shoulder and Elbow Surgeons score, Simple Shoulder Test (SST), Shoulder Activity Scale, Visual Analog Scale, and Patient-Reported Outcome Measurement Information System (PROMIS) computer adaptive tests at the time of presentation. Radiographs were graded according to the Samilson–Prieto classification. Mean pain and functional scores were compared between the radiographic grades of osteoarthritis (OA) and demographic factors.

Results: There were 6 shoulders rated as grade 1 OA, 41 shoulders as grade 2, 149 shoulders as grade 3a, and 65 shoulders as grade 3b. There was excellent interobserver reliability in grade of OA ($\kappa = 0.77$). There were no significant differences in patient-reported pain or any validated measure of clinical function between radiographic grades of OA ($P > .05$). Males reported higher function and lower pain scores than females ($P = .001–.066$), although only the values for the SST and PROMIS physical function test were clinically relevant.

Discussion: While gender correlated with pain and function, the clinical relevance is limited. Radiographic severity of GHOA does not correlate with patient-reported pain and function, and symptoms should remain the primary determinants of surgical decision-making. Further investigation is necessary to examine whether radiographic severity of OA influences improvement following operative intervention in this population.

Introduction

The shoulder is a complex joint with a variety of pathologies that can cause pain and dysfunction, with the reported lifetime prevalence of shoulder pain ranging from 7% to 67%.[1] Nearly 20 million people in the United States reported shoulder pain in 2008 alone, making it the third most common musculoskeletal complaint following back and knee pain, and that number is expected to increase as the population continues to age.[2]

Glenohumeral osteoarthritis (GHOA) is a common cause of shoulder pain and dysfunction that is associated with lower functional levels than patients with other shoulder conditions, independent of age.[3,4] As a result, patients experience a decreased quality of life and pose a

[1]Department of Orthopaedic Surgery, Washington University in Saint Louis, Saint Louis, Missouri

Corresponding Author:
Eitan M Kohan, Department of Orthopaedic Surgery, Washington University in St. Louis, 660 S. Euclid Ave., Campus Box 8233, St. Louis, MO 63110, USA.
Email: emkohan@gmail.com

substantial socioeconomic burden.[5] While several radiographic classifications have been developed to describe the degenerative changes seen in GHOA, it is unclear to what extent the radiographic severity of GHOA should factor into clinical decision-making and timing of surgery.[6–13]

Over the past 2 decades, there has been increased focus on understanding patients' perceptions of their illness and its effect on their quality of life, and patient-reported outcomes (PROs) have become increasingly used to aid decision-making and to monitor response to treatments.[14] This focus on PROs has led to the development of over 30 outcome measure instruments for shoulder pathology alone, although the data comparing radiographic severity of GHOA with PRO measures (PROMs) are limited.[15,16]

Increasing osteophyte size has been shown to correlate with decreased range of motion; however, radiographic severity does not correlate with reported pain scores.[6] Pain and range of motion alone do not entirely capture how patients function in their daily lives, and the correlation of radiographic severity with patient perceptions of their illness remains incompletely answered. The purpose of this study is to evaluate the association between radiographic severity of GHOA and patient demographic factors with pain and function as determined by PROMs.

Methods

Patient Selection

This was a retrospective cross-sectional case series. All clinic visits at our high-volume regional-referral shoulder arthroplasty center were reviewed from August 2015 to August 2017. Patients with an isolated International Classification of Diseases, 10th Revision (ICD-10) code of M19.011 (primary osteoarthritis [OA], right shoulder), M19.012 (primary OA, left shoulder), and M19.019 (primary OA, unspecified shoulder) were included for review. Clinic notes and radiographs were reviewed to ensure that included patients had an isolated diagnosis of GHOA. Exclusion criteria were patients with coexisting diagnoses (eg, full-thickness rotator cuff tear, inflammatory arthritis, cervical radiculopathy), ipsilateral shoulder surgery within the past year, outcome measures collected more than 1 year from the time of radiographs, and incomplete medical records.

Data Collection

All patient information was collected from at the initial clinic visit. Patient demographics collected included gender, age, body mass index (BMI), previous surgery, smoking status, hand dominance, and laterality of GHOA. PROMs collected included Visual Analog Scale (VAS, range 0–10), American Shoulder and Elbow Surgeon score (ASES, range 0–100),[17] Simple Shoulder Test (SST, range 0–12),[18] Shoulder Activity Scale (SAS, range 0–20),[19] Patient-Reported Outcome Measurement Information System (PROMIS) physical function (PF), PROMIS upper extremity (UE), and PROMIS pain interference (PI).[20,21] PROMIS instruments are administered as computer adaptive tests (CATs). Responses to the first prompt guide the system's choice of the next question. The CAT continues until either the standard error drops below a T score metric of 3.0 or the patient has answered 12 questions, whichever occurs first. PROMIS instruments are calibrated with a score of 50 as the average for the U.S. general population with a standard deviation of 10. For all outcome measures collected, a higher score corresponds to more of the concept being measured (ie, a higher ASES score translates to higher function, a lower PROMIS PI score translates to less pain).

Radiographs of the affected shoulder were obtained at the initial clinic visit. Three observers analyzed all radiographs: 1 board-certified shoulder and elbow surgeon and 2 orthopedic surgery residents. Radiographic severity was rated according to the Samilson–Prieto scale.[10] The true anterior–posterior (Grashey) radiograph was used to measure the size of the inferior osteophyte. Inferior osteophytes measuring $<3\,mm$ were rated as grade 1, osteophytes between 3 mm and 7 mm were rated as grade 2, and osteophytes $>7\,mm$ were rated as grade 3. As many patients presenting with GHOA have large inferior osteophytes, the grade 3 osteophytes were divided into 2 subgroups: inferior osteophytes measuring between 8 mm and 15 mm were rated as grade 3a, and inferior osteophytes measuring $>15\,mm$ were rated as grade 3b.

Statistical Analysis

No a priori sample size determination was performed as this was a retrospective study and all available patients were included.

Interobserver reliability was assessed using intraclass correlation coefficient (kappa, κ). Interobserver reliability <0.4 was rated as poor, between 0.40 and 0.59 as fair, between 0.60 and 0.74 as good, and >0.75 as excellent.[22] In the setting of interobserver disagreement, radiographic grade for statistical analysis was determined by the majority opinion among the 3 observers.

Data normality was assessed using the Shapiro–Wilk test. Equality of variance between groups was assessed with Levene test. One-way analysis of variance testing with a post hoc Tukey analysis was performed between grades of radiographic arthritis to assess for differences in patient-reported pain and function, as measured by

PROMs. The influence of gender, dominant extremity involvement, BMI >30 kg/m^2, and history of previous surgery on PROMs were compared by the Student t test or Mann–Whitney U test based on data normality. A P value of <.05 was considered to be significant.

Results

Patient Demographics

During the study period, 1497 patients presented with an ICD-10 code corresponding to GHOA. After excluding those with concomitant diagnoses and those with incomplete survey data, 256 shoulders in 246 patients were eligible for inclusion. Reasons for exclusion are shown in Figure 1. Overall demographics are shown in Table 1. A slight majority of patients were male and the dominant extremity was involved in approximately half. The majority of patients had no prior history of shoulder surgery on the affected extremity. Of the 42 who underwent prior surgery, 3 had an isolated biceps procedure, 1 had a capsular release, 1 had a distal clavicle resection, 14 had arthroscopic debridements, 3 had an arthroscopic labral repair, 1 had an open capsular shift, 13 had arthroscopic rotator cuff repairs, 3 had arthroscopic superior labrum tear from anterior to posterior (SLAP) repairs, and 3 had "soft tissue stabilization" procedures. The previous surgeries were done at an average of 12 years prior to presentation (range 1–49 years).

Radiographic Severity

Of the 256 shoulders, 6 were rated as grade 1, 41 shoulders as grade 2, 146 shoulders as grade 3a, and 63 shoulders as grade 3b. There was excellent interobserver reliability in the radiographic grading of GHOA ($\kappa = 0.77$).

Association With PROs

Overall mean and standard deviations for ASES, SAS, SST, VAS, PROMIS PF, PROMIS UE, and PROMIS PI are shown in Table 2. While there were consistent measurable disabilities and functional limitations of the shoulder in this cohort, there were no significant differences in any patient-reported pain or functional outcome measures among the different radiographic grades of GHOA ($P = .16$–1.0). The results of the Tukey post hoc test are shown in Table 3.

Compared to females, males showed higher function and lower pain scores across all outcome measures except VAS (though VAS approached statistical significance). The mean difference between genders was 7.5 points for ASES, 3.7 points for SAS, 2.0 points for SST, 4.9 points for PROMIS PF, 3.4 points for PROMIS UE, and 2.2 points for PROMIS PI.

Table 1. Patient Demographics.

Gender	Male = 141 (57.3%)
	Female = 105 (42.7%)
Age (mean ± SD)	66.1 ± 10.2
Laterality	Right = 131 (51.2%)
	Left = 125 (48.8%)
% Dominant extremity involved	48.8%
Previous surgery	Yes = 42 (16.4%)
	No = 214 (83.6%)
Body mass index (mean ± SD)	31.8 ± 7.5
Smoking status (%)	Active = 4.3%
	Former = 37.1%
	Nonsmoker = 58.6%

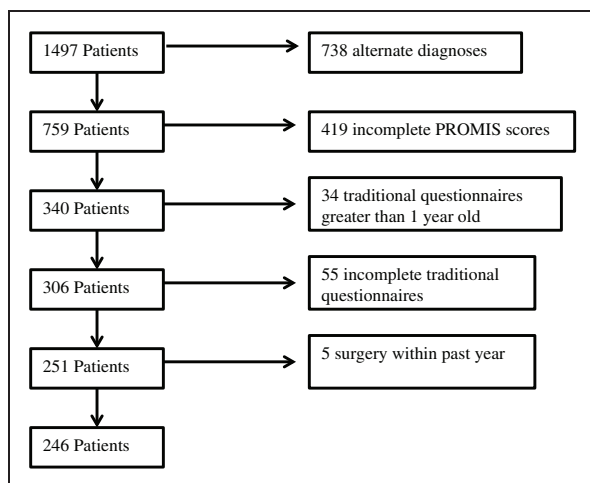Abbreviation: SD, standard deviation.

Figure 1. Exclusion criteria. PROMIS, Patient-Reported Outcome Measurement Information System.

Table 2. Descriptive Statistics of Outcome Measures.

Outcome Measure	Mean ± SD
ASES	35.3 ± 17.2
SAS	7.8 ± 5.2
SST	4.2 ± 2.8
VAS	6.9 ± 2.0
PROMIS PF	41.5 ± 8.4
PROMIS UE	32.8 ± 6.9
PROMIS PI	61.2 ± 5.9

Abbreviations: ASES, American Shoulder and Elbow Surgeons; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcome Measurement Information System; SAS, Shoulder Activity Scale; SD, standard deviation; SST, Simple Shoulder Test; UE, upper extremity; VAS, Visual Analog Scale.

Table 3. Tukey Post Hoc Analysis Based on Radiographic Osteoarthritis.

Outcome Measure	Radiographic Grade	n	Mean	SD	95% CI for Mean Difference 1	2	3a	Tukey HSD Comparisons, P 1	2	3a
ASES	1	6	27.8	6.7						
	2	41	37.4	21.4	−9.7 to 29.0			.57		
	3a	146	36.6	15.3	−9.5 to 27.3	−8.6 to 7.1		.60	.99	
	3b	63	31.5	18.5	−15.2 to 22.6	−14.8 to 2.9	−11.9 to 1.5	.96	.31	.19
SAS	1	6	10.2	6.3						
	2	41	7.2	4.9	−8.9 to 2.9			.56		
	3a	146	7.6	5.1	−8.1 to 3.1	−1.9 to 2.8		.65	.96	
	3b	63	8.3	5.6	−7.6 to 3.9	−1.6 to 3.8	−1.4 to 2.7	.83	.72	.85
SST	1	6	4.0	2.8						
	2	41	5.0	3.3	−2.2 to 4.2			.84		
	3a	146	4.1	2.7	−2.9 to 3.2	−2.2 to 0.4		1.00	.28	
	3b	63	4.0	2.8	−3.2 to 3.1	−2.6 to 0.4	−1.3 to 0.9	1.00	.24	.98
VAS	1	6	7.7	1.4						
	2	41	6.7	2.3	−3.3 to 1.3			.70		
	3a	146	6.7	2.0	−3.2 to 1.2	−1.0 to 0.9		.66	1.00	
	3b	63	7.3	2.0	−2.6 to 1.9	−0.4 to 1.7	−0.2 to 1.4	.98	.42	.16
PROMIS PF	1	6	41.5	10.4						
	2	41	40.7	11.4	−10.4 to 8.6			1.00		
	3a	146	42.2	7.5	−8.4 to 9.7	−2.3 to 5.4		1.00	.72	
	3b	63	40.5	7.9	−10.3 to 8.2	−4.5 to 4.2	−5.0 to 1.5	.99	1.00	.51
PROMIS UE	1	6	31.6	9.6						
	2	41	33.7	9.5	−5.8 to 10.0			.90		
	3a	146	33.0	5.9	−6.1 to 8.9	−3.9 to 2.5		.96	.94	
	3b	63	31.8	6.9	−7.4 to 7.9	−5.4 to 1.8	−3.8 to 1.6	1.00	.56	.71
PROMIS PI	1	6	62.6	6.1						
	2	41	61.4	6.6	−7.9 to 5.5			.97		
	3a	146	60.6	5.7	−8.4 to 4.3	−3.5 to 1.9		.84	.87	
	3b	63	62.3	6.0	−6.8 to 6.2	−2.1 to 4.0	−0.6 to 4.0	1.00	.86	.21

Abbreviations: ASES, American Shoulder and Elbow Surgeons; CI, confidence interval; HSD, honest significant difference; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcome Measurement Information System; SAS, Shoulder Activity Scale; SD, standard deviation; SST, Simple Shoulder Test; UE, upper extremity; VAS, Visual Analog Scale.

Dominant extremity involvement was associated with higher PROMIS PF (mean difference of 2.2 points), but otherwise showed no differences among outcome measures. A history of prior surgery was associated only with a higher SST compared to those without any prior surgery on the affected shoulder (mean difference 1.1 points). A BMI over $30 \, kg/m^2$ was associated with a lower SST, lower PROMIS PF, and higher PROMIS PI compared to those with a lower BMI (mean difference 0.7 points, 2.7 points, and 2.4 points, respectively). Results of the Student t tests are shown in Tables 4 to 7.

Discussion

PROMs have been increasingly used to quantify patient pain and function preoperatively and monitor improvement following surgery. While factors such as mental

Table 4. Patient Outcomes by Gender.

Outcome Measure	Male n = 141 Mean ± SD	Female n = 105 Mean ± SD	P
ASES	38.5 ± 17.4	31.1 ± 16.0	.001
SAS	9.4 ± 5.0	5.7 ± 4.7	<.001
SST	5.1 ± 2.8	3.1 ± 2.5	<.001
VAS	6.7 ± 2.1	7.1 ± 2.0	.066
PROMIS PF	43.6 ± 7.4	38.8 ± 8.8	<.001
PROMIS UE	34.2 ± 6.6	30.8 ± 7.0	<.001
PROMIS PI	60.2 ± 5.5	62.5 ± 6.2	.003

Abbreviations: ASES, American Shoulder and Elbow Surgeons; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcome Measurement Information System; SAS, Shoulder Activity Scale; SD, standard deviation; SST, Simple Shoulder Test; UE, upper extremity; VAS, Visual Analog Scale.

Table 5. Patient Outcomes by Extremity Involved.

Outcome Measure	Dominant n = 125 Mean ± SD	Nondominant n = 131 Mean ± SD	P
ASES	36.9 ± 17.8	33.7 ± 16.6	.137
SAS	8.0 ± 5.2	7.6 ± 5.2	.466
SST	4.5 ± 3.0	4.0 ± 2.7	.191
VAS	6.8 ± 2.1	7.0 ± 2.0	.451
PROMIS PF	42.7 ± 8.1	40.4 ± 8.5	.034
PROMIS UE	33.4 ± 6.8	32.2 ± 7.1	.161
PROMIS PI	61.2 ± 5.4	61.1 ± 6.4	.898

Abbreviations: ASES, American Shoulder and Elbow Surgeons; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcome Measurement Information System; SAS, Shoulder Activity Scale; SD, standard deviation; SST, Simple Shoulder Test; UE, upper extremity; VAS, Visual Analog Scale.

Table 6. Patient Outcomes by Previous Surgery.

Outcome Measure	Prior Surgery n = 42 Mean ± SD	No Prior Surgery n = 214 Mean ± SD	P
ASES	36.1 ± 18.4	35.1 ± 17.0	.736
SAS	9.2 ± 4.9	7.5 ± 5.2	.052
SST	5.2 ± 2.9	4.0 ± 2.8	.018
VAS	7.0 ± 2.1	6.9 ± 2.0	.666
PROMIS PF	42.6 ± 6.9	41.3 ± 8.6	.345
PROMIS UE	33.2 ± 6.5	32.7 ± 7.0	.682
PROMIS PI	61.2 ± 5.4	61.2 ± 6.0	.991

Abbreviations: ASES, American Shoulder and Elbow Surgeons; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcome Measurement Information System; SAS, Shoulder Activity Scale; SD, standard deviation; SST, Simple Shoulder Test; UE, upper extremity; VAS, Visual Analog Scale.

Table 7. Patient Outcomes by BMI.

Outcome Measure	BMI > 30 n = 131 Mean ± SD	BMI < 30 n = 125 Mean ± SD	P
ASES	34.2 ± 17.4	36.4 ± 17.0	.306
SAS	7.3 ± 5.4	8.3 ± 4.9	.147
SST	3.9 ± 2.8	4.6 ± 2.9	.035
VAS	6.8 ± 2.0	6.9 ± 2.1	.686
PROMIS PF	40.2 ± 8.6	42.9 ± 8.0	.009
PROMIS UE	32.1 ± 6.8	33.5 ± 7.0	.116
PROMIS PI	62.4 ± 6.3	60.0 ± 5.3	.001

Abbreviations: ASES, American Shoulder and Elbow Surgeons; BMI, body mass index; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcome Measurement Information System; SAS, Shoulder Activity Scale; SD, standard deviation; SST, Simple Shoulder Test; UE, upper extremity; VAS, Visual Analog Scale.

health have been shown to influence outcome measures in patients with GHOA, the correlation between radiographic severity of GHOA with traditional PROMs and PROMIS is not well defined.[23] Previous studies have noted that the size of the inferior osteophyte correlates with decreased functional range of motion.[6] However, patient pain, clinical function, and ability to perform daily activities are still considered to be the predominant indications to pursue arthroplasty. We demonstrate that the radiographic severity of GHOA does not correlate with pain scores or any of the other commonly used PROMs in this study.

Patient demographics of our cohort are representative of the typical population that develops GHOA. The majority of patients presenting for evaluation had more severe radiographic findings, although nearly 20% of patients were graded as Samilson–Prieto I or II. Mean PROs were also typical of what would be expected in a population presenting with OA, reporting higher pain and worse PF compared to the general U.S. population.[24]

When evaluating for differences in PROMs between categories of radiographic severity, we were unable to identify any significant differences. This includes measures looking both at clinical function (ASES, SST, SAS, PROMIS PF, and PROMIS UE) and pain (VAS, PROMIS PI). This expands on prior studies that showed poor correlation between radiographic severity and constant scores.[6] A lack of correlation between severity of radiographic OA and PROM scores contrasts to OA in the lower extremity, where increasing severity of knee OA is associated with higher pain scores and lesser quality of life.[25] This discrepancy may be, in part, related to the different weight-bearing requirements among the joints. The knee is a weight-bearing joint that commonly withstands 2 to 3 times the force of body weight with routine activities.[26] Conversely, shoulder joint reaction force is <50% of body weight in most activities of daily living, with the maximum contact force experienced only 1.6 times body weight for a sit-to-stand task.[27] Patients may therefore be able to tolerate GHOA that is significantly worse radiographically before it causes symptoms that prompt physician evaluation. Additionally, while the Samilson–Prieto scale is a common radiographic classification of OA used in research, it does not account for loss of joint space, posterior subluxation of the humeral head, or glenoid erosion. It is possible that factors not assessed in the Samilson–Prieto classification play important factors in function and may explain the absence of correlation between radiographic severity and PROM scores. Future investigations utilizing classification systems that assess deformity in other planes, such as the Walch classification, or joint space narrowing, such as the Kellgren–Lawrence classification, may identify a

correlation between radiographic deformity and outcome scores.[28]

Interestingly, males reported statistically significantly higher function and lower pain scores than females in our cohort. However, with large patient cohorts, it is important to distinguish between statistical significance and clinical relevance. Prior studies of shoulder arthritis populations have identified minimal clinically important differences (MCIDs) of 13.6 points for the ASES, 1.5 points for the SST, and 1.6 points for the VAS.[29,30] No studies have yet been performed to establish the MCID for the PROMIS instruments in a shoulder arthritis population, although several studies have identified MCIDs in the range of 3.5 to 5.0 points for the PROMIS instruments in other disease processes.[31–33] Using these MCID values as proxies for clinical relevance shows that despite the statistically significant differences in outcome measures between males and females, only the SST and PROMIS PF values are likely to be clinically relevant. Similarly, the few statistically significant differences in outcome measures based on dominant extremity involvement, prior surgery, and BMI over 30 kg/m^2 are likely not clinically relevant. We were unable to identify specific parameters within the SST and PROMIS PF questionnaires that would explain why these instruments identified what are likely clinically significant differences between genders while the other instruments did not.

A major strength of this study is the relatively large number of patients included in the cohort. To our knowledge, this is the largest study evaluating the association of PROs and radiographic severity of GHOA. Study limitations include a retrospective design and cross-sectional analysis. The retrospective nature required identification of patients via an ICD-10 code corresponding to GHOA. This may have led to the exclusion of patients whose visits were not coded appropriately. The cross-sectional design allows us to determine association but not causation. Additionally, we are unable to draw conclusions on responsiveness to treatment, such as whether or not radiographic severity influences outcomes following operative intervention. Another limitation is that we did not assess shoulder function directly, such as range of motion or strength, and advanced imaging was not routinely performed and thus we could not exclude other pain generators such as undiagnosed rotator cuff tears or biceps disease. Another limitation is the uneven distribution of patients in different radiographic severity groups. The majority of patients had more advanced GHOA (grade 3a or 3b), with relatively low number of patients evaluated with less radiographically severe (grade 1 or 2) GHOA. As such, the study may be underpowered to demonstrate differences in PROMs by radiographic severity of GHOA and raises the possibility of a Type II error.

It is possible that patients with less severe radiographic OA may be less symptomatic and less commonly warrant evaluation by a shoulder and elbow surgeon. Interestingly, however, those patients with less radiographically severe OA who were symptomatic enough to present to our clinic for evaluation, there was no difference in patient-reported pain and function compared to those with more severe radiographic findings. Further investigations currently underway are evaluating the association between preoperative radiographic severity of GHOA and outcomes following operative intervention.

Conclusion

Radiographic severity of GHOA does not correlate with patient-reported pain and function, and patient demographic factors do not show a correlation that is clinically relevant. While radiographs are important to identify the etiology of shoulder pain and help inform treatment options, patient pain and clinical function should remain the predominant indications for surgery regardless of the severity of radiographic arthrosis of the joint.

ORCID iDs

Eitan M Kohan https://orcid.org/0000-0003-1326-1747
Jay D Keener https://orcid.org/0000-0002-1665-4346

References

1. Luime JJ, Koes BW, Hendriksen IJ, et al. Prevalence and incidence of shoulder pain in the general population: a systematic review. *Scand J Rheumatol.* 2004;33:73–81.

2. Jacobs JJ, et al. United States Bone and Joint Decade: The Burden of Musculoskeletal Diseases in the United States. 2nd ed. Rosemont, IL: American Academy of Orthopaedic Surgeons; 2010.

3. Felson DT. An update on the pathogenesis and epidemiology of osteoarthritis. *Radiol Clin North Am.* 2004;42:1–9.

4. van der Heijden GJ. Shoulder disorders: a state-of-the-art review. *Baillieres Best Pract Res Clin Rheumatol.* 1999;13:287–309.

5. Virta L, Joranger P, Brox JI, Eriksson R. Costs of shoulder pain and resource use in primary health care: a cost-of-illness study in Sweden. *BMC Musculoskelet Disord.* 2012;13:17.

6. Kircher J, Morhard M, Magosch P, Ebinger N, Lichtenberg S, Habermeyer P. How much are radiological parameters related to clinical symptoms and function in osteoarthritis of the shoulder? *Int Orthop.* 2010;34:677–681.

7. Guyette TM, Bae H, Warren RF, Craig E, Wickiewicz TL. Results of arthroscopic subacromial decompression in patients with subacromial impingement and glenohumeral degenerative joint disease. *J Shoulder Elbow Surg.* 2002;11:299–304.

8. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis.* 1957;16:494–502.

9. Neer CS II. Replacement arthroplasty for glenohumeral osteoarthritis. *J Bone Joint Surg Am.* 1974;56:1–13.

10. Samilson RL, Prieto V. Dislocation arthropathy of the shoulder. *J Bone Joint Surg Am.* 1983;65:456–460.

11. Boenisch U, Lembcke O, Naumann T. Classification, clinical findings and operative treatment of degenerative and posttraumatic shoulder disease: what do we really need to know from an imaging report to establish a treatment strategy? *Eur J Radiol.* 2000;35:103–118.

12. Prescher A. Anatomical basics, variations, and degenerative changes of the shoulder joint and shoulder girdle. *Eur J Radiol.* 2000;35:88–102.

13. Walch G, Boulahia A, Boileau P, Kempf JF. Primary glenohumeral osteoarthritis: clinical and radiographic classification. The Aequalis Group. *Acta Orthop Belg.* 1998;64:46–52.

14. Bayley KB, London MR, Grunkemeier GL, Lansky DJ. Measuring the success of treatment in patient terms. *Med Care.* 1995;33:AS226–AS235.

15. Wright RW, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg.* 2010;18:436–444.

16. Wylie JD, Beckmann JT, Granger E, Tashjian RZ. Functional outcomes assessment in shoulder surgery. *World J Orthop.* 2014;5:623–633.

17. Richards RR, An KN, Bigliani LU, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg.* 1994;3:347–352.

18. Lippitt S, Harryman D, Matsen F. A practical tool for evaluating function: the Simple Shoulder Test. In: Matsen F, Fu F, Hawkins R, eds. *The Shoulder: A Balance of Mobility and Stability.* Rosemont, IL: The American Academy of Orthopaedic Surgeons; 1993: 545–559.

19. Brophy RH, Beauvais RL, Jones EC, Cordasco FA, Marx RG. Measurement of shoulder activity level. *Clin Orthop Relat Res.* 2005;439:101–108.

20. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care.* 2007;45:S3–S11.

21. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol.* 2010;63:1179–1194.

22. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess.* 1994;6:284–290.

23. Kohan EM, Hill JR, Schwabe M, Aleem AW, Keener JD, Chamberlain AM. The influence of mental health on Patient-Reported Outcomes Measurement Information System (PROMIS) and traditional outcome instruments in patients with symptomatic glenohumeral arthritis. *J Shoulder Elbow Surg.* 2018;28:e40–e48.

24. Minoughan CE, Schumaier AP, Fritch JL, Grawe BM. Correlation of PROMIS Physical Function Upper Extremity Computer Adaptive Test with American Shoulder and Elbow Surgeons shoulder assessment form and Simple Shoulder Test in patients with shoulder arthritis. *J Shoulder Elbow Surg.* 2018;27:585–591.

25. Pereira D, Severo M, Santos RA, et al. Knee and hip radiographic osteoarthritis features: differences on pain, function and quality of life. *Clin Rheumatol.* 2016;35:1555–1564.

26. Kutzner I, Heinlein B, Graichen F, et al. Loading of the knee joint during activities of daily living measured in vivo in five subjects. *J Biomech.* 2010;43:2164–2173.

27. Klemt C, Prinold JA, Morgans S, et al. Analysis of shoulder compressive and shear forces during functional activities of daily life. *Clin Biomech (Bristol, Avon).* 2018;54:34–41.

28. Walch G, Badet R, Boulahia A, Khoury A. Morphologic study of the glenoid in primary glenohumeral osteoarthritis. *J Arthroplasty.* 1999;14:756–760.

29. Simovitch R, Flurin PH, Wright T, Zuckerman JD, Roche CP. Quantifying success after total shoulder arthroplasty: the minimal clinically important difference. *J Shoulder Elbow Surg.* 2018;27:298–305.

30. Wong SE, Zhang AL, Berliner JL, Ma CB, Feeley BT. Preoperative patient-reported scores can predict

postoperative outcomes after shoulder arthroplasty. *J Shoulder Elbow Surg*. 2016;25:913–919.

31. Chen RE, Papuga MO, Voloshin I, et al. Preoperative PROMIS scores predict postoperative outcomes after primary ACL reconstruction. *Orthop J Sports Med*. 2018;6:2325967118771286.

32. Gausden EB, Levack A, Nwachukwu BU, Sin D, Wellman DS, Lorich DG. Computerized adaptive testing for patient reported outcomes in ankle fracture surgery. *Foot Ankle Int*. 2018;39:1192–1198.

33. Ho B, Houck JR, Flemister AS, et al. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. *Foot Ankle Int*. 2016;37:911–918.