J Shoulder Elbow Surg (2022) 31, e332–e345

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

A validated algorithm using current literature to judge the appropriateness of anatomic total shoulder arthroplasty utilizing the RAND/UCLA appropriateness method

Stephen Le Breton, BS[a,b,*], Stephen Sylvia, MD[a], Sundeep Saini, DO[a], Albert Mousad, BS[a,b], Matthew Chilton, BS[a,b], Sonia Lee, BS[a,b], Lambert Li, BS[a], Meghan MacAskill, BS[a], Glen Ross, MD[a], Joseph Gentile, MD[c], Randall J. Otto, MD[d], Scott G. Kaar, MD[d], Sridhar Pinnamaneni, MD[e], Andrew Jawa, MD[a], Jacob Kirsch, MD[a], Gabriella Ode, MD[f], William Aibinder, MD[g], R. Michael Greiwe, MD[h], Joseph DeAngelis, MD[a], Joseph J. King, MD[i], Sarav S. Shah, MD[a]

[a]*New England Baptist Hospital, Boston, MA, USA*
[b]*Tufts University School of Medicine, Boston, MA, USA*
[c]*Novant Health Orthopedics & Sports Medicine, Huntersville, NC, USA*
[d]*Department of Orthopaedic Surgery, St. Louis University, St. Louis, MO, USA*
[e]*Signature Medical Group, St Louis, MO, USA*
[f]*Department of Orthopaedic Surgery, Prisma Health-Upstate, Greenville, SC, USA*
[g]*Department of Orthopaedic Surgery and Rehabilitation Medicine, SUNY Downstate Health Sciences University, Brooklyn, NY, USA*
[h]*OrthoCincy Orthopaedics & Sports Medicine, St Elizabeth Healthcare, The Orthopaedic Research & Innovation Foundation, Inc., Edgewood, KY, USA*
[i]*Department of Orthopaedic Surgery and Rehabilitation, University of Florida, Gainesville, FL, USA*

Background: Currently, appropriateness criteria evaluating when to perform total shoulder arthroplasty (TSA) is lacking. In the absence of society guidelines and limited quality evidence, the RAND/University California in Los Angeles (UCLA) method provides a suitable alternative to evaluate appropriateness and assist in clinical decision making. Given the rise in utilization, appropriateness criteria for TSA have the potential to be an extremely powerful tool for improving quality of care and controlling costs. Thus, the goal of this study was to test explicit criteria to assess the appropriateness of TSA decision making using the RAND/UCLA appropriateness method.
Methods: A review of recent scientific literature to gather available evidence about the use, effectiveness, efficiency, and the risks involved in surgical intervention was performed by a shoulder/elbow fellowship trained physician. Based on pertinent variables including age, rotator cuff status, previous surgical management, mobility, symptomatology, and imaging classifications, 186 clinical

Investigation performed at New England Baptist Hospital, Boston, MA, USA.
Institutional review board approval was not required for this survey study.

*Reprint requests: Stephen Le Breton, BS, New England Baptist Hospital, 125 Parker Hill Ave Boston, MA 02120, USA.
E-mail address: Stephen.le_breton@tufts.edu (S. Le Breton).

https://doi.org/10.1016/j.jse.2021.12.025

scenarios were created. Appropriateness criteria for TSA were developed using a modified Delphi method with a panel consisting of American Shoulder and Elbow Surgeons (ASES) members. A second panel of ASES members rated the same scenarios, with reliability testing performed to compare groups.

Results: Panel members reached agreement in 40 (64%) indications. TSA was appropriate in 15 (24%) of indications. For patients with severe symptomatology, TSA was often appropriate for patients aged <75 years and inconclusive or inappropriate for patients aged >75 years. Among patients aged <65 years, TSA varied between appropriate and inconclusive, often dependent on Walch classification. For patients with moderate symptomatology, TSA was inappropriate or inconclusive for patients aged <65 or >75 years. When compared to the second panel's results, moderate agreement was obtained with a weighted kappa statistic of 0.56.

Conclusions: Using the RAND/UCLA method, ASES members created an appropriateness decision tree for pertinent patient variables. This presents the data in a manner that streamlines the clinical decision-making process and allows for rapid and more reliable determination of appropriateness for practitioners. The decision tree is based on a combination of clinical experience from high-volume ASES-member surgeons and a comprehensive review of current evidence. This tool can be used as part of a broader set of factors, including individual patient characteristics, prior studies, and expert opinion, to inform clinical decision making, improve quality of care, and control costs.

Level of evidence: Survey Study using Experts; Delphi Method

Keywords: Shoulder; total shoulder arthroplasty; TSA; appropriateness; RAND/UCLA; algorithm; decision tree; Delphi technique

Over the past decade, there has been a dramatic rise in the number of anatomic total shoulder arthroplasties (TSAs) performed, with approximately 41,000 completed each year.[40] Although TSA is a cost-effective operation, there is high variability in the way patients may present, and some TSA-eligible patients may be better nonoperative candidates or more suitable for other procedures such as reverse shoulder arthroplasty (RSA).[31,45] There is limited quality evidence regarding efficacy to justify appropriateness of surgical interventions for TSA and a dearth of guidelines providing clear criteria for appropriateness of TSA.

Any determination of the extent to which TSA surgery is appropriate requires the use of valid appropriateness criteria. Researchers from the RAND Corporation and the University California in Los Angeles (UCLA) developed a modified Delphi method that combines available scientific evidence with expert opinion.[5] This has been widely used to assess the appropriateness of other orthopedic interventions, including hip replacement, knee replacement, lumbar laminectomy, lower back pain, and lumbar disc herniation.[11,20,32,33,35,38] Currently, appropriateness criteria evaluating when to perform TSA surgery is lacking.[16] In the absence of society guidelines and limited quality evidence, the RAND/UCLA method provides a suitable alternative to evaluate appropriateness and assist in clinical decision making.[38]

Given the rise in utilization, appropriateness criteria for TSA have the potential to be an extremely powerful tool for improving quality of care and controlling costs. Thus, the goal of this study was to test explicit criteria to assess the appropriateness of TSA decision making using the RAND/UCLA appropriateness method.

Materials and methods

Overview of the RAND/UCLA method

This is a study determining appropriateness of total shoulder arthroplasty across various indications using the RAND/UCLA method. The basic methodology of the RAND/UCLA method has been described by Fitch et al[12] in the *RAND/UCLA Appropriateness Method User's Manual*. First, a detailed literature review is performed to synthesize the latest evidence on the procedure to be rated. A list of clinical scenarios that categorize patients based on relevant clinical factors (eg, symptoms, imaging) are generated. A panel of experts on the specific procedure is formed. Panelists are provided with the literature review and the list of indications and will rate the benefit-to-harm ratio of the procedure on a scale of 1-9, with 1 indicating the harms outweigh the benefits and 9 indicating the benefits outweigh the risk.

The panelists rate each of the indications twice, in a 2-round ''modified Delphi'' process. In the first round, panelists rate the indications individually. In the second round, panelists meet in a group alongside a moderator to discuss their individual ratings, focusing on areas of disagreement, and are given the opportunity to modify the original list of indications. Finally, each indication is classified as ''appropriate,'' ''inconclusive," or ''inappropriate'' for the procedure under review in accordance with the panelists' median score and the level of disagreement among the panelists. Indications with median scores of 1-3 are classified as inappropriate, 4-6 as inconclusive, and 7-9 as appropriate. All indications rated ''with disagreement,'' irrespective of the median value, are classified as inconclusive. ''Disagreement'' indicates a lack of consensus, often due to group polarization (eg, half of the panel rating as ''inappropriate'' whereas the other half rates as ''appropriate''). Various methods of identifying ''disagreement'' have been cited in prior literature, are study-specific, and depend on the composition of the panel.[11,33]

Table I Complete bibliography provided to panelists

First author	Date	Journal	Title	Key takeaways
Concern for age, RTC tear/complications post TSA and outcomes vs. RSA				
Polisetty[30]	2021	*JBJS*	Value Analysis of Anatomic and Reverse Shoulder Arthroplasty for Glenohumeral Osteoarthritis with an Intact Rotator Cuff	• Minimum 2-yr f/u comparison of 252 TSAs vs. 63 RSAs • Total hospitalization costs, charges, and reimbursements along with outcome improvements in units of MCIDs and patient satisfaction did not differ between groups
Young[45]	2012	*JBJS*	Secondary Rotator Cuff Dysfunction Following Total Shoulder Arthroplasty for Primary Glenohumeral Osteoarthritis: Results of a Multicenter Study With More Than Five yr of Follow-up	• 16.8% of patients had secondary cuff dysfunction at an average 8.6-yr follow-up and exhibit worse clinical and radiographic outcomes compared to those with a normal functioning rotator cuff • Survivorship free of secondary cuff dysfunction was 100% at 5 yr, 84% at 10 yr
Wright[44]	2020	*JAAOS*	Comparison of Clinical Outcomes After Anatomic Total Shoulder Arthroplasty and Reverse Shoulder Arthroplasty in Patients 70 yr and Older With Glenohumeral Osteoarthritis and an Intact Rotator Cuff	• No difference in PROMs (ASES or Western Ontario Osteoarthritis of the Shoulder) comparing TSA to RSA in patients with OA aged >70 yr • Self-reported to be underpowered • Compared mean values with high standard deviations for the PROMs, thus leaving the comparison vulnerable to only a small percentage of outcome events needed to change the study's statistical significance
Poondla[31]	2021	*JSES*	Anatomic and Reverse Shoulder Arthroplasty in Patients 70 yr of Age and Older: A Comparison Cohort at Early to Midterm Follow-up	• Final ASES scores and improvement from preoperative ASES score between the age groups were not significantly different (TSA $P = .520$; RSA $P = .065$). • Given satisfactory results following TSA in patients aged ≥70 yr, the authors do not routinely perform reverse total shoulder arthroplasty for primary osteoarthritis with an intact rotator cuff solely based on age.
Shah[37]	2021	*Orthopedics*	The Comparative Effect of Age on Clinical Outcomes Following Anatomic Total Shoulder Arthroplasty and Reverse Total Shoulder Arthroplasty	• In patients aged ≥75 yr, the percentage of patients achieving MCID were similar; however, a higher percentage of TSA patients achieved Substantial Clinical Benefit vs. RSA (90.5% vs. 76.9%; $P = .01$). • "This study highlights the importance of indications and a shared decision-making model to ensure patient satisfaction"
Brewley[4]	2020	*JSES*	Defining the Younger Patient: Age as a Predictive Factor for Outcomes in Shoulder Arthroplasty	• In patients aged <65 yr, TSA was associated with a 3.4-fold increased risk of revision • TSA (n = 518; mean age, 68.1 yr)
Simovitch[39]	2017	*JBJS*	Rate of Improvement in Clinical Outcomes With Anatomic and Reverse Total Shoulder Arthroplasty	• TSA more reliably improves range of motion • RSA had significantly less improvement in active and passive external rotation
Levy[22]	2014	*JSES*	Speed of Recovery After Shoulder Arthroplasty: A Comparison of Reverse and Anatomic Total Shoulder Arthroplasty	• TSA can anticipate more consistent and effective recovery of pain, function, and shoulder rotation • RSA can expect a variable length of recovery

(*continued on next page*)

Table I Complete bibliography provided to panelists (*continued*)

First author	Date	Journal	Title	Key takeaways
Raiss[34]	2012	*JBJS*	Results of Cemented Total Shoulder Replacement With a Minimum Follow-up of 10 yr	• Survivorship of glenoid was 100% after 13 yr with revision as end point
Schoch[36]	2019	*JSES*	Glenoid Component Lucencies Are Associated With Poorer Patient-Reported Outcomes Following Anatomic Shoulder Arthroplasty	• Peri-implant glenoid lucencies following TSA are associated with lower forward elevation and PROs • Lucencies >grade 2 associated with clinically important losses in overhead motion
Neyton[26]	2019	*JSES*	Mid- to Long-Term Follow-up of Shoulder Arthroplasty for Primary Glenohumeral Osteoarthritis in Patients Aged 60 or Under	• TSA resulted in significantly better functional and subjective outcomes with no significant difference in survivorship compared with patients treated with HA
Leschinger[21]	2017	*AOTS*	Predictors of Medium-term Clinical Outcomes After Total Shoulder Arthroplasty	• Higher Walch classification is a negative predictor for postoperative clinical function • "Emphasizes the importance of the preoperative evaluation of humeral head subluxation and glenoid erosion"
Radiographic parameters in shoulder osteoarthritis				
Kircher[19]	2010	*Int Orthop*	How Much Are Radiological Parameters Related to Clinical Symptoms and Function in Osteoarthritis of the Shoulder?	• Joint space narrowing and development of osteophytes are reliable but independent parameters of primary shoulder arthrosis • Size of the caudal humeral osteophyte is a predictive factor for function • Samilson classification can be meaningful supplemented with measurements of joint space width and the size of osteophytes • "In our opinion, patients should be carefully informed about possible disadvantages in postponing such an operation when in advanced stages of osteoarthritis." • Primary clinical feature, pain, as the main indication for surgery is not related to radiologic parameters
Glenoid bone loss				
Walker[42]	2018	*JBJS*	Progression of Glenoid Morphology in Glenohumeral Osteoarthritis	• Asymmetric bone loss rarely develops in A1 glenoids, whereas initial posterior translation of the humeral head (B1 glenoids) may be associated with subsequent development and progression of posterior glenoid bone loss over time
Donohue[10]	2018	*JBJS*	The Association Between Rotator Cuff Muscle Fatty Infiltration and Glenoid Morphology in Glenohumeral Osteoarthritis	• Significant difference in the Goutallier score for the supraspinatus, infraspinatus, and teres minor muscles between Walch subtypes • Higher fatty infiltration of the infraspinatus, teres minor, and combined posterior rotator cuff muscles was associated with increasing glenoid retroversion
Aleem[1]	2019	*JBJS*	Association Between Rotator Cuff Muscle Size and Glenoid Deformity in Primary Glenohumeral Osteoarthritis	• Asymmetric glenoid wear and humeral-head subluxation in osteoarthritis are associated with asymmetric atrophy within the rotator cuff transverse plane.

(*continued on next page*)

Table I Complete bibliography provided to panelists (*continued*)

First author	Date	Journal	Title	Key takeaways
				• Increased posterior rotator cuff muscle area compared with anterior rotator cuff muscle area is associated with greater posterior glenoid wear and subluxation.
Updated Walch classification				
Bercik[2]	2016	*JSES*	A modification to the Walch classification of the glenoid in primary glenohumeral osteoarthritis using three-dimensional imaging	• A1—centered humeral head, minor erosion • A2—centered humeral head, major central glenoid erosion • B1—posterior subluxated head, no bony erosion • B2—posterior subluxated head, posterior erosion with biconcavity of the glenoid • B3—glenoid as monoconcave and posteriorly worn, with at least 15° of retroversion or at least 70% posterior humeral head subluxation, or both. (B3 glenoid with posterior subluxation without significant retroversion differs from the B1 by the presence of posterior wear.) • C—dysplastic glenoid with at least 25° of retroversion regardless of erosion • D—any level of glenoid anteversion or with humeral head subluxation of less than 40% (ie, anterior subluxation)

RTC, rotator cuff; *TSA*, total shoulder arthroplasty; *RSA*, reverse shoulder arthroplasty; *JBJS*, *Journal of Bone and Joint Surgery*; *MCIDs*, minimal clinically important differences; *JAAOS*, *Journal of the American Academy of Orthopaedic Surgeons*; *PROMs*, patient-reported outcome measures; *ASES*, American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form; *JSES*, *Journal of Shoulder and Elbow Surgery*; *HA*, hemiarthroplasty; *AOTS*, *Archives of Orthopaedic and Trauma Surgery*; *Int Orthop*, *International Orthopaedics*.

Following the 2-round modified Delphi process, authors may use patient data to apply appropriateness criteria retrospectively and measure procedure overuse, or apply necessity criteria to measure underuse.[12,33] Of note, neither add-on approach was performed in this study and are ongoing in a follow-on study. Instead, the authors opted to include a second panel, which performed a single-round individual rating that was used to evaluate inter-rater reliability.[11,33]

Bibliography review

A comprehensive review of recent literature from 2010 to 2020 was performed by a board-certified ASES fellowship director and Shoulder and Elbow Fellowship–trained physician to gather available evidence about the use, effectiveness, efficiency, and risks involved in surgical intervention. The search terms used were *total shoulder arthroplasty*, *complications*, and *patient-reported outcomes*. The purpose of the summarized review was to provide the panel members with an updated and uniform source of information that would assist in rating the scenarios and supplement the panelists' extensive clinical experience. Please see Table I for the complete bibliography provided to panelists.

Explicit appropriateness criteria scenario development

A list of clinically relevant variables were identified from the literature review (Table II) and used to write an extensive and mutually exclusive set of 62 clinical indications based on the number of variables for TSA. Variables included age, rotator cuff (RTC) status, previous surgical interventions, joint mobility, symptomology, and imaging classifications (Samilson & Prieto, Walch).

Age was separated into 3 groups (<65 years, 65-75 years, >75 years) based on recent literature evaluating outcomes of shoulder arthroplasty according to age.[4,37] RTC status/prior open surgery was included given prior literature demonstrating diminished functional improvements in patients with previous ipsilateral shoulder surgery, but was only applicable to 2 scenarios exploring patients with full-thickness RTC tears or prior RTC repair/open Bankart repairs. Walch classification was included given glenoid morphology has been associated with the progression of symptoms and outcomes following TSA.[2,18,19,41] Although SP classification has been shown not to correlate with symptomatology, it was included for completeness and historical use in informing TSA.[18,19] The full list of criteria can be found in Table II.

Three clinically equivalent scenarios were generated per indication, for a total of 186 scenarios. Three scenarios per indication were used to calculate intrarater reliability. Several example scenarios have been provided in Figure 1.

Panelist rating

Following development of the clinical scenarios, a national panel of 5 American Shoulder and Elbow Surgeons (ASES) candidate level or above members was convened (panel 1). The panel consisted of 2 shoulder fellowship directors, a chief of shoulder surgery at a high-volume urban center, and 2 shoulder specialists performing a high volume of TSA at urban academic centers. Panelists were provided with a synopsis of the literature review (Table I), the list of all the variables, and the list of 186 patient scenarios consisting of all variables. Then, each panelist individually rated each of the 186 scenarios for the appropriateness of TSA. In accordance with prior literature, appropriateness was defined as a situation in which the expected health benefit exceeds the expected negative consequences by a sufficiently wide margin to make TSA worth performing.[33]

Appropriateness ratings were scored on a 9-point scale. The use of TSA for a specific indication was considered appropriate if the panel's mean score (rounded to the nearest whole number) was between 7 and 9 without disagreement, inconclusive if the panel's mean score was between 4 and 6 or any mean with disagreement, and inappropriate if the mean rating was between 1 and 3 without disagreement. Additionally, agreement was established if 4 panelists scored in the same category, or fewer than 3 panelists scored outside the category of the mean value. Disagreement was considered when 2 or more panelists scored an indication between 1 and 3 and at least another 2 rated it between 7 and 9. An agreement rating of ''Indeterminate'' encompassed all other situations (Fig. 2).

The scenario ratings took place in 2 rounds using a modified Delphi method. The first round was performed confidentially prior to the primary panel meeting, with each of the 5 panel members rating all 186 scenarios. Ratings were collected via an online survey. In the second round, the results of the first round were presented to the panelists at a 1-day meeting in order to establish the level of agreement among the panelists. Prior to the meeting, each panelist received a comprehensive package that included the literature review summary, unblinded ratings from the 5 panelists, and the average rating for each scenario. Panelists discussed 19 scenarios where disagreement occurred. After extensive discussion, the 5 panelists revised a total of 25 of 930 scenario ratings to align with the aforementioned definition of ''appropriateness.'' A total of 6 scenarios changed appropriateness rating following the revision.

Then, a second national panel of 7 ASES candidate level or above members was convened (panel 2), with the purpose of using the results to compare with those from the primary panel for the reliability analysis. Panel 2 consisted of 1 director of orthopedic research at a high-volume institution, 1 shoulder fellowship director, an internationally trained shoulder specialist, and 4 surgeons practicing at high-volume urban centers. A nearly identical process was followed for the second panel with the exception of a follow-up meeting.[11,33] As a result, an additional 2 members were included in the second panel. Scenarios were scored on a 9-point scale for appropriateness. Appropriateness was determined if 5 of 7 panelists were in agreement of the category. Agreement was established when 5 of 7 panelists scored in the same category or fewer than 4 panelists scored outside the category of the mean value. Disagreement was considered when 3 or more panelists scored an indication between 1 and 3 and at least another 3 rated it between 7 and 9. An agreement rating of ''Indeterminate'' encompassed all other situations.

Once all scenarios had been rated by both panels, appropriateness was determined for each indication. The median value across each of the 3 scenarios was used to form a single appropriateness rating for each indication for each panelist. For example, a panelist that provided ratings of 7, 8, and 9 for the 3 scenarios of a given indication would have an overall score of 8 for that indication. The same 9-point scale and agreement-determination methodology as described above were used for each panel.

Statistical analysis

Reliability studies were performed comparing the results from the primary and second panels, with weighted kappa statistics used to evaluate the level of agreement. Appropriateness ratings for each indication were compared between the 2 panels. Intrarater reliability was calculated to identify the accuracy with which panelists rated the 3 scenarios within a single indication. A generalized linear model was applied across the median panelists' ratings to determine validity. This model included a stepwise regression process to compare the relative contribution to the panel score variability of each indication variable taken into consideration, using likelihood ratio χ^2. Statistics were conducted using SPSS 28.0 software (IBM Corp., Armonk, NY, USA).

A decision tree for all patients was also created to present the data in a manner that allows for rapid determination of appropriateness for practitioners (Figs. 2-4). Appropriateness for each indication was calculated using indication-level appropriateness ratings from panel 1.

Results

Reliability analysis

Of the 62 clinically relevant indications evaluated by the primary panel, 15 (24%) were characterized as appropriate use of TSA. Panel members reached agreement in 40 (64%) indications. Twenty indications were characterized as inappropriate, with panel members reaching agreement in 18 (90%) of those indications. The remaining 27 indications (44%) were inconclusive (Table III).

Of the 186 scenarios scored by the primary panel, 53 (28%) were characterized as appropriate use of TSA. Panel members reached agreement in 129 (69%) of all scenarios. Sixty-eight (37%) were characterized as inappropriate, with panel members reaching agreement in 65 (96%) of those scenarios. The remaining 65 scenarios (35%) remained inconclusive (Table III).

Table II Variables of the algorithm

Age
<65; 65-75; >75 yr old
Rotator cuff (RTC) status/previous surgical management
Full-thickness RTC tear
Prior open or arthroscopic RTC repair/prior Bankart repair
Intact RTC/no prior management/no prior open or arthroscopic procedure for rotator cuff repair
Mobility
Preserved mobility:
Equivalent to a minimum active range of movement from 140° to 160° of forward flexion, 140° of abduction, 40° of external rotation, and internal rotation to the L4 spinous process or higher.
Limited mobility:
Equivalent to an active range of movement of 100°-140° of forward flexion, 100°-140° of abduction, <40° of external rotation, and internal rotation to the L5 spinous process or lower.
Symptomatology
Slight:
Sporadic in nature
Allows daily activities to be carried out (those requiring great physical effort may be limited);
Medication: aspirin, acetaminophen, nonsteroidal anti-inflammatory drugs (NSAIDs) control the pain without any side effects
Moderate:
Occasional pain
Some limitation of daily activities
Medication: aspirin, acetaminophen, NSAIDs control the pain with no/few side effects
Severe:
Pain of almost continuous nature
Daily activities are limited significantly, if not constantly
Medication: continuous use of NSAIDs or analgesic-narcotics for treatment to take effect
Radiographic presentation (Samilson and Prieto [SP] classification)
SP grade I: <3 mm
SP grade II: 3-7 mm
SP grade III: >7 mm
Computed tomography, Walch classification
Walch A1: Centered humeral head, minor erosion
Walch A2: Centered humeral head, major central glenoid erosion in which a line drawn from the anterior to posterior rims of the native glenoid transects the humeral head
Walch B1: Posterior subluxated head, no/minimal bony erosion of the glenoid
Walch B2: Posterior subluxated head, posterior erosion with biconcavity of the glenoid
Walch B3: At least 70% posterior humeral head subluxation, and/or monoconcave and posterior glenoid wear with at least 15° of retroversion
Walch C: Dysplastic glenoid with at least 25° of retroversion regardless of erosion
Walch D: Any level of glenoid anteversion or with humeral head subluxation of less than 40% (ie, anterior subluxation)

Intrarater reliability studies performed on the 186 scenarios scored by panel 1 demonstrated a high level of reliability between scenario ratings with the same indication, with an average 77% reliability across the 5 raters in primary panel. Following review of 19 scenarios with disagreement, 1 scenario previously rated as inconclusive was changed to appropriate, whereas 5 scenarios were changed from inconclusive to inappropriate. No scenarios changed from appropriate to inappropriate, or vice versa.

Panel 2 rated 186 scenarios, encompassing 62 indications, in a single round. The inter-rater reliability study of the indication-level ratings between the 2 panels gave a weighted kappa value of 0.56 (Table IV). There was no statistically significant difference in the number of indications that panel 2 rated as inappropriate (12 vs. 20; $P = .10$), inconclusive (34 vs. 27; $P = .21$), or appropriate (16 vs. 15; $P = .84$) relative to panel 1. However, panel 2 did rate a significantly lower number of scenarios as inappropriate (38 vs. 68; $P < .001$) and a significantly greater number of scenarios as inconclusive (97 vs. 65; $P = .001$) relative to panel 1. There was no significant difference in the number of scenarios rated as appropriate (51 vs. 53; $P = .82$).

Validity analysis

A generalized linear model (intercept-only default) was used to confirm criterion validity and assess the role that each clinical variable played in influencing panel 1's final appropriateness ratings (Table V). Age accounted for the

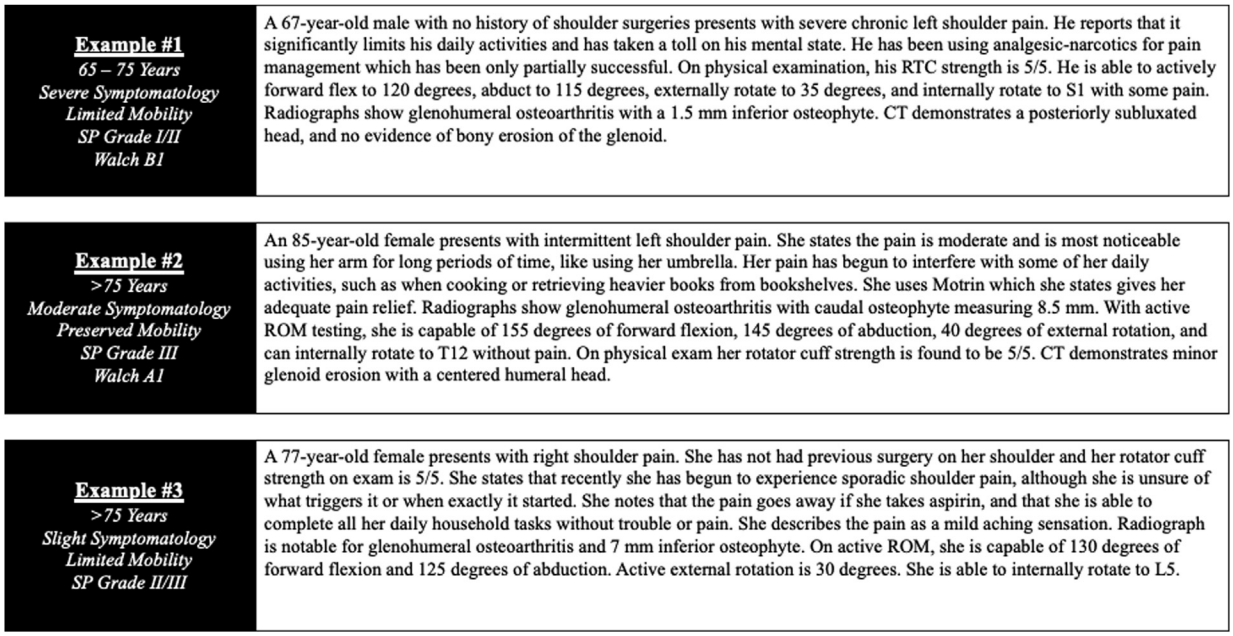

Figure 1 Three example scenarios for patients with slight, moderate, and severe symptomatology. The box on the left-hand side highlights which indication is applicable to the scenario. For example, the first scenario would be one of 3 scenarios applicable to a patient who was 65-75 years old with severe symptomatology, limited mobility, SP grade I/II, and Walch B1. Three similar scenarios were written for a given indication, with a total of 186 scenarios written for 62 clinically relevant indications.

most variance, followed by computed tomography, symptomatology, radiographic presentation, and mobility.

Synthesis of results

Findings from panel 1 were mapped to a decision tree. For patients with slight symptomatology, TSA was rated as either inconclusive or inappropriate (Fig. 3). Panelists did not perceive there to be any indications with slight symptomatology where TSA was definitively appropriate, reflecting that the perceived risk of harm in terms of complications was greater than the functional benefit that could be gained from the procedure for patients of slight symptomatology. For these indications, panelists suggested surgery may be premature and that cortisone injections may reasonably address the slight symptomatology. In the same category of slight symptomatology, use of TSA in patients with limited mobility that were either SP1 or SP2-3 were considered inappropriate. Again, in patients with mild symptomatology but also osteophytes, the panelists suggested trying alternative strategies (eg, comprehensive arthroscopic management) before TSA.

For patients with moderate symptomatology, appropriateness was correlated with age (Fig. 4). Among patients aged <65 years, use of TSA was considered either inappropriate or inconclusive, depending on mobility and radiographic presentation. In patients aged 65-75 years, use of TSA was either appropriate or inconclusive. The panelists cited pain and function as the 2 driving forces for surgery in this population, with an emphasis on activities of daily living. For patients rated inconclusive, the panelists were split between whether TSA was more appropriate than rTSA. In these patients, panelists leaned toward a more severe Walch classification (Walch ≥B2) as a cutoff for when to use rTSA rather than TSA. Additionally, panelists suggested that inconclusive indications may benefit from a more robust evaluation of the impact on activities of daily living to inform appropriateness of TSA, with greater impact resulting in higher appropriateness.

For patients with severe symptomatology, appropriateness of TSA was often appropriate for patients aged <75 years and inconclusive or inappropriate for patients aged >75 years (Fig. 5). Among patients aged <65 years, use of TSA varied between appropriate and inconclusive, often dependent on Walch classification. Inconclusive ratings were often driven by concern over patient age, with panelists suggesting a bridge to TSA (eg, comprehensive arthroscopic management procedure, injection). In patients aged 65-75 years, a majority of indications were deemed appropriate. In patients aged >75 years, use of TSA alternated between inappropriate and inconclusive. For patients in this group with SP1-2, Walch classification ≥B2 was deemed inappropriate, whereas all remaining indications were inconclusive, irrespective of Walch or mobility status. In patients with SP3, Walch A1, and limited mobility, TSA was deemed inappropriate.

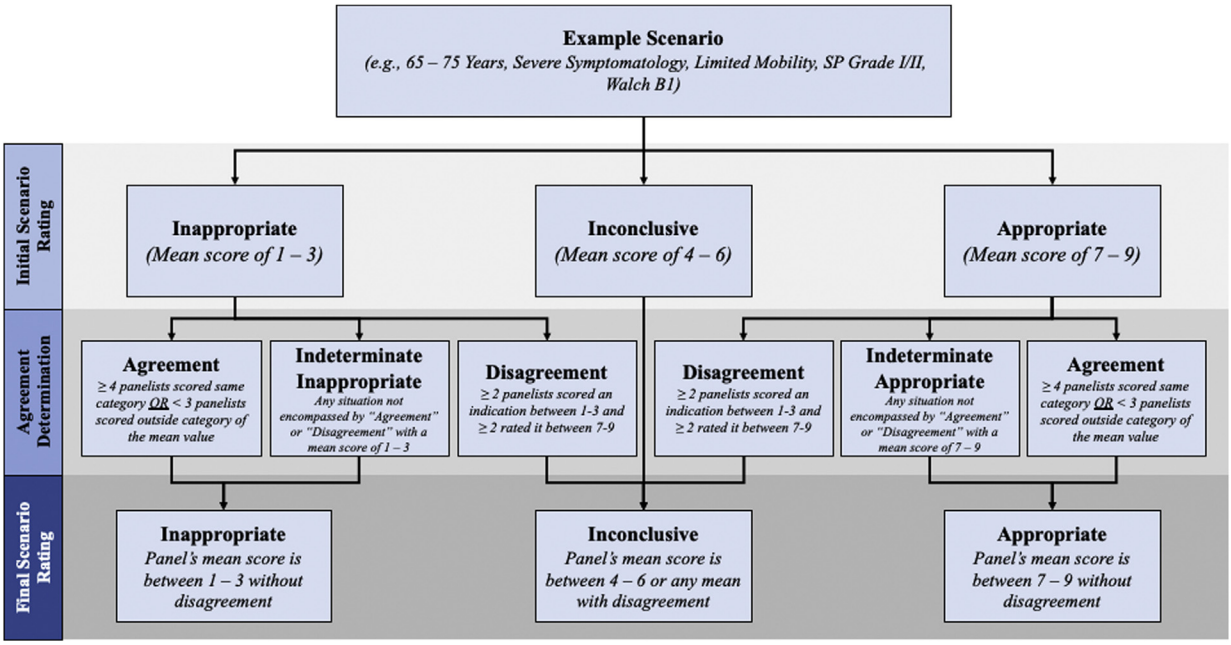

Figure 2 Flowchart demonstrating the rating methodology for each scenario, as well as the role of agreement/disagreement in determining the final indication rating. Agreement was established if 4 panelists scored in the same category, or fewer than 3 panelists scored outside the category of the mean value. Disagreement was considered when 2 or more panelists scored an indication between 1 and 3 and at least another 2 rated it between 7 and 9. An agreement rating of ''Indeterminate'' encompassed all other situations.

Discussion

Using the RAND/UCLA method, ASES members created an appropriateness decision tree for pertinent patient variables. The decision tree was formed based primarily on expert opinion that was further bolstered by a thorough updated literature review. Our results demonstrate the importance of age and Walch classification as objective criteria that inform appropriateness of TSA in the clinical setting, with these 2 factors being the greatest contributors to variability among panelists. Our study presents the data in a manner that allows for rapid and more reliable determination of appropriateness for practitioners. In the absence of national and society guidelines, the RAND method has been described in the past as a solution to evaluate appropriateness.[33,38] These results ultimately support the use of this decision tree as an aid in informing clinical practice guidelines for TSA, which may be a powerful tool for improving quality of care and controlling cost.[28] Guidelines that promote interventions of proven benefit and discourage ineffective ones have the potential to reduce morbidity and mortality and improve quality of life, at least for some conditions.[43] Guidelines can also improve the consistency of care, as the frequency with which procedures are performed varies dramatically among geographical regions.[8]

Age and symptomatology may be a dependent variable for outcomes following TSA.[4,14,15,24,26,31] Among patients aged <65 years with severe symptomatology, use of TSA varied between appropriate and inconclusive, often dependent on Walch classification. This indication for TSA in what may have been previously thought of as ''young'' patients may be related to recent literature supporting better implant survival and convertibility to RSA.[3,27,29] Additionally, posterior glenoid wear morphology has been implicated in the progression of symptoms and worsened outcomes following TSA.[21] A series of studies have shown the progression of bone loss over time and asymmetric glenoid wear and humeral head subluxation in osteoarthritis to be associated with asymmetric atrophy within the rotator cuff; thus, the idea of early operative management may be warranted.[1,42] TSA was often inconclusive or inappropriate for patients aged >75 years, irrespective of symptomatology. This may be related to current national trends that support RSA for patients aged >75 years, despite new evidence that suggests a higher percentage of TSA patients achieved substantial clinical benefit vs. RSA patients at 2-year follow-up.[37] Similarly, our results demonstrate the importance of symptomatology as a key subjective criterion to inform appropriateness in the clinical setting, especially given radiographic measures may not directly correlate with perceived pain and function.[18]

Our study demonstrates the changing role of imaging in determining appropriateness of TSA and is reflective of current evidence. Our results demonstrate that SP classification was not a statistically significant contributor to the variability among panelist scores. This is consistent with findings from Kirscher et al, which demonstrated that whereas the size of the humeral caudal osteophyte is a predictive factor for function, pain is the main clinical

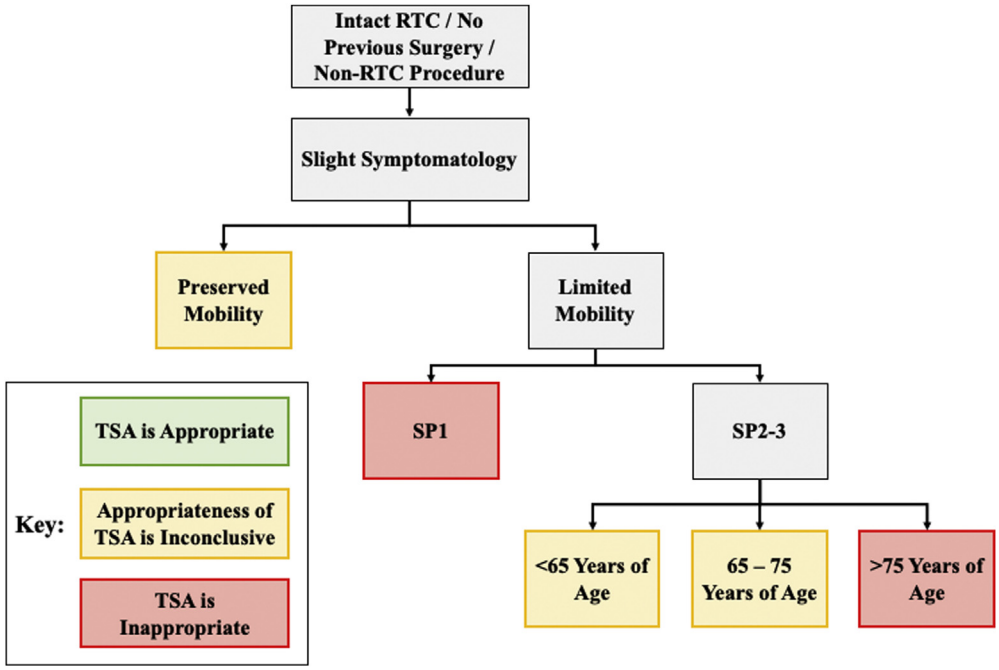

Figure 3 Decision tree for patients with slight symptomatology and an intact rotator cuff.

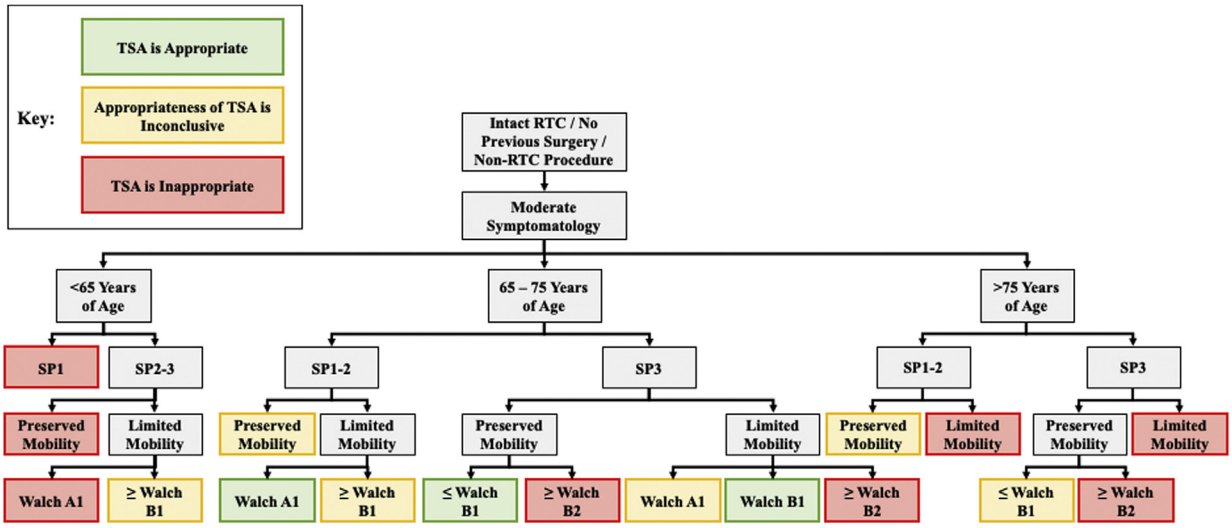

Figure 4 Decision tree for patients with moderate symptomatology and an intact rotator cuff.

indication for arthroplasty procedures and is not associated with radiologic parameters.[19] Additionally, Joyce et al[18] demonstrated that radiographic severity via SP classification may not parallel the pain levels of patients. As such, there may be a reduced clinical relevance of this classification as an objective criterion to determine appropriateness of TSA.

There was a high proportion of both scenarios and indications classified as inconclusive by panel 1 (35% and 44%, respectively) and panel 2 (52% and 55%, respectively); this is higher than comparable orthopedic appropriateness studies exploring nonshoulder procedures.[11,20,32,33,35,38] In select cases, indications that were expected to be rated inappropriate or appropriate were rated inconclusive. This skew toward inconclusive ratings may be driven by several factors. First, our study methodology required that agreement occur in order to achieve an appropriate rating. This step in the methodology was not included in comparable studies but was moderately counteracted by our definition of agreement.[11,33] Additionally, descriptions of the scenarios

Table III Appropriateness by agreement of panel 1*

	Appropriate, n (%)	Inconclusive, n (%)	Inappropriate, n (%)	Total, n (%)
Scenarios				
Agreement	53 (28.5)	11 (5.9)	65 (34.9)	129 (69.4)
Indeterminate	0 (0)	49 (26.3)	3 (1.6)	52 (28)
Disagreement	0 (0)	5 (2.7)	0 (0)	5 (2.7)
Total	53 (28.5)	65 (34.9)	68 (36.6)	186 (100)
Indications				
Agreement	15 (24.2)	7 (11.3)	18 (29)	40 (64.5)
Indeterminate	0	15 (24.2)	2 (3.2)	17 (27.4)
Disagreement	0	5 (8.1)	0 (0)	5 (8.1)
Total	15 (24.2)	27 (43.5)	20 (32.3)	62 (100)

* Panel 1 made up of 5 ASES candidate level or above members.

Table IV Comparison of the rating of the 2 panels*

	Panel 2			
	Appropriate, n (%)	Inconclusive, n (%)	Inappropriate, n (%)	Total, n (%)
Panel 1				
Appropriate	10 (16.1)	5 (8.0)	0 (0)	15 (24.1)
Inconclusive	6 (9.6)	20 (35.5)	1 (1.6)	27 (43.5)
Inappropriate	0 (0)	9 (14.5)	11 (17.7)	20 (32.2)
Total	16 (25.8)	34 (54.8)	12 (19.4)	62 (100)

Weighted kappa = 0.56.
* Panel 1 was made up of 5 ASES candidate level or above members; panel 2 was made up of 7 ASES candidate level or above members.

Table V Panel 1 validity study

Indication variables*,†	Generalized linear model (intercept-only)		
	Likelihood ratio χ^2	Difference	*P* value
Age	36.660		<.001‡
Computed tomography	68.651	31.991	<.001‡
Symptomatology	85.357	16.706	.004‡
Radiographic presentation	93.510	8.153	.160
Mobility	94.367	0.857	.026‡
Rotator cuff status	94.367	0.000	.150

* Evaluated the panel 1 median score as the dependent variable, with differences indicating the contribution of each indication variable to the likelihood ratio χ^2 statistic.
† All cumulative effects were significant at $\alpha = 0.01$.
‡ Indication variables were significant in independent univariate generalized linear model analyses with $\alpha = 0.05$.

were variable based on a range (eg, age of patients for scenarios in the >75-year age category ranged from ages 76-89 years), which may have altered the responses for some of the surgeons. All surgeons were supplied with the list of variables and decision tree in an attempt to avoid discrepancies. Finally, the high proportion of inconclusive ratings may be explained by continued controversy regarding appropriateness of TSA across a number of indications, despite a growing body of evidence exploring outcomes across age, symptomatology, prior surgery, and radiologic parameters.[4,13,14,19,23,24,26,31,34,36,44] Consistent with the literature, panelists qualitatively discussed bridging to surgery in patients with higher preoperative functioning and lower pain and may consider injections or less aggressive surgical interventions prior to TSA. Despite this qualitative perspective, our study found that appropriateness of TSA in patients with slight symptomatology and preserved mobility was inconclusive. Despite the higher proportion of

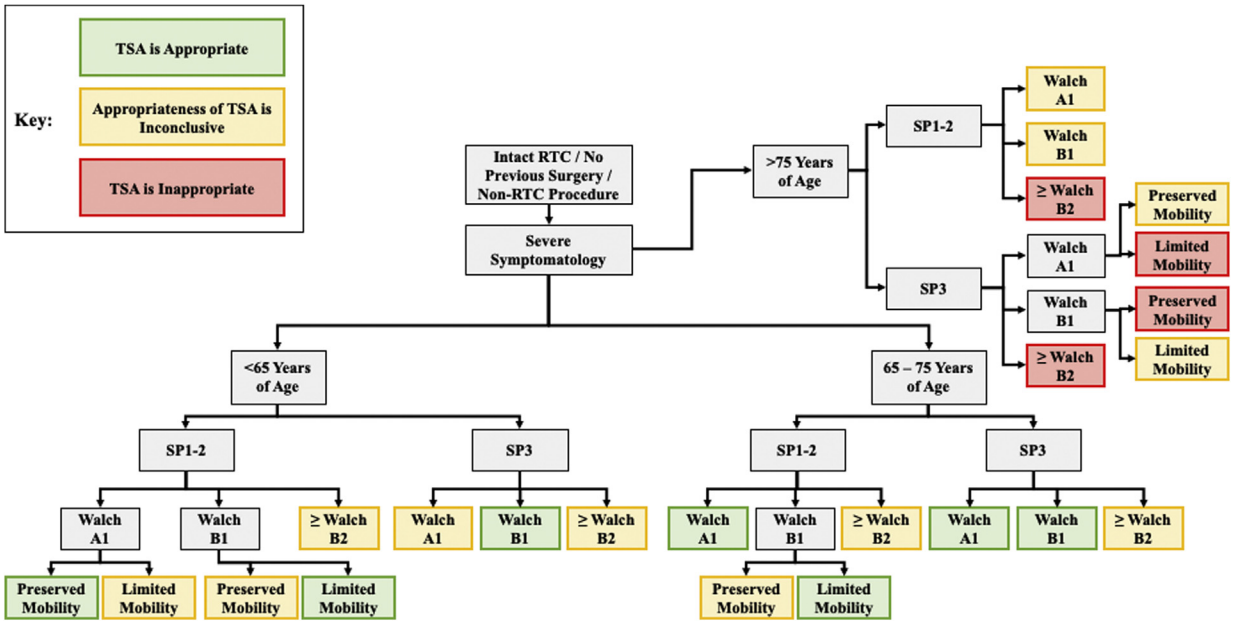

Figure 5 Decision tree for patient with severe symptomatology and an intact rotator cuff. Of note, mobility status was evaluated in all indications with Walch A1 and Walch B1. In cases where mobility status was not explicitly mentioned (eg, patients aged <65 years with severe symptomatology and SP3), panelists determined the appropriateness of TSA to be the same irrespective of mobility status. For full decision tree, please see Supplementary Figures S1 and S2.

inconclusive ratings, our results demonstrate trends that are consistent with current clinical practice based on the expertise of ASES members and available scientific evidence.

As with all studies, there are limitations. Statistical analyses were not performed in the generation of the decision tree, and results were not tested against actual patients. Finally, the present study is subject to the inherent limitations of the RAND method, including the method being anchored in expert opinion.[6,7,9,11,17,25,33] However, despite these limitations, this study represents a novel use of the RAND/UCLA method to assess the appropriateness of TSA. The results of this study ultimately support the use of this decision tree as an aid in informing future clinical practice guidelines. This tool should be used as part of a broader set of factors, including individual patient characteristics, prior studies, and expert opinion, to inform clinical decision making and future practice guidelines. Ongoing research seeks to evaluate these criteria in the setting of clinical practice, including evaluating outcomes of patients electing to undergo TSA relative to their respective indication in the decision tree.[33]

Conclusion

Using the RAND/UCLA method, ASES members developed a decision tree to determine appropriateness of total shoulder arthroplasty across a range of pertinent indications. This presents the data, which are based on a combination of clinical experience from high-volume ASES-member surgeons and a comprehensive review of current evidence, in a manner that streamlines the clinical decision-making process and assists in rapid and reliable determination of appropriateness for practitioners. The decision tree can be used as part of the physician toolbox, alongside individual patient factors, current literature, and expert opinion, to inform clinical decision making, improving quality of care and controlling cost.

Disclaimers:

Funding: No funding was disclosed by the authors.
Conflicts of interest: Glen Ross receives presenter and consultant payments from Arthrex and Stryker. He also receives consultant payments from Tornier. Joseph Gentile receives presenter and consultant payments from Catalyst Orthoscience. Randall J. Otto receives presenter and consultant payments from DJ Orthopaedics. Scott G. Kaar is a paid consultant for Mitek. Andrew Jawa is a board/committee member for the American Academy of Orthopaedic Surgeons (AAOS) and American Shoulder and Elbow Surgeons (ASES). He is on the editorial or governing board for the *Journal of Shoulder and Elbow Surgery*. He receives presenter and consultant payments,

as well as research support, from DJ Orthopaedics. He receives intellectual property royalties and stock/stock options from Ignite Orthopaedics. He receives other financial or material support from DePuy and Boston Outpatient Surgical Suites. William Aibinder is on the editorial or governing board for The European Journal of Orthopaedic Surgery and Traumatology. He receives consulting payments from Exactech, Inc. R. Michael Greiwe is on the editorial or governing board for *Acta Experentia Orthopaedica*. He is a board or committee member for ASES. He receives publishing royalties and financial or material support from Elsevier. He receives consultant payments from CONMED Linvatec and Stryker. He receives stock or stock options from Catalyst and OrthoLive. Joseph DeAngelis is a board or committee member for AAOS. He is on the editorial or governing board for *BMC Musculoskeletal Disorders*, *Orthopaedic Journal of Harvard Medical School*, and *Techniques in Orthopaedics*. He receives consulting payments from Anika Therapeutics, DePuy, Miach Orthopaedics, and Patient Apps Inc. Joseph J. King receives consulting payments from Exactech, Inc. He receives stock or stock options from Pacira Pharmaceuticals. Sarav Shah is a board or committee member for AAOS. He is on the editorial or governing board for arthroscopy. He receives consulting payments from Exactech, Inc. The other authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jse.2021.12.025.

References

1. Aleem AW, Chalmers PN, Bechtold D, Khan AZ, Tashjian RZ, Keener JD. Association between rotator cuff muscle size and glenoid deformity in primary glenohumeral osteoarthritis. J Bone Joint Surg Am 2019;101:1912-20. https://doi.org/10.2106/JBJS.19.00086
2. Bercik MJ, Kruse K, Yalizis M, Gauci MO, Chaoui J, Walch G. A modification to the Walch classification of the glenoid in primary glenohumeral osteoarthritis using three-dimensional imaging. J Shoulder Elbow Surg 2016;25:1601-6. https://doi.org/10.1016/j.jse.2016.03.010
3. Black EM, Roberts SM, Siegel E, Yannopoulos P, Higgins LD, Warner JJ. Reverse shoulder arthroplasty as salvage for failed prior arthroplasty in patients 65 years of age or younger. J Shoulder Elbow Surg 2014;23:1036-42. https://doi.org/10.1016/j.jse.2014.02.019
4. Brewley EE, Christmas KN, Gorman RA, Downes KL, Mighell MA, Frankle MA. Defining the younger patient: age as a predictive factor for outcomes in shoulder arthroplasty. J Shoulder Elbow Surg 2020;29:S1-8. https://doi.org/10.1016/j.jse.2019.09.016
5. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. Int J Technol Assess Health Care 1986;2:53-63.
6. Burnand B, Vader JP, Froehlich F, Dupriez K, Larequi-Lauber T, Pache I, et al. Reliability of panel-based guidelines for colonoscopy: an international comparison. Gastrointest Endosc 1998;47:162-6.
7. Campbell SM, Hann M, Roland MO, Quayle JA, Shekelle PG. The effect of panel membership and feedback on ratings in a two-round Delphi survey: results of a randomized controlled trial. Med Care 1999;37:964-8.
8. Chassin MR, Brook RH, Park RE, Keesey J, Fink A, Kosecoff J, et al. Variations in the use of medical and surgical services by the Medicare population. N Engl J Med 1986;314:285-90.
9. Coulter ID, Marcus M, Freed JR. Consistency across panels of ratings of appropriateness of dental care treatment procedures. Community Dent Health 1998;15:97-104.
10. Donohue KW, Ricchetti ET, Ho JC, Iannotti JP. The association between rotator cuff muscle fatty infiltration and glenoid morphology in glenohumeral osteoarthritis. J Bone Joint Surg Am 2018;100:381-7. https://doi.org/10.2106/JBJS.17.00232
11. Escobar A, Quintana JM, Aróstegui I, Azkárate J, Güenaga JI, Arenaza JC, et al. Development of explicit criteria for total knee replacement. Int J Technol Assess Health Care 2003;19:57-70. https://doi.org/10.1017/s0266462303000060
12. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA appropriateness method user's manual. Santa Monica, CA: RAND Corporation; 2001.
13. Frank RM, Lee S, Sumner S, Griffin J, Leroux T, Verma NN, et al. Shoulder arthroplasty outcomes after prior non-arthroplasty shoulder surgery. JB JS Open Access 2018;3:e0055. https://doi.org/10.2106/JBJS.OA.17.00055
14. Garcia GH, Liu JN, Sinatro A, Wu HH, Dines JS, Warren RF, et al. High satisfaction and return to sports after total shoulder arthroplasty in patients aged 55 years and younger. Am J Sports Med 2017;45:1664-9. https://doi.org/10.1177/0363546517695220
15. Gauci MO, Bonnevialle N, Moineau G, Baba M, Walch G, Boileau P. Anatomical total shoulder arthroplasty in young patients with osteoarthritis: all-polyethylene versus metal-backed glenoid. Bone Joint J 2018;100-B:485-92. https://doi.org/10.1302/0301-620X.100B4.BJJ-2017-0495.R2
16. Gyftopoulos S, Rosenberg ZS, Roberts CC, Bencardino JT, Appel M, Baccei SJ, et al. ACR appropriateness criteria imaging after shoulder arthroplasty. J Am Coll Radiol 2016;13:1324-36. https://doi.org/10.1016/j.jacr.2016.07.028
17. Hicks NR. Some observations on attempts to measure appropriateness of care. BMJ 1994;309:730-3.
18. Joyce CD, Gutman MJ, Hill BW, Singh AM, Sherman M, Abboud JA, et al. Radiographic severity may not be associated with pain and function in glenohumeral arthritis. Clin Orthop Relat Res 2021;480:354-63. https://doi.org/10.1097/CORR.0000000000001950
19. 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;34:677-81. https://doi.org/10.1007/s00264-009-0846-6
20. Larequi-Lauber T, Vader JP, Burnand B, Brook RH, Kosecoff J, Sloutskis D, et al. Appropriateness of indications for surgery of lumbar disc hernia and spinal stenosis. Spine (Phila Pa 1976) 1997;22:203-9.
21. Leschinger T, Raiss P, Loew M, Zeifang F. Predictors of medium-term clinical outcomes after total shoulder arthroplasty. Arch Orthop Trauma Surg 2017;137:187-93. https://doi.org/10.1007/s00402-016-2602-x
22. Levy JC, Everding NG, Gil CC, Stephens S, Giveans MR. Speed of recovery after shoulder arthroplasty: a comparison of reverse and anatomic total shoulder arthroplasty. J Shoulder Elbow Surg 2014;23:1872-81. https://doi.org/10.1016/j.jse.2014.04.014
23. Mahony GT, Werner BC, Chang B, Grawe BM, Taylor SA, Craig EV, et al. Risk factors for failing to achieve improvement after anatomic

total shoulder arthroplasty for glenohumeral osteoarthritis. J Shoulder Elbow Surg 2018;27:968-75. https://doi.org/10.1016/j.jse.2017.12.018
24. Matsen FA, Russ SM, Vu PT, Hsu JE, Lucas RM, Comstock BA. What factors are predictive of patient-reported outcomes? A prospective study of 337 shoulder arthroplasties. Clin Orthop Relat Res 2016;474: 2496-510. https://doi.org/10.1007/s11999-016-4990-1
25. McKee M, Priest P, Ginzler M, Black N. How representative are members of expert panels? Qual Assur Health Care 1991;3:89-94.
26. Neyton L, Kirsch JM, Collotte P, Collin P, Gossing L, Chelli M, et al. Mid- to long-term follow-up of shoulder arthroplasty for primary glenohumeral osteoarthritis in patients aged 60 or under. J Shoulder Elbow Surg 2019;28:1666-73. https://doi.org/10.1016/j.jse.2019.03.006
27. Otte RS, Naylor AJ, Blanchard KN, Cancienne JM, Chan W, Romeo AA, et al. Salvage reverse total shoulder arthroplasty for failed anatomic total shoulder arthroplasty: a cohort analysis. J Shoulder Elbow Surg 2020;29:S134-8. https://doi.org/10.1016/j.jse.2020.04.013
28. Panteli D, Legido-Quigley H, Reichebner C, Ollenschläger G, Schäfer C, Busse R. Clinical practice guidelines as a quality strategy. In: Busse R, Klazinga N, Panteli D, Quentin W, editors. Improving healthcare quality in Europe. Copenhagen, Denmark: World Health Organization; 2019. p. 233-57.
29. Patel DN, Young B, Onyekwelu I, Zuckerman JD, Kwon YW. Reverse total shoulder arthroplasty for failed shoulder arthroplasty. J Shoulder Elbow Surg 2012;21:1478-83. https://doi.org/10.1016/j.jse.2011.11.004
30. Polisetty TS, Colley R, Levy JC. Value analysis of anatomic and reverse shoulder arthroplasty for glenohumeral osteoarthritis with an intact rotator cuff. J Bone Joint Surg Am 2021;103:913-20. https://doi.org/10.2106/JBJS.19.01398
31. Poondla RK, Sheth MM, Heldt BL, Laughlin MS, Morris BJ, Elkousy HA, et al. Anatomic and reverse shoulder arthroplasty in patients 70 years of age and older: a comparison cohort at early to midterm follow-up. J Shoulder Elbow Surg 2021;30:1336-43. https://doi.org/10.1016/j.jse.2020.08.030
32. Porchet F, Vader JP, Larequi-Lauber T, Costanza MC, Burnand B, Dubois RW. The assessment of appropriate indications for laminectomy. J Bone Joint Surg Br 1999;81:234-9.
33. Quintana JM, Aróstegui I, Azkarate J, Goenaga JI, Elexpe X, Letona J, et al. Evaluation of explicit criteria for total hip joint replacement. J Clin Epidemiol 2000;53:1200-8.
34. Raiss P, Schmitt M, Bruckner T, Kasten P, Pape G, Loew M, et al. Results of cemented total shoulder replacement with a minimum follow-up of ten years. J Bone Joint Surg Am 2012;94:e1711-10. https://doi.org/10.2106/JBJS.K.00580
35. Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. Arthritis Rheum 2014; 66:2134-43. https://doi.org/10.1002/art.38685
36. Schoch BS, Wright TW, Zuckerman JD, Bolch C, Flurin PH, Roche C, et al. Glenoid component lucencies are associated with poorer patient-reported outcomes following anatomic shoulder arthroplasty. J Shoulder Elbow Surg 2019;28:1956-63. https://doi.org/10.1016/j.jse.2019.03.011
37. Shah SS, Fu MC, Ling D, Wong A, Warren RF, Dines DM, et al. The comparative effect of age on clinical outcomes following anatomic total shoulder arthroplasty and reverse total shoulder arthroplasty. Orthopedics 2021;44:e600-6. https://doi.org/10.3928/01477447-20210618-24
38. Shekelle PG, Schriger DL. Evaluating the use of the appropriateness method in the Agency for Health Care Policy and Research Clinical Practice Guideline Development process. Health Serv Res 1996;31: 453-68.
39. Simovitch RW, Friedman RJ, Cheung EV, Flurin PH, Wright T, Zuckerman JD, et al. Rate of improvement in clinical outcomes with anatomic and reverse total shoulder arthroplasty. J Bone Joint Surg Am 2017;99:1801-11. https://doi.org/10.2106/JBJS.16.01387
40. Wagner ER, Farley KX, Higgins I, Wilson JM, Daly CA, Gottschalk MB. The incidence of shoulder arthroplasty: rise and future projections compared with hip and knee arthroplasty. J Shoulder Elbow Surg 2020;29:2601-9. https://doi.org/10.1016/j.jse.2020.03.049
41. Walch G, Badet R, Boulahia A, Khoury A. Morphologic study of the glenoid in primary glenohumeral osteoarthritis. J Arthroplasty 1999; 14:756-60.
42. Walker KE, Simcock XC, Jun BJ, Iannotti JP, Ricchetti ET. Progression of glenoid morphology in glenohumeral osteoarthritis. J Bone Joint Surg Am 2018;100:49-56. https://doi.org/10.2106/JBJS.17.00064
43. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999;318:527-30.
44. Wright MA, Keener JD, Chamberlain AM. Comparison of clinical outcomes after anatomic total shoulder arthroplasty and reverse shoulder arthroplasty in patients 70 years and older with glenohumeral osteoarthritis and an intact rotator cuff. J Am Acad Orthop Surg 2020; 28:e222-9. https://doi.org/10.5435/JAAOS-D-19-00166
45. Young AA, Walch G, Pape G, Gohlke F, Favard L. Secondary rotator cuff dysfunction following total shoulder arthroplasty for primary glenohumeral osteoarthritis: results of a multicenter study with more than five years of follow-up. J Bone Joint Surg Am 2012;94:685-93. https://doi.org/10.2106/JBJS.J.00727