Establishing User Error on the Patient-Reported Component of the American Shoulder and Elbow Surgeons Shoulder Score

Jonathan Bourget-Murray,* MD, CM, Ariana Frederick,* MSc, Lisa Murphy,* BSc, Jacqui French,[†] BPhys, Shane Barwood,[†] MBBS, FRACS(Orth), and Justin LeBlanc,*[‡] MD, MSc, FRCSC
Investigation performed at the University of Calgary, Calgary, Alberta, Canada, and the Melbourne Orthopaedic Group, Melbourne, Australia

Background: The American Shoulder and Elbow Surgeons (ASES) score is a patient-reported outcome (PRO) questionnaire developed to facilitate communication among international investigators and to allow comparison of outcomes for patients with shoulder disabilities. Although this PRO measure has been deemed easy to read and understand, patients may make mistakes when completing the questionnaire.

Purpose: To evaluate the frequency of potential mistakes made by patients completing the ASES score.

Study Design: Cross-sectional study; Level of evidence, 3.

Methods: A prospective cross-sectional study was performed for 600 ASES questionnaires completed by patients upon their first visit to 1 of 2 clinic locations (Australian vs Canadian site). Two categories of potential errors were predefined, and then differences in error rates were compared based on demographics (age, sex, and location). To determine whether these methods were reliable, an independent, third reviewer evaluated a subset of questionnaires separately. The interrater reliability was evaluated through use of the Cohen kappa.

Results: The mean patient age was 49.9 years, and 63% of patients were male. The Cohen kappa was high for both evaluation methods used, at 0.831 and 0.918. On average, 17.9% of patients made at least 1 potential mistake, while an additional 10.4% of patients corrected their own mistakes. No differences in total error rate were found based on baseline demographics. Canadians and Australians had similar rates of error.

Conclusion: To ensure the accuracy of the ASES score, this questionnaire should be double checked, as potential mistakes are too frequently made. This attentiveness will ensure that the ASES score remains a valid, reliable, and responsive tool to be used for further shoulder research.

Keywords: ASES; shoulder score; patient-reported outcomes

Patient-reported outcome (PRO) forms are self-administered questionnaires that play an important role in quantifying patients' pain and function. More than 20 different PRO forms are currently used to assess patients with shoulder disabilities.[3,26] Many of these questionnaires have undergone extensive evaluation of their psychometric properties in order to determine their validity, reliability, and responsiveness. However, to the best of our knowledge, there are no published reports of error rates made by patients while completing these questionnaires or standardized methods of how to deal with potential errors.

The American Shoulder and Elbow Surgeons (ASES) Standardized Shoulder Assessment Form was originally designed as a baseline measure consisting of a patient-derived self-evaluation section and a physician evaluation section.[18] The ASES was created as a standardized scale to allow comparisons of outcomes in patients with shoulder problems.[18] It has been demonstrated to be valid,[2,9-13,15,17] reliable,[7,9,12,14,15,20] and responsive[1,10,13,23] for many shoulder conditions, including osteoarthritis, rheumatoid arthritis, arthroplasty, rotator cuff disease, and instability.

The ASES patient-reported component is normally completed without any instructions other than those found on

the questionnaire itself. This section consists of a 100-mm visual analog scale (VAS) used to measure the subjective intensity or frequency of the patient's pain, where the ends of a horizontal line are defined as the extreme limits of pain; from left (*no pain at all*) to right (*pain as bad as it can be*). This is followed by a 10-item Activities of Daily Living (ADL) function section filled separately for each arm. These 10 items represent one's ability to perform specific activities as graded on a 4-point Likert scale, where scores range from 0 on the left (*unable to do*) to 3 on the right (*not difficult*).[18] For research purposes, equal weight is given for the VAS and ADL sections. Calculation of the final ASES score is calculated through an independent formula:

$$[(10 - \text{VAS}) \times 5] + [(5/3) \times \text{ADL}],$$

where the raw scores of the VAS and ADL sections are multiplied by a predefined coefficient to get a final score out of 100.[18] A higher score indicates a favorable outcome.

Although this questionnaire is simple to complete, we have noticed throughout our practice that there are common inconsistencies when patients complete the 10-item ADL section of the ASES. Specifically, we believe that questions 7 (*lift 10 lb above shoulder*) and 8 (*throw a ball overhand*) should, theoretically, receive equal or worse scores than questions 5 (*comb hair*) and 6 (*reach a high shelf*), given the added functional demand required to perform lifting and throwing activities. Although this may appear arbitrary, we believe that if patients can comfortably lift 10 lb above their shoulders or throw a ball overhead (demonstrating some degree of glenohumeral stability and rotator cuff function), then they should be capable of combing their hair or reaching a high shelf. We have also noticed that occasionally, the affected arm received better scores than the unaffected arm, suggesting that patients either were getting the direction of the scoring mixed up or were not paying attention to which arm was which. These mistakes bring into question the accuracy of the ASES score in terms of patients' symptoms and functional status.

Although the ASES is commonly used, no study has established the potential user error on the patient-reported component of the form, nor has user error in completing PRO measures been addressed in the orthopaedic literature. Accuracy in patient responses is important for understanding the extent of patients' symptoms and function and for minimizing bias in the conclusions of studies. Although identifying the accuracy of a patient's perception of his or her symptoms is more complicated to address, identifying nonlogical responses is straightforward, and such user errors should be addressed.[8,25] The purpose of this study was to evaluate the frequency of potential mistakes in the completion of a standard ASES questionnaire.

METHODS

Study Design and Participants

A prospective cross-sectional study was used to establish the potential user error on the patient-reported component of the ASES as administered at 1 of 2 clinic locations. Approval from institutional research ethics boards was obtained from both locations before the start of any data collection. To be eligible to participate in the study, patients needed to have been referred for isolated unilateral shoulder symptoms. Patients were excluded if they had previously been seen at our clinic or if they had a past medical history notable for shoulder abnormality or surgery. Anyone not fluent in English was also excluded from participating.

A total of 400 consecutive patients with unilateral shoulder symptoms who were seen at a private orthopaedic surgery sports medicine clinic in Australia between 2013 and 2014 were initially enrolled to participate in the study. All patients were asked to complete the patient-reported component of the ASES questionnaire upon arrival to their appointment but before seeing the surgeon. No verbal instructions were provided, and patients were asked to refer to the instructions presented on the ASES questionnaire.

The following 2 scoring discrepancies were considered as potential errors:

- Error category 1: Questions 5 and 6 received a worse score than questions 7 and 8 (Figure 1A).
- Error category 2: The patient scored the affected arm better than the unaffected arm (Figure 1B).

In addition, all questionnaires were reviewed to determine how many patients noticed that a category 1 or 2 mistake was made and corrected it before handing in the document. This was identified by answers that were crossed out and reanswered, based on the above comparison of questions 5 and 6 and questions 7 and 8, or when scores were switched between the left and right, as in Figure 1C.

Accuracy and Generalizability

To determine whether these methods were reliable and generalizable to a broad population, an additional 200 patients from a public university hospital practice in Canada were added to the study between 2015 and 2016 and independently reviewed by 2 authors (L.M. and J.L.).
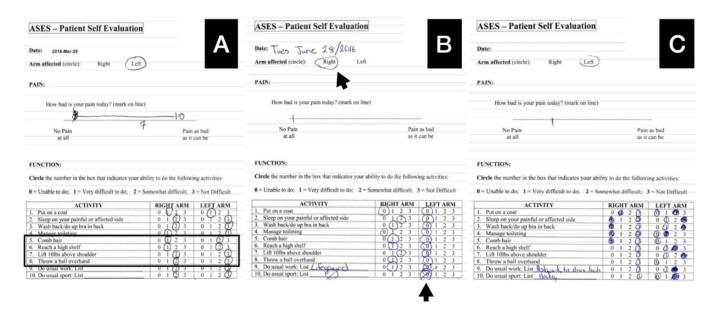
Figure 1. Examples of errors made by patients while completing the American Shoulder and Elbow Surgeons (ASES) score. (A) Indicating inconsistent scores between questions 5 and 6 and questions 7 and 8 (box). (B) Mistakenly scoring the (wrong) opposite/nonaffected arm (arrow). Note: Example A also demonstrates scoring the nonaffected arm as worse than the affected arm. (C) Patients correcting mistakes on their own.

One of the reviewers who assessed the Australian questionnaires also reviewed the Canadian sample for consistency. The same inclusion and exclusion criteria used in Australia were used to select the Canadian patients.

Data on age and sex were collected for each patient. Chisquare and t tests were performed to determine whether there were any differences between patients who completed the questionnaire with potential error and those who made no such errors. Differences were considered significant at $P < .05$.

Interrater reliability was tested through use of the Cohen kappa. To compute the Cohen kappa, each questionnaire was marked as having a potential category 1 error if questions 5 and 6 were scored worse than questions 7 and 8. All variations of this between the 4 questions were pooled together. A potential category 2 error was if the affected arm scored better than the unaffected arm. No distinction was made between whether the affected arm was on the left or right, only whether the scoring pattern matched. If only a portion of the scores followed this pattern, the majority of the scores for the affected arm had to be scored better than the unaffected arm to be considered a potential error. If a difference in agreement occurred between the 2 raters, the questionnaire was reviewed by both raters together and an agreement was made as to whether an actual potential error had occurred.

RESULTS

Demographics

After all 600 ASES questionnaires were reviewed, 5 patients from Canada (CAD) and 8 from Australia (AUS) were removed because they were younger than 18 years. No

TABLE 1
Baseline Demographics

	Canada (n = 195)	Australia (n = 392)	Total (N = 587)	P
Age, y				
Range	18-83	18-87	18-87	
Mean ± SD	48.9 ± 16.25	50.9 ± 16.20	49.9 ± 16.21	.303
Male, %	66	62	63	.356

patients declined participation, and all questionnaires were returned fully complete. The mean age of all included patients at the time they completed the questionnaire was 49.9 years (range, 18-87 years), and 63% of participants were male (CAD: mean age 48.9 years, 66% male; AUS: mean age 50.9 years, 62% male). No differences in age ($t_{585} = -1.03$; $P = .303$) or sex distribution ($\chi^2_{1, n=587} = 0.853$; $P = .356$) were found between the 2 sample populations (Table 1).

Questionnaire Review

The interrater reliability was high for both methods used to evaluate whether questionnaires were completed, with potential error at 0.831 and 0.918 for error categories 1 and 2, respectively (Table 2).

Total potential error rate was similar for both populations (15.9% CAD; 18.9% AUS; $\chi^2_{1, n=587} = 0.787$; $P = .375$). The discrepancy between questions 5 and 6 versus questions 7 and 8 had a similar identification rate between Canadians and Australians; however, Australians had a higher rate of mismatch between the circled "affected arm" and ADL functional scores (ie, the normal arm was

TABLE 2
Interrater Reliability Testing Across Methods Used
to Assess for Incorrect Questionnaires

Error	Cohen Kappa
Category 1: Questions 5 and 6 scored worse than questions 7 and 8	0.831
Category 2: Affected arm scored better than the unaffected arm	0.918
Corrected mistakes	0.914

TABLE 3
Potential Error Rates[a]

	Canada (n = 195)	Australia (n = 392)	Total (N = 587)	P
Any category of error	31 (15.9)	74 (18.9)	105 (17.9)	.375
Category 1	24 (12.3)	60 (15.3)	84 (14.3)	.329
Category 2	13 (6.7)	66 (16.8)	79 (13.5)	.001
Both category 1 and 2	6 (3.1)	52 (13.2)	58 (9.8)	<.001
Self-corrected	27 (13.8)	34 (8.7)	61 (10.4)	.053

[a]Values are expressed as n (%).

TABLE 4
Demographics of Patients Who Submitted Incorrect Scores

	Canada (n = 31)	Australia (n = 74)	Total (n = 105)	P
Age, y				
Range	18-64	18-81	18-81	
Mean ± SD	43.9 ± 15.14	49.3 ± 16.20	47.7 ± 16.62	.128
Male, %	74	59	64	.152

marked with worse scores compared with the truly symptomatic arm) (Table 3). This also indicates that more participants in the Australian cohort made both categories of errors (CAD, 6/195, 3%; AUS, 52/392, 13%; $\chi^2_{1, n=587}$ = 15.18; $P < .001$). No statistical differences were observed in baseline demographics including age, sex, or geographical location between patients who completed their questionnaires correctly or with potential error (Table 4).

In addition to the 2 categories of errors, about 10% (CAD, 13.8%; AUS, 8.7%; $\chi^2_{1, n=587}$ = 3.74; $P = .053$) of questionnaires were initially completed incorrectly but then were revised by the patient before handing in the questionnaire (Table 3).

DISCUSSION

The current study adds to the literature by providing data on the potential user error of the patient self-reporting section of the ASES. We report that our patients were potentially completing their ASES forms with an average error rate of 17.9%, plus an additional 10% of patients who self-corrected any potential errors. This study included 2 English-speaking populations located in 2 different countries: Canada and Australia. This is a surprisingly high potential user error rate, considering that the ASES is among the most commonly implemented PRO measure for patients with shoulder disability[19] and has the best overall psychometric properties.[22] The ASES is commonly used as an initial screening tool in both clinical and research aspects of orthopaedic surgery. Understanding and establishing user error across completed ASES questionnaires are therefore paramount for minimizing bias in the conclusions of studies and in making informed decisions in clinical practice. Unfortunately, the practical implementation and user error of the ASES have never been formally studied. Although the ASES form has been deemed easy to read and understand,[3,16] this study shows that mistakes are a legitimate concern.

The patient self-evaluation section of the ASES is often used in shoulder research as a means to subjectively assess patient symptoms. A reliable PRO measure should be valid (the score measures what it is supposed to measure), reliable (results are precise and reproducible), and responsive (the instrument can measure clinical change).[12] These parameters have previously been evaluated for the ASES, and it has demonstrated content, construct, criterion, and discriminant validity.[2,9-13,15,17] The ASES questionnaire has also been translated and validated for use in numerous languages, including German and Italian.[9,15] It is a reliable outcome measure showing good test-retest scores, with an intraclass correlation coefficient ranging from 0.84 to 0.96,[4,9,11-13,15,21] and good internal consistency, with a Cronbach α of 0.61 to 0.96.[7,9,11-13,15] The ASES has demonstrated good responsiveness, as measured by its statistically minimal detectable change of 9.4 to 15.5 points[12,13,20] within a 90% confidence interval of 6.7 to 11 points[12,13] and a minimal clinically important difference (from a patient perspective) ranging from 6.4 to 21 points.[12,20,23,24] Normative scores for the ASES have been reported to be 92 to 99.[6,21]

Inaccuracies and oversights in completing a PRO measure are usually tested and minimized when the PRO measure's reliability and validity are established. For this reason, once a PRO measure becomes an accepted tool, clinicians and researchers often presume that it is easy to complete and therefore truly quantifies how patients feel or what they are able to do in the context of their health status, thus reflecting the voice of the patient. Michener and Leggin[12] suggested that practical considerations for an outcome score include ease of administration, time to completion, and time to score. The ASES takes, on average, 4 minutes to complete and 2 additional minutes to score.[12] It is easy to administer, as it is meant to be self-explanatory, with all necessary instructions written on the document. As with most validated PRO measures, the ASES is written at a reading level below a high school level—a target deemed acceptable by many governing bodies.[16] However, patient responses can be influenced by individual linguistic, cultural, and educational backgrounds or emotional factors such as stress.[5] These factors can influence a patient to rush or not read directions completely, possibly contributing to blatant errors or oversights in responses. Such inaccuracies can ultimately influence the

relevance of the PRO measure in reflecting how patients feel or their functional limitations in the context of their health.

Of the nearly 600 questionnaires reviewed in this study, almost 28% were completed with some degree of potential inaccuracy (with errors) or uncertainty (self-corrected). Neither age, sex, nor geographic location showed any significant bearing on our results, as baseline patient characteristics were similar between participants who completed their questionnaire correctly or incorrectly. A limitation of the study is that extensive baseline patient characteristics were not recorded. Variables such as ethnicity, level of education, severity of injury or symptoms, and previous experience with questionnaires could all be factors correlated with rates of error. In addition, we are unable to confirm whether any patients sought clarification elsewhere (ie, secretary, family member, mobile device) while completing their forms. Another possible explanation for the mistakes may stem from the difference in numerical scoring between the VAS and the Likert scale used in the ASES shoulder score: for the VAS, zero represents "no pain at all" (the best score), whereas on the Likert scale, zero represents "unable to do" (the worst score). This may lead to a misinterpretation by some patients if instructions are not read carefully.

Ensuring data accuracy is imperative, and as large databases and electronic accessibility of PRO scores become more commonplace, it becomes increasingly more important to find ways to ensure low error rates. Examining logistically impossible (or improbable) responses is valid and easy, and this is a commonly suggested method for checking data quality after chart reviews.[8,25] Addressing error rates in PRO scores has not been explored in the literature, but we demonstrate here that it should be taken into consideration in study design. To the best of our knowledge, no previous studies have discussed using an assessor to provide verbal cues or have addressed the usefulness of reviewing the completed questionnaire with patients to ensure that the score truly corroborates the patients' presentation. Our study is further limited by the lack of data on verification of participants' ASES scores. Verifying scores while patients are in clinic can serve as a spot-check to clarify any answers where a user error may have been made.

We believe that the ASES is a valuable tool for investigators to use. However, given the herein demonstrated potential for mistakes when the form is administered as designed (without verbal explanation), we recommend that all ASES scores be carefully reviewed with patients to ensure more accurate results. Reviewing the scores of the affected versus the unaffected arm and reviewing the responses to questions 5 and 6 versus questions 7 and 8 would be quick and easy to achieve. Additional suggestions moving forward include providing brief verbal instructions when administering the questionnaire, or changing the numerical values of the ADL section so they reflect those of the VAS, whereby worse symptoms are associated with a higher number.

CONCLUSION

The ASES is commonly used in clinical and research aspects of orthopaedic surgery. It has been shown to be a valid, reliable, and responsive questionnaire for patients with shoulder symptoms. However, inaccuracies or oversights made by patients when negotiating the form could have significant implications for the accuracy of research outcomes or in making informed decisions in the clinical practice setting. Establishing the accuracy of patient responses would be of interest to shoulder surgeons or any clinicians who routinely use the ASES in their practice—whether to guide treatment, monitor the progress of care, or improve the quality of health care services. This study suggests that on average, 17.8% of patients may incorrectly complete the ASES form, whereas an additional 10.4% of patients make a mistake while completing the form but recognize it and correct it themselves. These values can be dramatically reduced with the following recommendations.

Recommendations

Some vetting at the point of care is needed after completion of the patient self-report section of the ASES form to ensure that accurate information is gathered. Anyone using the ASES score for clinical or research purposes should consider the answers to questions 5 and 6 in the ADL section and how these compare with the answers to questions 7 and 8. Additionally, ensuring that the scores for the affected arm are lower (ie, worse) than those for the unaffected arm will improve the accuracy of the ASES score to reflect patient symptoms and improve the accuracy of research outcomes. This takes little time (<10 seconds) and can easily be done by the clinician, a nurse, or a research assistant (depending on the resources available for each site) during the patient's visit. This will ensure that the ASES score remains valid, reliable, responsive, and accurate.

Future studies should address participant error rates upon implementation within a clinical setting. Considering the lack of information on this topic, most patient-completed questionnaires should take participant error into consideration, along with appropriate measures for identifying and correcting potential errors (such as quick verification points). Future research could also consider factors that could lead to a higher user error and consider recording extensive demographic factors such as level of education, language ability, and severity of injury.

REFERENCES

1. Angst F, Goldhahn J, Drerup S, Aeschlimann A, Schwyzer HK, Simmen BR. Responsiveness of six outcome assessment instruments in total shoulder arthroplasty. *Arthritis Rheum*. 2008;59(3):391-398.
2. Angst F, Pap G, Mannion AF, et al. Comprehensive assessment of clinical outcome and quality of life after total shoulder arthroplasty: usefulness and validity of subjective outcome measures. *Arthritis Rheum*. 2004;51(5):819-828.
3. Angst F, Schwyzer HK, Aeschlimann A, Simmen BR, Goldhahn J. Measures of adult shoulder function: Disabilities of the Arm, Shoulder, and Hand Questionnaire (DASH) and its short version (QuickDASH), Shoulder Pain and Disability Index (SPADI), American Shoulder and Elbow Surgeons (ASES) Society Standardized Shoulder Assessment Form, Constant (Murley) Score (CS), Simple Shoulder Test (SST), Oxford Shoulder Score (OSS), Shoulder Disability Questionnaire

(SDQ), and Western Ontario Shoulder Instability Index (WOSI). *Arthritis Care Res (Hoboken)*. 2011;63(suppl 11):S174-S188.

4. Beaton D, Richards RR. Assessing the reliability and responsiveness of 5 shoulder questionnaires. *J Shoulder Elbow Surg*. 1998;7(6): 565-572.

5. Chang EM, Gillespie EF, Shaverdian N. Truthfulness in patient-reported outcomes: factors affecting patients' responses and impact on data quality. *Patient Relat Outcome Meas*. 2019;10:171-186.

6. Clarke MG, Dewing CB, Schroder DT, Solomon DJ, Provencher MT. Normal shoulder outcome score values in the young, active adult. *J Shoulder Elbow Surg*. 2009;18(3):424-428.

7. Cook KF, Roddey TS, Olson SL, Gartsman GM, Valenzuela FF, Hanten WP. Reliability by surgical status of self-reported outcomes in patients who have shoulder pathologies. *J Orthop Sports Phys Ther*. 2002;32(7):336-346.

8. Davis M. Data cleaning. In: Salkind NJ, ed. *Encyclopedia of Research Design*. Thousand Oaks, CA: SAGE Publications; 2010. https://dx.doi.org/10.4135/9781412961288.n100/. Accessed August 3, 2019.

9. Goldhahn J, Angst F, Drerup S, Pap G, Simmen BR, Mannion AF. Lessons learned during the cross-cultural adaptation of the American Shoulder and Elbow Surgeons shoulder form into German. *J Shoulder Elbow Surg*. 2008;17(2):248-254.

10. Kemp KA, Sheps DM, Beaupre LA, Styles-Tripp F, Luciak-Corea C, Balyk R. An evaluation of the responsiveness and discriminant validity of shoulder questionnaires among patients receiving surgical correction of shoulder instability. *Scientific World Journal*. 2012;2012: 410125.

11. Kocher MS, Horan MP, Briggs KK, Richardson TR, O'Holleran J, Hawkins RJ. Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease, and glenohumeral arthritis. *J Bone Joint Surg Am*. 2005;87(9):2006-2011.

12. Michener LA, Leggin BG. A review of self-report scales for the assessment of functional limitation and disability of the shoulder. *J Hand Ther*. 2001;14(2):68-76.

13. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons standardized shoulder assessment form, patient self-report section: reliability, validity, and responsiveness. *J Shoulder Elbow Surg*. 2002;11(6):587-594.

14. Oh JH, Jo KH, Kim WS, Gong HS, Han SG, Kim YH. Comparative evaluation of the measurement properties of various shoulder outcome instruments. *Am J Sports Med*. 2009;37(6):1161-1168.

15. Padua R, Padua L, Ceccarelli E, Bondi R, Alviti F, Castagna A. Italian version of ASES questionnaire for shoulder assessment: cross-cultural adaptation and validation. *Musculoskelet Surg*. 2010;94(suppl 1):S85-S90.

16. Perez JL, Mosher ZA, Watson SL, et al. Readability of orthopaedic patient-reported outcome measures: is there a fundamental failure to communicate? *Clin Orthop Relat Res*. 2017;475(8):1936-1947.

17. Razmjou H, Bean A, van Osnabrugge V, MacDermid JC, Holtby R. Cross-sectional and longitudinal construct validity of two rotator cuff disease-specific outcome measures. *BMC Musculoskelet Disord*. 2006;7:26.

18. Richards RR, An KN, Bigliani LU, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg*. 1994;3(6): 347-352.

19. Roe Y, Soberg HL, Bautz-Holter E, Ostensjo S. A systematic review of measures of shoulder pain and functioning using the international classification of functioning, disability and health (ICF). *BMC Musculoskelet Disord*. 2013;14:73.

20. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. *Arthritis Rheum*. 2009;61(5):623-632.

21. Sallay PI, Reed L. The measurement of normative American Shoulder and Elbow Surgeons scores. *J Shoulder Elbow Surg*. 2003;12(6): 622-627.

22. Schmidt S, Ferrer M, Gonzalez M, et al. Evaluation of shoulder-specific patient-reported outcome measures: a systematic and standardized comparison of available evidence. *J Shoulder Elbow Surg*. 2014;23(3):434-444.

23. Tashjian RZ, Deloach J, Green A, Porucznik CA, Powell AP. Minimal clinically important differences in ASES and simple shoulder test scores after nonoperative treatment of rotator cuff disease. *J Bone Joint Surg Am*. 2010;92(2):296-303.

24. Tashjian RZ, Hung M, Keener JD, et al. Determining the minimal clinically important difference for the American Shoulder and Elbow Surgeons score, simple shoulder test, and visual analog scale (VAS) measuring pain after shoulder arthroplasty. *J Shoulder Elbow Surg*. 2017;26(1):144-148.

25. Van den Broeck J, Cunningham SA, Eeckels R, Herbst K. Data cleaning: detecting, diagnosing, and editing data abnormalities. *PLoS Med*. 2005;2(10):e267.

26. Wright RW, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg*. 2010;18(7):436-444.