The American Shoulder and Elbow Score Is Highly Correlated With the Western Ontario Rotator Cuff Index and Has Less Responder and Administrator Burden

Keith M. Baumgarten, M.D., Brett J. Barthman, M.S., and Peter S. Chang, M.D.

**Purpose:** To compare the correlation, responsiveness, and responder and administrator burdens of the American Shoulder and Elbow Score (ASES) with the Western Ontario Rotator Cuff Index (WORC) for patients undergoing arthroscopic rotator cuff repair to determine whether one score is superior to the other to limit the use of multiple scoring measures when tracking patient outcomes. **Methods:** A retrospective review of a database of patients undergoing arthroscopic rotator cuff repair was reviewed where the ASES was simultaneously recorded with the WORC. Correlations were determined using the Pearson coefficient. Subgroup analysis was performed to determine if correlations differed in (1) preoperative outcome and (2) postoperative outcome determinations. Responsiveness was determined by calculating the standardized response mean and the effect size of both scores. Responder and administrator burden was examined using 50 consecutively scored WORC and ASES scores by determining the number of questions accurately answered and the length of time taken to score each questionnaire. **Results:** Correlation was excellent for the ASES and WORC ($r = 0.90$). The correlation of preoperative scores was strong-moderate ($r = 0.69$), and the correlation of postoperative scores was excellent ($r = 0.86$). The standardized response mean WORC = 2.3; ASES = 2.2) and the effect size (WORC = 2.9; ASES = 2.8) demonstrated comparable responsiveness. In total, 71.5% of the WORC questions were able to be scored compared with 93.3% for the ASES ($P < .0001$). The mean time to score the WORC was significantly greater than the ASES (154 vs 23 seconds; $P < .0001$). **Conclusions:** There is excellent correlation and comparable responsiveness between the ASES and WORC. Since there is greater responder and administrative burden for the WORC score, the authors recommend using the ASES over the WORC in patients undergoing rotator cuff repair. **Level of Evidence:** Level IV, diagnostic series.

In an era of value-based medicine, it is essential that both clinicians and researchers track patient outcomes to ensure that we are providing value to our patients. Porter and Lee\(^1\) suggested that the rigorous measurement of value (outcomes and costs) is perhaps the single most important step in improving health care. Patient-determined outcome scores have become the standard for determining outcomes after treatment interventions. In a review of shoulder rating scales published in 2010, at least 26 different shoulder outcome scores were available to choose from to track patient outcomes.\(^2\) At this time, there is no consensus as to which scales should be considered for either clinical or research purposes. It would be ideal to identify one single quality-of-life outcome score that could be universally accepted among shoulder clinicians and researchers to track outcomes in patients with shoulder pathology. However, at this time, that one ideal patient outcomes measure either does not exist or has not been identified. Thus, it is important to identify scores that may be redundant and unnecessarily increase responder burden without adding additional...
psychometric benefit. Scores that are highly correlated may be considered redundant, and an investigator may want to avoid using them simultaneously.

Some have recommended including both a joint-specific outcome score and a disease-specific outcome score when assessing patient outcomes to provide a comprehensive assessment of patient-determined outcomes. The American Shoulder and Elbow Score (ASES) is an example of a joint-specific outcome score, whereas the Western Ontario Rotator Cuff Index (WORC) is an example of a disease-specific score. Both scores are commonly used simultaneously to assess outcomes in patients with rotator cuff tears.

Although it is a more comprehensive assessment, patients are burdened when asked to answer multiple scales simultaneously. In 1 study, patients required an average of 18-20 minutes to complete the multiple scales used. We have found that the WORC score has an average of 18-20 minutes to complete the multiple scales simultaneously. In 1 study, patients required an average of 18-20 minutes to complete the outcome measures. In this study, patients were burdened when asked to answer multiple analog scales required by this outcomes instrument. It is not unusual for patients to incorrectly complete the questionnaire, thus invalidating the results of the patient outcomes tool. It has been our anecdotal experience that approximately 20% of patients incorrectly answer the WORC form regardless of the instructions being highlighted on the survey itself. To obtain valid results in these circumstances when patients have a difficult time accurately completing the outcome score, the physician/researcher is further burdened to instruct the patient on how to correctly complete the form, and the patients are additionally burdened when they have to complete the form twice.

The purpose of this study was to compare the correlation, responsiveness, and responder and administrator burdens of the ASES with the WORC undergoing arthroscopic rotator cuff repair to determine whether 1 score was superior to the other to limit the use of multiple scoring measures when tracking patient outcomes. Eliminating the simultaneous use of redundant scoring systems would further decrease responder and administrator burden and may potentially make shoulder outcomes research more reliable, accurate, and decrease the percentage of subjects lost to follow-up. We hypothesized that the ASES rating would highly correlate with the WORC for patients undergoing arthroscopic rotator cuff repair and that the WORC and ASES scores would both be highly responsive and comparable with each other in regard to sensitivity to change of clinical status.

Methods

After intuitional review board approval (University of South Dakota IRB-19-36), the senior author’s (K.M.B.) registry of patients undergoing arthroscopic rotator cuff repair from January 2008 until December 31, 2020, was deidentified and then retrospectively reviewed. Data from this registry have been included in previously published manuscripts. Inclusion criteria were patients that had an arthroscopic rotator cuff repair and had a preoperative and/or postoperative ASES score recorded simultaneously with a WORC score. All outcome scores were completed in a pen-and-paper format. Patients with procedures performed concomitant to the arthroscopic rotator cuff repair such as subacromial decompressions, acromioclavicular joint resections, biceps tenodesis, and labral tears also were included. Patients who had a procedure that did not include an arthroscopic rotator cuff repair or those that did not have simultaneously recorded ASES and WORC scores were excluded. Indications for arthroscopic rotator cuff repair were patients with a symptomatic, full-thickness rotator cuff tear in which a patient-determined, shared-decision making model was used with the surgeon (K.M.B.) providing education, guidance, and informed consent to the treatment options available to the patient.

Correlations between scores were determined using a Pearson correlation coefficient ($r > 0.7$ excellent, $r = 0.61-0.7$ strong-moderate, $r = 0.31-0.6$ moderate, $r = 0.2-0.3$ poor). Correlation analysis also was stratified for preoperative scores and postoperative scores. Responsiveness was measured by determining (1) the standardized response mean (SRM) and (2) the effect size. The SRM was calculated by dividing the mean change in the outcome scores by the standard deviation of the amount of change of each score. Greater magnitudes in the SRM demonstrate greater sensitivity to change. The effect size was calculated by subtracting the mean preoperative score from the mean postoperative score and dividing it by the standard deviation of the preoperative score. Large effect sizes were considered $>0.80$.

To investigate the aims of examining responder and investigator burden, a predefined, matched comparison group of 50 consecutive simultaneously recorded WORC and ASES scores were examined by the senior author (K.M.B.). The number of incorrectly scored questions was recorded. An incorrectly scored question was defined as a question left blank and unanswered, a question with duplicate answers, or an answer to a visual analog question of which the senior author could not accurately measure. In addition, the time to score each completed outcome measure was recorded with a digital stopwatch. The Student $t$ test was used to examine differences between normally distributed groups and Mann–Whitney $U$ test was performed for non-normally distributed groups, and $\chi^2$ testing to analyze discrete variables. The level of significance was set at .05.
Results

Correlation was excellent for the ASES and WORC for all patients undergoing arthroscopic rotator cuff repair \((n = 1321; r = 0.90; P < .0001)\) (Fig 1). The correlation of preoperative scores was strong-moderate \((n = 823; r = 0.69; P < .0001)\) and the correlation of postoperative scores was excellent \((n = 498; r = 0.86; P < .0001)\). Postoperative score assessment occurred at 6 weeks \((n = 37; r = 0.75; P < .0001)\), 12 weeks \((n = 49; r = 0.78; P < .0001)\), 18 weeks \((n = 38; r = 0.87; P < .0001)\), 6 months \((n = 46; r = 0.79; P < .0001)\), 1 year \((n = 8; r = 0.88; P < .0001)\), and final follow-up \((\text{mean 3.7 years)}\), which ranged from 2 years to 7.5 years \((n = 280; r = 0.81; P < .0001)\). There was not a statistically significant difference between mean ASES and WORC scores respectively at any time point recorded (Table 1).

Both scores were determined to be highly responsive and comparable with the standardized response, mean of the WORC was 2.3 and the ASES was 2.2. The effect size of the WORC was 2.9 and the ASES was 2.8.

The mean number of questions unable to be accurately scored per outcomes measure was 5.9 for the WORC \((21 \text{ total questions})\) and 0.74 for the ASES score \((11 \text{ total questions})\) \((P = .01)\). The percentage of questions able to be scored for the WORC was 71.5% compared with 93.3% for the ASES score \((P < .0001)\). Eleven patients responded incorrectly to all 21 of the WORC questions, whereas the greatest number of questions left unanswered on the ASES was 3 (Table 2). The majority of the questions that were not completed in the ASES were the sports and work questions \((28 \text{ of the 37 unanswered questions})\), which may reflect that this population does not typically participate in sports and are no longer working. The mean duration of time taken for the investigator to score the WORC was significantly greater than the ASES \((154 \text{ vs } 23 \text{ seconds}; P < .0001)\).

Discussion

This study demonstrates that in general the ASES score \((\text{a joint-specific score})\) and the WORC score \((\text{a disease-specific score})\) have excellent correlation in patients undergoing arthroscopic rotator cuff repair. Thus, it appears that their simultaneous use may be redundant, and one may consider eliminating one of these scores when tracking patient outcomes to decrease responder and administrator burden. It appears that it is not essential to simultaneously record both a joint-specific score and a disease-specific score if either the ASES or WORC is used. This study also confirmed that both scores were highly responsive, and it does not appear that one is superior to the other in detecting change in clinical status after arthroscopic

| Table 1. Comparison of Mean WORC and ASES Scores |
|-----------------------------------------------|
| Preoperative \((n = 823)\) | 42 ± 17 | 43 ± 17 | .38 |
| 6 weeks \((n = 37)\) | 57 ± 15 | 54 ± 14 | .17 |
| 12 weeks \((n = 49)\) | 73 ± 18 | 70 ± 16 | .16 |
| 18 weeks \((n = 38)\) | 81 ± 17 | 79 ± 16 | .35 |
| 6 months \((n = 46)\) | 86 ± 16 | 85 ± 16 | .56 |
| 12 months \((n = 48)\) | 90 ± 14 | 89 ± 14 | .24 |
| Final follow-up | 88 ± 16 | 88 ± 17 | .85 |
| Mean 3.7 years \((n = 498)\) | |

NOTE. Values are shown as mean ± standard deviation.

ASES: American Shoulder and Elbow Score; WORC: Western Ontario Rotator Cuff Index.
Table 2. Successful Scoring of the WORC Compared With the ASES

| Successful scoring of all questions | WORC (21 Scored Questions) | ASES (11 Scored Questions) |
|-------------------------------------|-----------------------------|-----------------------------|
| Unable to score 1 question           | 5                           | 10                          |
| Unable to score 2 questions          | 2                           | 6                           |
| Unable to score 3 questions          | 3                           | 5                           |
| Unable to score 4 questions          | 1                           |                              |
| Unable to score 5 questions          | 3                           |                              |
| Unable to score 6 questions          | 1                           |                              |
| Unable to score 19 questions         | 1                            |                              |
| Unable to score 21 questions         | 11                           |                              |
| Total                               | 50                          | 50                          |

NOTE. An incorrectly scored question was defined as a question left unanswered, a question with duplicate answers, or an answer to a visual analog question of which could not be accurately measured.

ASES American Shoulder and Elbow Score; WORC, Western Ontario Rotator Cuff Index.

rotator cuff repair. In addition, this study demonstrated it is more difficult for the patient to successfully complete the WORC compared with the ASES. This is due to some patients having difficulty in comprehending how to answer the visual analog scales that are inherent to the Western Ontario shoulder scoring systems (Fig 2). Twenty-two percent of patients answered all 21 questions on the WORC measure incorrectly, suggesting that rather than some questions not being applicable to their life situation (eg, questions about work in retired people), the patients could not comprehend how to complete the visual analog scale—based questionnaire. This occurred regardless that the instructions were highlighted on the questionnaire. In these situations, our clinician/researcher was required to ask the patient to complete another WORC form after verbal and visual demonstration of how to complete the outcome measure which increased administrator and responder burden. Lastly, this study demonstrates that it takes a significantly longer duration of time for the clinician/researcher to score the WORC score compared with the ASES.

The ASES score was developed in 1994 to be a general shoulder score that could be applied to all patients regardless of the diagnosis.2,5 The self-assessment section has 3 domains that includes a single, numeric-based visual analog scale and a 10 question activities of daily living scale.7 The duration required for patients to complete the ASES ranged from 1.8 to 10 minutes, which was significantly longer than the durations found in our study.4,18-23 The ease of scoring has been considered difficult,24 although Razmjou et al.25 has stated that the superiority of the ASES is its “practicality of being administered in and scored under 5 minutes as compared to 10 to 15 minutes for more lengthy measures” and Roy et al.26 suggested that the administrative burden was low. The ASES has been shown to be a reliable, valid, and responsive outcome tool.18,26 A systematic review determined that the ASES had one of the smallest absolute error of measurements compared to multiple outcome scores including the WORC.27 When examining responsiveness, the minimal clinically important difference of the ASES was 11.1 suggesting that a 11.1% change in the total score will demonstrate an important difference.28 Kocher et al.17 reported that the floor and ceiling effects of the ASES were acceptable. The American Shoulder and Elbow Surgeons Value Committee has recommended it as a outcomes tool that should be used for all patients with shoulder pathology.29

The Western Ontario disease-specific shoulder scores were developed by Kirkley et al.30 using similar methodologies including item reduction and validation using multiple 100-mm visual analog scales anchored by both ends by the extremes of the items being measured. In 2003, the WORC was developed as a tool to measure quality of life in patients with rotator cuff disease and includes 21 questions divided into 5 domains.4,30 The WORC has been shown to be reliable and valid.30 When examining responsiveness, the minimal clinically important difference of the WORC was 245.26, suggesting that an 11.7% change in the total score will demonstrate an important difference.28,30 The ease of scoring has been considered only moderate, with 10 to 15 minutes required to complete the questionnaire.24,27

Since there are excellent correlations between the scores, it appears that simultaneous use of these scores is unnecessary. Thus, the authors recommend using the ASES score instead of the WORC score for several reasons. First, it is our experience that patients have more difficulty correctly completing the WORC score compared with the ASES score, which (1) increases responder and physician burden and (2) decreases the validity of measuring patient outcomes (Fig 2). A systematic review authored by St-Pierre agreed with our assessment that the WORC had a greater administration burden than the ASES.27 Second, it takes longer for the clinician/researcher to score the WORC compared with the ASES. Third, it appears that historically the ASES is more commonly used score than the WORC. When searching a publicly available online database and searching for the combined terms “rotator cuff” and “WORC score,” 77 references were found (https://www.ncbi.nlm.nih.gov/pubmed/?term=rotator+cuff+and+WORC+score). When searching for the combined terms “rotator cuff” and “ASES score,” 365 references were found (https://www.ncbi.nlm.nih.gov/pubmed/?term=rotator+cuff+and+ASES+score). Razmjou et al.25 has suggested that the ASES is the most widely used subjective joint-specific score. Fourth, it does not appear that one score is clinically more responsive than
the other. In addition to this study demonstrating comparability of the responsiveness between the WORC and the ASES, when we examined responsiveness in relation to the minimal clinically important difference, both scores require approximately 11% change to determine the minimal clinically important difference. Lastly, the American Shoulder and Elbow Surgeons Value Committee endorses the use of the ASES as an outcomes tool that should be used for all patients with shoulder pathology, with only providing a recommendation for use of the WORC for research purposes. Similar to our recommendation, Vidt et al. recommended the use of the ASES over the WORC, since it correlated with functional strength outcomes better than the WORC.

Several previous studies demonstrated strong to excellent correlations between the WORC and ASES. In contrast, another study that included 80 subjects in the military with rotator cuff tears did not find a strong correlation between the WORC and ASES. However, this study differed from our study, since it only included patients at initial presentation and did not record scores after surgical intervention.
The strengths of our study include the large sample size, which significantly exceeds similar previous studies,\(^4,15,16,25,31\) the inclusion of both preoperative and postoperative scores, which is rare in previous studies,\(^16,25\) and examination of a homogenous population of only patients who had a rotator cuff tear that required arthroscopic rotator cuff repair, whereas most other published studies included patients with a varying degrees of rotator cuff pathologies.\(^4,15,16,25,31\)

**Limitations**

The limitations of this study include its retrospective design. In addition, other than correlation, responsiveness, and responder and administrator burdens, this study was not structured to determine other psychometric properties such as reliability, internal consistency, or minimal clinically important differences. In addition, this study does not assess a computer-administered format for the WORC and ASES, which may affect both responder and administrator burdens.

**Conclusions**

There is excellent correlation and comparable responsiveness between the ASES and WORC. Since there is greater responder and administrative burden for the WORC score, the authors recommend using the ASES over the WORC in patients undergoing rotator cuff repair.

**References**

1. Porter ME, Lee TH. The strategy that will fix health care. *Harvard Business Rev* 2013;91:50-70.
2. Wright RW, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg* 2010;18:436-444.
3. Richards RR, An KN, LU Biglani, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 1994;3:347-352.
4. Kirkley A, Alvarez C, Griffin S. The development and evaluation of a disease-specific quality-of-life questionnaire for disorders of the rotator cuff: The Western Ontario Rotator Cuff Index. *Clin J Sport Med* 2003;13:84-92.
5. Baumgarten KM, Chang PS, Dannenbring TM, Foley EK. Does arthroscopic rotator cuff repair improve patients’ activity levels? *J Shoulder Elbow Surg* 2018;27:2167-2174.
6. Baumgarten KM, Osborn R, Schweinle WE Jr, Zens MJ, Helsper EA. Are pulley exercises initiated 6 weeks after rotator cuff repair a safe and effective rehabilitative treatment? A randomized controlled trial. *Am J Sports Med* 2016;44:1844-1851.
7. Gagnier JJ, Robbins C, Bedi A, Carpenter JE, Miller BS. Establishing minimally important differences for the American Shoulder and Elbow Surgeons score and the Western Ontario Rotator Cuff Index in patients with full-thickness rotator cuff tears. *J Shoulder Elbow Surg* 2018;27:e160-e166.
8. Watson ST, Robbins CB, Bedi A, Carpenter JE, Gagnier JJ, Miller BS. Comparison of outcomes 1 year after rotator cuff repair with and without concomitant biceps surgery. *Arthroscopy* 2017;33:1928-1936.
9. Saraswat MK, Styles-Tripp F, Beaupre LA, et al. Functional outcomes and health-related quality of life after surgical repair of full-thickness rotator cuff tears using a mini-open technique: A concise 10-year follow-up of a previous report. *Am J Sports Med* 2015;43:2794-2799.
10. Razmjou H, Davis AM, Jaglal SB, Holby R, Richards RR. Disability and satisfaction after rotator cuff decompression or repair: A sex and gender analysis. *BMC Musculoskelet Disord* 2011;12:66.
11. Ravindra A, Barlow JD, Jones GL, Bishop JY. A prospective evaluation of predictors of pain after arthroscopic rotator cuff repair: Psychosocial factors have a stronger association than structural factors. *J Shoulder Elbow Surg* 2018;27:1824-1829.
12. MacDonald P, McRae S, Leiter J, Mascarenhas R, Lapner P. Arthroscopic rotator cuff repair with and without acromioplasty in the treatment of full-thickness rotator cuff tears: A multicenter, randomized controlled trial. *J Bone Joint Surg Am* 2011;93:1953-1960.
13. Burks RT, Crim J, Brown N, Fink B, Greis PE. A prospective randomized clinical trial comparing arthroscopic single- and double-row rotator cuff repair: Magnetic resonance imaging and early clinical evaluation. *Am J Sports Med* 2009;37:674-682.
14. Lapner PL, Sabri E, Rakhra K, et al. A multicenter randomized controlled trial comparing single-row with double-row fixation in arthroscopic rotator cuff repair. *J Bone Joint Surg Am* 2012;94:1249-1257.
15. Provencer MT, Frank RM, Macian D, et al. An analysis of shoulder outcomes scores in 275 consecutive patients: Disease-specific correlation across multiple shoulder conditions. *Mil Med* 2012;177:975-982.
16. Holby R, Razmjou H. Measurement properties of the Western Ontario rotator cuff outcome measure: A preliminary report. *J Shoulder Elbow Surg* 2005;14:506-510.
17. 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;87:2006-2011.
18. 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;11:587-594.
19. Skutek M, Freamery RW, Zeichen J, Bosch U. Outcome analysis following open rotator cuff repair. Early effectiveness validated using four different shoulder assessment scales. *Arch Orthop Trauma Surg* 2000;120:432-436.
20. 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;14:68-76.
21. Cunningham G, Ladermann A, Denard PJ, Kherad O, Burkhart SS. Correlation between American Shoulder and Elbow Surgeons and Single Assessment Numerical Evaluation Score after rotator cuff or SLAP repair. *Arthroscopy* 2015;31:1688-1692.
22. Robins RJ, Anderson MB, Zhang Y, Presson AP, Burks RT, Greis PE. Convergent validity of the patient-reported outcomes measurement information system’s physical function computerized adaptive test for the knee and shoulder injury sports medicine patient population. *Arthroscopy* 2017;33:608-616.

23. Nicholson AD, Kassam HF, Pan SD, Berman JE, Blaine TA, Kovacevic D. Performance of PROMIS Global-10 compared with legacy instruments for rotator cuff disease. *Am J Sports Med* 2019;47:181-188.

24. Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: A systematic review of the literature. *Ann Rheum Dis* 2004;63:335-341.

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

26. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: A systematic review of four questionnaires. *Arthritis Rheum* 2009;61:623-632.

27. St-Pierre C, Desmeules F, Dionne CE, Fremont P, MacDermid JC, Roy JS. Psychometric properties of self-reported questionnaires for the evaluation of symptoms and functional limitations in individuals with rotator cuff disorders: A systematic review. *Disabil Rehabil* 2016;38:103-122.

28. Cvetanovich GL, Gowd AK, Liu JN, et al. Establishing clinically significant outcome after arthroscopic rotator cuff repair. *J Shoulder Elbow Surg* 2019;28:939-948.

29. Hawkins RJ, Thigpen CA. Selection, implementation, and interpretation of patient-centered shoulder and elbow outcomes. *J Shoulder Elbow Surg* 2018;27:357-362.

30. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003;19:1109-1120.

31. Vidt ME, Santiago AC 2nd, Hegedus EJ, et al. Can self-report instruments of shoulder function capture functional differences in older adults with and without a rotator cuff tear? *J Electromyogr Kinesiol* 2016;29:90-99.