Validity and Responsiveness of the Single Alpha-numeric Evaluation for Shoulder Patients

Charles A. Thigpen,*†‡ PT, PhD, ATC, Ellen Shanley,†‡ PT, PhD, OCS, Amit M. Momaya,§ MD, Michael J. Kissenberth,‖ MD, Stefan J. Tolan,‖ MD, John M. Tokish,¶ MD, and Richard J. Hawkins,‡ MD, FRCSC

Investigation performed at Greenville Health System, Steadman Hawkins Clinic of the Carolinas, Greenville, South Carolina, USA

Background: There is an ever-increasing demand for widespread implementation of patient-reported outcomes. However, adoption is slow owing to limitations in clinical infrastructure and resources within busy orthopaedic practices. Prior studies showed the single alpha-numeric evaluation (SANE) score to correlate at a single point in time with the American Shoulder and Elbow Surgeons (ASES) score. However, no study has validated the SANE in terms of test-retest reliability, responsiveness, or clinical utility.

Purpose: To validate SANE with the ASES across a sample of patients with common orthopaedic shoulder diagnoses.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: Patients undergoing rotator cuff repair (n = 77), total shoulder replacement (n = 55), or physical therapy (n = 80) for signs and symptoms of subacromial impingement syndrome (n = 61) or adhesive capsulitis (n = 19) were administered the SANE and ASES at baseline and again at their 3-month follow-up from initial care or surgery (N = 212, mean ± SD age = 52.6 ± 1.2 years, n = 145 women). Interclass correlation coefficient ($ICC_{2,1}$) and standard error of the measurement (SEm) were used to evaluate the test-retest reliability of the SANE and the validity between the SANE and ASES scores. Analysis of variance (treatment group × time) was used to evaluate the responsiveness to treatment, and a receiver operating characteristic curve was used to establish the minimal clinically important difference (MCID) for the SANE as compared with the ASES ($\alpha = .05$). Floor and ceiling effects were evaluated as the percentage of patients who scored the highest or lowest score on each tool.

Results: The SANE demonstrated good pretreatment reliability ($ICC_{2,1} = 0.84$, SEM = 3.8), similar to the ASES ($ICC_{2,1} = 0.82$, SEM = 3.4). The SANE also showed good agreement with the ASES before and after treatment across all treatment groups (rotator cuff repair, $ICC_{2,1} = 0.85$, SEM = 3.4; total shoulder replacement, $ICC_{2,1} = 0.72$, SEM = 5.2; physical therapy: $ICC_{2,1} = 0.82$, SEM = 2.9). The SANE and ASES displayed similar responsiveness after treatment, with similar mean change and SD within each treatment group. The receiver operating characteristic curve revealed an area under the curve of 0.79 (SE, 0.62; $P < .001$) and a cutoff of 15% on the SANE, with a sensitivity of 85% to establish the MCID. Acceptable and similar floor and ceiling effects were observed for the ASES (4%) and SANE (9%).

Conclusion: The study demonstrates that the SANE is valid for a range of common shoulder diagnoses to assess patient outcomes across operative and nonoperative treatment for shoulder complaints. The MCID of 15% is similar to that of the ASES (11%), suggesting that the SANE is a simple and efficient tool to assess treatment effects for shoulder disorders. Future studies are warranted to confirm these results and compare across other body parts and diagnoses.

Keywords: patient-reported outcome measures; shoulder; performance measures; functional scores

Patient-oriented outcomes (PROs) have traditionally been used as research tools to evaluate treatment effectiveness. As patient-centered models of care have emerged, PROs can also be used to guide treatment decisions, determine prognosis, and evaluate treatment response.[20,23,24] However, emergence of value-based payment models, such as those put forth by the Centers for Medicare and Medicaid Services, may use PROs as performance measures and indicators of quality of health care delivery.[17] However, most PROs were developed for research and for specific patient populations and were validated for the purpose of assessing treatment effectiveness.[3]

There are 11 well-studied region-specific shoulder-related outcome scores that have reliability and validity measures reported.[14] However, like most PROs, the

The American Journal of Sports Medicine
1–6
DOI: 10.1177/0363546518807924

surveys require patient understanding and completion, followed by clinician scoring and interpretation, thus resulting in efficiency challenges for a busy orthopaedic clinic.[11] While these measures are valid and have many strengths, implementation at a population level has several barriers, including the lack of appropriate infrastructure or resources, loss of clinician and patient time, and complexity/interpretation of multiquestion forms.[2] The ideal outcome score should be simple and valid and should entail limited resources and time. The single alphanumeric evaluation (SANE) score is an alternative method for assessing patient self-reported function by asking, "How would you rate your shoulder today as a percentage of normal (0% to 100% scale with 100% being normal)?"[25]

The SANE score was validated in comparison with the International Knee Documentation Committee (IKDC) score for knee disorders[19,26] and was recently shown to be associated with postoperative American Shoulder and Elbow Surgeons (ASES) score in the setting of rotator cuff repair, SLAP repair (superior labral anterior and posterior), and total shoulder arthroplasty.[5,18,25] Although the SANE score was shown to correlate with the ASES score at a single point in time, this only establishes construct validity and does not address all necessary components of a measurement tool.

In addition to validity, other psychometric properties of a tool should be established, including the test-retest reliability, responsiveness, error estimates for clinical utility, and floor and ceiling effects across the population of patients likely to be sampled.[6] Test-retest reliability establishes the consistency of the measure and the likelihood that it will yield a similar result when a patient's condition is stable. Error estimates, including the minimal detectable change (MDC) and the minimal clinically important difference (MCID), should be determined to aid in interpretation of statistically and clinically meaningful changes. Finally, a tool should display adequate responsiveness to the population of interest to detect if clinical change has occurred. To our knowledge, no study has evaluated the SANE in terms of test-retest reliability, responsiveness, and clinical utility error estimates, including floor and ceiling effects, across the most common orthopaedic shoulder cases. Therefore, the purpose of this study was to validate the SANE as compared with the ASES across a sample of patients with the most common orthopaedic shoulder diagnoses by examining test-retest reliability, responsiveness, and clinical utility in terms of MCID as well as floor and ceiling effects.

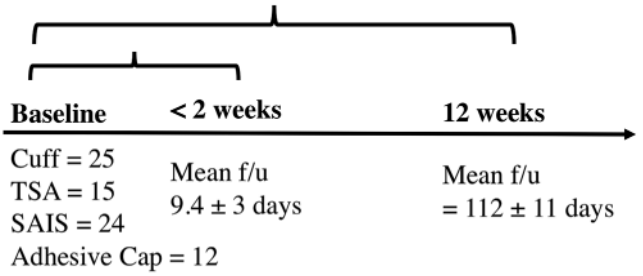

Figure 1. All patients completed the SANE and ASES at baseline and 12 weeks. Patients completed the test-retest reliability within 2 weeks of their surgical date or initial physician visit (for nonoperative cases). ASES, American Shoulder and Elbow Surgeons; cap, capsulitis; f/u, follow-up; SAIS, subacromial impingement syndrome; SANE, single alphanumeric evaluation; TSA, total shoulder arthroplasty.

METHODS

A total of 212 consecutive patients undergoing shoulder treatment were recruited (mean ± SD age, 52.6 ± 1.2 years; n = 145 women): primary arthroscopic rotator cuff repair (n = 77), total shoulder replacement (n = 55) for glenohumeral osteoarthritis, or physical therapy (n = 80) for signs and symptoms of subacromial impingement syndrome (n = 61) or adhesive capsulitis (n = 19). All patients were administered the SANE and ASES at baseline and at their 3-month follow-up from initial care or surgery (Figure 1). Patients receiving physical therapy were diagnosed per McClure and Michener[12] and followed a standardized treatment approach.[21] The self-report portion of the ASES score was used and is a condition-specific shoulder scale intended to measure the functional limitations and pain of the shoulder.[16] This portion of the ASES takes approximately 5 to 7 minutes to complete 2 dimensions: a pain scale and 10 questions quantifying activities of daily living (eg, carrying or lifting above shoulder level). The composite ASES score was used in this study: 50% pain

*Address correspondence to Charles A. Thigpen, PT, PhD, ATC, ATI Physical Therapy, 200 Patewood Dr, Suite 150C, Greenville, SC 29615, USA (email: charles.thigpen@atipt.com).
†ATI Physical Therapy, Greenville, South Carolina, USA.
‡South Carolina Center for Effectiveness Research in Orthopedics, Greenville, South Carolina, USA.
§University of Alabama Birmingham, Birmingham, Alabama, USA.
‖Steadman Hawkins Clinic of the Carolinas, Greenville, South Carolina, USA.
¶Mayo Clinic, Phoenix, Arizona, USA.
Presented at the annual meeting of the AOSSM, Toronto, Ontario, Canada, July 2017.
One or more of the authors has declared the following potential conflict of interest or source of funding: A.M.M. has received education and travel expenses from Smith & Nephew and a grant from DJO. M.J.K. is a paid consultant for Arthrex; has received financial or material support from Arthrex, ArthroCare, Arthrosurface, Breg, DJ Orthopaedics, Neurotech, Pacira, and Smith & Nephew; and has received payments for travel and lodging from Peerless Surgical Inc. S.J.T. is a paid consultant for Stryker; has received payments for speaking, education, and travel from Arthrex; has received travel expenses from Peerless Surgical and Pacira Pharmaceuticals; and has received honoraria and travel expenses from Arthrosurface. J.M.T. is a consultant for Arthrex, DePuy, Mitek, and Zimmer Biomet; is a paid speaker for Arthrex and Mitek; has received education and travel payments from Peerless Surgical; and has received intellectual property royalties from Arthrex. R.J.H. receives intellectual property royalties from Ossur; is a paid consultant for Arthrex; has received grants and royalties from Encore Medical; has received travel payments from Peerless Surgical and DePuy Synthes; and has received consulting fees from Pacira Pharmaceuticals. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

TABLE 1
Descriptive Statistics and Psychometric Properties of the SANE Score for Operative and Nonoperative Patients With SAIS or Adhesive Capsulitis[a]

Time: Treatment	Mean	SD	$ICC_{2,1}$	MDC, %	MCID, %
Baseline					
Rotator cuff	46.6	12.4	0.84	6.7	11.8
TSA	39.3	17.5	0.80	8.1	16.8
SAIS	64.5	16.5	0.88	7.7	16.0
Adhesive capsulitis	58.5	13.5	0.90	7.1	14.7
Total	52.3	15.5	0.84	7.5	14.9
Follow-up, 12 wk					
Rotator cuff	65.1	13.7	0.83	7.2	13.1
TSA	66.4	18.7	0.81	8.6	18.1
SAIS	84.8	14.6	0.86	7.6	14.0
Adhesive capsulitis	76.8	18.2	0.88	8.4	17.5
Total	70.8	15.7	0.83	7.8	15.0

[a]ICC, interclass correlation coefficient; MCID, minimal clinically important difference; MDC, minimal detectable change; SAIS, subacromial impingement syndrome; SANE, single alpha-numeric evaluation; TSA, total shoulder arthroplasty.

and 50% daily function.[10,15] The SANE has been widely reported and simply asks, "How would you rate your shoulder today as a percentage of normal (0% to 100% scale with 100% being normal)?"[25]

Reliability and Validity

The SANE score's test-retest reliability—that is, the stability of the measure during a time when it should not have changed[6]—was assessed with the interclass correlation coefficient ($ICC_{2,1}$) and standard error of the measure (SEM) between 2 baseline measures of the SANE and ASES. A subsample of patients (n = 78) completed the SANE and ASES at baseline and within 14 days (9.4 ± 3 days) to assess test-retest reliability (Figure 1). Construct validity of the SANE was assessed by comparison with the relative gold standard of the ASES, which has well-established psychometrics with $ICC_{2,k}$. The ASES score is also the most commonly used shoulder outcome measure in North America and was suggested as the regional outcome measure of choice by the American Academy of Orthopaedic Surgeons and the American Shoulder and Elbow Surgeons.[1,7] We also assessed the criterion validity by evaluating the precision of the SANE by calculating the SEM and MDC, which is the smallest amount of change statistically likely to be real change in the score.[6,8]

Clinical Utility

Responsiveness and MCID were calculated for the SANE to clarify its clinical application to evaluate patients' reported function over time.[6,8] Responsiveness is the ability of an instrument to accurately detect change as compared with an external criterion indicative of clinically meaningful change, which is most often quantified as the MCID and can be calculated via anchor- or distribution-based methodologies.[9] We chose a distribution-based approach to define the MCID for the ASES and SANE across all patients and then for each treatment group. Additionally, a mixed-model analysis of variance (treatment group × time) was used to evaluate the responsiveness to treatment, and a receiver operating characteristic curve was used to establish the MCID for the SANE comparison with the ASES ($\alpha = .05$). Effect sizes were calculated across all patients and then individually for rotator cuff repair, total shoulder replacement, and physical therapy with nonspecific shoulder pain. To provide a measure of responsiveness, effect sizes were interpreted as follows: 0.2, a small effect; 0.5, a moderate effect; and ≥0.8, a large effect.[4]

Floor or ceiling effects were evaluated and considered present if >15% of patients achieved the lowest or highest possible score at baseline or follow-up.[13,22] The presence of floor or ceiling effects suggests that items at the extreme have limited content validity. This would result in the inability to differentiate 2 scores from each other at the lowest or highest ranges; thus, reliability is reduced. This would also limit the responsiveness because changes cannot be measured for these patients. A sample size of at least 50 patients was suggested to adequately evaluate floor and ceiling effects.[22]

RESULTS

Reliability and Validity

The SANE demonstrated good pretreatment reliability ($ICC_{2,1}$ = 0.84, SEM = 3.4) similar to the ASES ($ICC_{2,1}$ = 0.82, SEM = 3.4) (Table 1). The ASES and SANE scores between baseline and the test-retest measure did not change, with score differences ranging from 0% to 8.8% across all patients. The SANE demonstrated good agreement with the ASES across all treatment groups before and after treatment (Table 1). The MDC ranged from 7% to 9%, depending on the timing and treatment group. This means that differences <9% are likely variability in

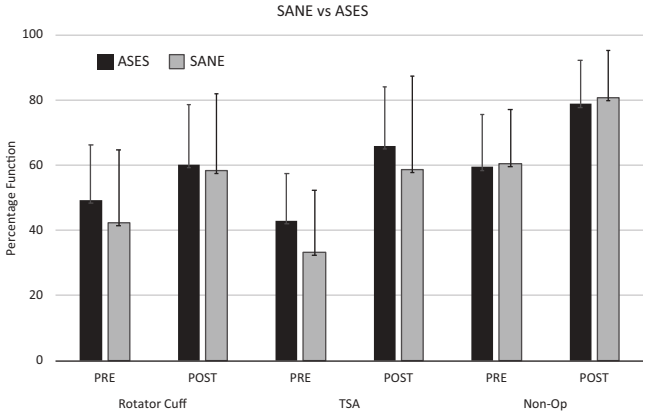

Figure 2. The SANE demonstrated similar responsiveness in the assessment of operative and nonoperative shoulder conditions. Values are presented as mean ± SD. ASES, American Shoulder and Elbow Surgeons; POST, postoperative; PRE, preoperative; SANE, single alpha-numeric evaluation; TSA, total shoulder arthroplasty.

reporting and not reflective of real change. The MCID ranged from 11.8% to 18%, depending on the diagnosis and time of assessment.

Clinical Utility

There was a significant group × time interaction effect on SANE score change among treatment groups ($P < .05$). According to post hoc testing, the nonoperative and total shoulder replacement groups showed different responsiveness over the course of care measured, whereas they were similar to rotator cuff repair. There were large effect sizes across groups, ranging from 1.3 to 1.5, meaning that the SANE change relative to the variation in SANE change was important beyond statistical significance alone. The SANE and ASES displayed similar responsiveness after treatment, with similar mean change and SD within each treatment group (Figure 2).

The receiver operating characteristic curve for the SANE on ASES response group for 12-week measures revealed an area under the curve of 0.79 (SE, 0.62; $P < .001$) and a cutoff of 15%, with a sensitivity of 85% and a specificity of 31% to establish the MCID for postoperative measures. The MCIDs varied by 5% to ≥10% across treatment groups, suggesting that population-specific meaningful changes in function, as measured by the SANE and ASES, appear to exist (Table 1).

Baseline and 3-month follow-up scores were evaluated for floor and ceiling effects by observing the frequency of scores that reached the lowest and highest scores at each time point. When all scores were evaluated, <3% of ASES scores and <8% of SANE scores represented the lowest and highest possible scores. The distribution was not different between baseline and follow-up scores between instruments, with no patients reporting 0 on either scale (lowest scores: 8 and 9, respectively). The maximum SANE and ASES scores were distributed evenly across the postoperative and nonoperative patient subgroups. This suggests that the SANE has lower ceiling effects as compared with the ASES but is well below the accepted cutoff of 15%.[13,22]

DISCUSSION

Our results show that the SANE is reliable, has acceptable precision, and can monitor patients' functional changes after rotator cuff repair, total shoulder arthroplasty, and nonoperative care for common shoulder conditions. The SANE showed pretreatment reliability similar to that of the ASES score; it also displayed similar responsiveness after treatment. Moreover, the MCID was 15%, similar to prior reports for the ASES.[10,15] Finally, the SANE demonstrated acceptable floor and ceiling effects, with rates <10% across all observations, suggesting that it is not limited by the lowest or highest scores for most patients.

Our study demonstrated that the SANE score correlates well with the ASES score, which is in agreement with previous studies. Williams et al[25] evaluated the SANE score and its correlation with the Rowe score and the ASES score in the setting of various shoulder surgical procedures. The authors reported good correlation between the SANE and other scores. Cunningham et al[5] compared the SANE score with the ASES score after rotator cuff repair, rotator cuff revision, and SLAP repair, demonstrating good correlation with postoperative outcome scores. Beyond shoulder pathology, the SANE score correlates well with the Lysholm score in the setting of anterior cruciate ligament reconstruction,[26] the modified Cincinnati Knee Rating System and IKDC in the setting of anterior cruciate ligament reconstruction or knee arthroscopy,[19] and the IKDC in the setting of a knee injury.[27]

Furthermore, the SANE has acceptable floor and ceiling effects, with <15% of the observations reaching the lowest or highest possible scores. This indicates that the SANE is an acceptable outcome measure for short-term follow-up for the majority of patients undergoing shoulder treatment. Our results agree and are in contrast to recent work evaluating the responsiveness and internal validity of the SANE score as compared with the Constant score, Western Ontario Osteoarthritis of the Shoulder (WOOS) index, and ASES score in the setting of total shoulder arthroplasty.[18] In contrast, Sciascia et al[18] demonstrated that the SANE score was the least responsive with the lowest effect sizes, lower than the ASES score and WOOS index. However, differences among the WOOS, ASES, and SANE were not beyond the MDC of these measures, which suggests that these measures were statistically equivalent. In addition, marked postoperative ceiling effects were seen with the SANE score and the ASES score. Interestingly, the authors also noted that many patients may have inadvertently inversed the SANE score, as compared with the standard measures. Our patients wrote a number down as opposed to marking a visual analog scale, so we did not see this in our data. The differences

in observed floor and ceiling effects may be due to the differences in follow-up (2 years vs 3 months) and may reflect higher levels of function in their study. This may also explain why we did not observe consistent floor and ceiling effects in our sample. We did not have a significant portion of high-functioning patients across this sample, and future studies are warranted to understand the SANE's application in higher-functioning populations. Beyond the short follow-up of 3 months, our sample included no patients <25 years old and no competitive athletes, thereby limiting the population to which our results can be generalized. This is an important feature that should be evaluated in future studies.

Our results indicate that the MCID for the SANE score is, on average, 15% across a variety of shoulder conditions. This MCID is similar to that shown for the ASES score (11%).[15] Only 1 previous study evaluated an MCID for the SANE after knee surgery, and it was 7% at 6-month follow-up and 19% at 12-month follow-up, similar to our results.[27] Interestingly, these results suggest that the MCID varies over time as recovery of a population begins to diverge across conditions (12% for rotator cuff vs 18% for total shoulder replacement). This is similar to our results, where the precision and responsiveness across condition varied at baseline and 3 months. Future studies should evaluate the stability of the SANE over time and emphasize the importance of population-specific MCIDs or other measures of recovery for interpretations of clinical outcomes. While the SANE provides similar precision and "percentage function" as compared with the ASES, classifying a patient as improved or failed when in reality one may report no change should be carefully considered. Patient function is a multidimensional construct and is not likely captured by a single outcome measure.[28] Therefore, future studies should evaluate how the SANE compares with other measures, such as the ASES, in assessing patient recovery and response to treatments. However, even in light of these limitations, our results suggest that, similar to the ASES, the SANE score is a simple and time-efficient PRO that can be used to track patients over time within a busy orthopaedic practice.

Our results should be interpreted within the context of our sample, which included a range of patients with varying degrees of shoulder pathology and treatment procedures, as well as a limited time frame. Therefore, the specific MCIDs are population and time specific and should be interpreted as such. Our sample is representative of the average patient presenting for shoulder care but does not include patients with shoulder instability or the athletic shoulder, which may not yield the same results. Additionally, we compared the SANE score with only the ASES score. We did not use other more specific tools, such as the Penn Shoulder Score, Western Ontario Rotator Cuff Index, and Western Ontario and McMaster Universities Osteoarthritis Index as these may provide different results, in particular within a specific disease. However, our goal was to evaluate clinical utility, and all of these measures have ≥25 questions, thereby representing a challenge to the common orthopaedic practice. Finally, the SANE and ASES did not perfectly agree, which suggests that the SANE may reflect other constructs not measured by pain and functional ability, including a patient's expectation of what "normal" is for his or her shoulder. This is a limitation when applied among populations but a strength for comparing individual patients over time.

CONCLUSION

Our study demonstrates that the SANE is valid for a range of common shoulder diagnoses to assess patient outcomes across operative and nonoperative treatment of shoulder complaints. The MCID of 15% is similar to that of the ASES (11%), suggesting that the SANE is a simple and efficient tool to assess treatment effects for shoulder disorders. Future studies are warranted to confirm these results and compare across other body parts, diagnoses, and time frames.

REFERENCES

1. American Academy of Orthopaedic Surgeons. Instruments for collection of orthopaedic quality data. http://www.aaos.org/quality/performance_measures/patient_reported_outcome_measures/. Accessed January 2018.
2. Burnham JM, Meta F, Lizzio V, Makhni EC, Bozic KJ. Technology assessment and cost-effectiveness in orthopedics: how to measure outcomes and deliver value in a constantly changing healthcare environment. *Curr Rev Musculoskelet Med*. 2017;10(2):233-239.
3. Chen MM, Rosenkrantz AB, Nicola GN, et al. The Qualified Clinical Data Registry: a pathway to success within MACRA. *AJNR Am J Neuroradiol*. 2017;38(7):1292-1296.
4. Cohen J, Cohen P, West S, Aiken LS. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Vol 3. Mahwah, NJ: Lawrence Erlbaum Associates; 2003.
5. 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(9):1688-1692.
6. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol*. 1992;45(12):1341-1345.
7. Hawkins RJ, Thigpen CA. Selection, implementation, and interpretation of patient-centered shoulder and elbow outcomes. *J Shoulder Elbow Surg*. 2018;27(2):357-362.
8. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415.
9. Jayadevappa R, Malkowicz SB, Wittink M, Wein AJ, Chhatre S. Comparison of distribution- and anchor-based approaches to infer changes in health-related quality of life of prostate cancer survivors. *Health Serv Res*. 2012;47(5):1902-1925.
10. 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.
11. Makhni EC, Hamamoto JT, Higgins JD, et al. How comprehensive and efficient are patient-reported outcomes for rotator cuff tears? *Orthop J Sports Med*. 2017;5(3):2325967117693223.
12. McClure PW, Michener LA. Staged approach for rehabilitation classification: Shoulder Disorders (STAR-Shoulder). *Phys Ther*. 2015;95(5):791-800.
13. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293-307.

14. 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.
15. 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.
16. Richards RR, Bigliani LU, Gartsman GM, Iannotti JP, Zuckerman JD. A standardized assessment of shoulder function. *J Shoulder Elbow Surg*. 1994;3(6):347-352.
17. Sayeed Z, El-Othmani M, Shaffer WO, Saleh KJ. The Medicare Access and CHIP Reauthorization Act (MACRA) of 2015: what's new? *J Am Acad Orthop Surg*. 2017;25(6):e121-e130.
18. Sciascia AD, Morris BJ, Jacobs CA, Edwards TB. Responsiveness and internal validity of common patient-reported outcome measures following total shoulder arthroplasty. *Orthopedics*. 2017;40(3):e513-e519.
19. Shelbourne KD, Barnes AF, Gray T. Correlation of a single assessment numeric evaluation (SANE) rating with modified Cincinnati knee rating system and IKDC subjective total scores for patients after ACL reconstruction or knee arthroscopy. *Am J Sports Med*. 2012;40(11):2487-2491.
20. Snyder AR, Parsons JT, Valovich McLeod TC, Curtis Bay R, Michener LA, Sauers EL. Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part I: disablement models. *J Athl Train*. 2008;43(4):428-436.
21. Tate AR, McClure PW, Young IA, Salvatori R, Michener LA. Comprehensive impairment-based exercise and manual therapy intervention for patients with subacromial impingement syndrome: a case series. *J Orthop Sports Phys Ther*. 2010;40(8):474-493.
22. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
23. Thigpen C, Shanley E. Clinical assessment of upper extremity injury outcomes. *J Sport Rehabil*. 2011;20(1):61-73.
24. Valovich McLeod TC, Snyder AR, Parsons JT, Curtis Bay R, Michener LA, Sauers EL. Using disablement models and clinical outcomes assessment to enable evidence-based athletic training practice, part II: clinical outcomes assessment. *J Athl Train*. 2008;43(4):437-445.
25. Williams GN, Gangel TJ, Arciero RA, Uhorchak JM, Taylor DC. Comparison of the Single Assessment Numeric Evaluation method and two shoulder rating scales: outcomes measures after shoulder surgery. *Am J Sports Med*. 1999;27(2):214-221.
26. Williams GN, Taylor DC, Gangel TJ, Uhorchak JM, Arciero RA. Comparison of the single assessment numeric evaluation method and the Lysholm score. *Clin Orthop Relat Res*. 2000;373:184-192.
27. Winterstein AP, McGuine TA, Carr KE, Hetzel SJ. Comparison of IKDC and SANE outcome measures following knee injury in active female patients. *Sports Health*. 2013;5(6):523-529.
28. Wright AA, Cook CE. Criterion validation of the rate of recovery, single alphanumeric measure, in patients with low back pain. *Physiother Res Int*. 2013;18(2):124-129.