J Shoulder Elbow Surg (2018) 27, 357–362

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

REVIEW ARTICLES

Selection, implementation, and interpretation of patient-centered shoulder and elbow outcomes

Richard J. Hawkins, MD[a,*], Charles A. Thigpen, PhD, PT, ATC[b,c]

[a]*Hawkins Foundation, Greenville, SC, USA*
[b]*ATI Physical Therapy, Greenville, SC, USA*
[c]*Program in Observational Clinical Research in Orthopedics, Center for Effectiveness in Orthopedic Research, Arnold School of Public Health, University of South Carolina, Greenville, SC, USA*

The movement toward a value-based health care market requires comparison of physicians, hospitals, and health systems. Traditionally, process-based measures such as infection and readmission rates have been used. However, these events are uncommon in shoulder and elbow surgery, thus limiting their utility. Patient-reported outcomes (PROs) are a promising measure of quality and have been proposed as a potential metric to compare surgeon performance. However, there are over 25 different PROs for shoulder and elbow conditions. Therefore, the American Shoulder and Elbow Surgeons Value Committee was established to recommend shoulder and elbow PROs in an effort to align their implementation for quality assessment. The committee developed criteria for assessing the outcome measures including that each measure should be patient reported, not requiring clinician input; have published validation and psychometrics; and be standardized and demonstrate ease of use for the patient and clinician. Two sets were suggested: one set for clinical implementation and a more robust set for research purposes. The final recommendation was that all patients should complete the Veterans Rand 12 for general health and the Single Assessment Numeric Evaluation for the specified body region. For patients with shoulder complaints, the American Shoulder and Elbow Surgeons score was recommended, and for those with elbow complaints, the Quick Disabilities of the Arm, Shoulder and Hand score was recommended. More robust disease-specific measures were provided for research purposes. Continued efforts should be made to align these measures across orthopedics to facilitate use of patient outcome measures as a component of value-based health care assessment.

Level of evidence: Narrative Review

Keywords: Shoulder outcomes; elbow outcomes; patient-centered outcomes; selection; implementation; interpretation of outcomes

The American Shoulder and Elbow Surgeons (ASES) Value Committee established a subcommittee on outcomes charged to make recommendations regarding selection, implementation, and interpretation of patient-centered shoulder and elbow outcome scores. In this report we will define quality; establish why we measure outcomes, as well as how they are measured; describe the process of outcome collection; and make recommendations back to the ASES Executive Committee.

This article is exempt from institutional review board review because it is a review article.

*Reprint requests: Richard J. Hawkins, MD, Hawkins Foundation, 200 Patewood Dr, Ste C100, Greenville, SC 29615, USA.

E-mail address: Richard.hawkins@hawkinsfoundation.com (R.J. Hawkins).

1058-2746/$ - see front matter
https://doi.org/10.1016/j.jse.2017.09.022

The outcomes subgroup members consisted of Bernard Morrey, John (JT) Tokish, Guido Marra, Scott Steinmann, Ted Schlegel, and Charles Thigpen. Ex officio members were Rob Bell and Bill Mallon as Past Presidents.

We are aware of the drive to move toward value-based compensation for physicians, where maximum value is the best outcome at the lowest cost. Although cost remains the primary target of improving value in the US medical system, most attempts to measure quality have aimed to reduce expensive cost drivers associated with orthopedic procedures, such as complications, readmissions, or reoperations. However, proponents of value-based medicine such as Michael Porter and Robert Kaplan from Harvard University suggest physicians will be defined by quality measured as patient-reported outcomes.[9,11,12] To date, patient-reported outcomes have not tended to be included in value-based payment models.

Why should we collect outcome scores?

1. Physicians will be judged on quality in the future.
2. It is helpful to know the outcome of treating our own patients.
3. If we do not do it, others will do it for us.
4. Payers and agencies, such as the Centers for Medicare & Medicaid Services (CMS), will demand it.
5. We need to get ahead of the government and other agencies (CMS) with what we, as surgeons, perceive as appropriate scores to measure the outcomes of the treatment of our patients.

Presently, there are many agencies and programs on which physicians are graded from consumer sites such as healthgrades.com, RateMDs.com, and Yelp. For the transition to pay for performance for hospital systems and office practices, surveys such as the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) and Consumer Assessment of Healthcare Providers & Systems (CAHPS) are used to assess the consumer perception of care. Reporting programs such as The Joint Commission's Surgical Care Improvement Project (SCIP) and the CMS Physician Quality Reporting System have financial implications for reporting and meeting thresholds related to surgical treatment. However, these approaches include patient-reported outcomes that have the potential to assess changes in patients' function based on their treatment. As emerging payment models move away from fee-for-service and toward performance-based reimbursement, accurate assessment of patient outcomes is crucial to a comprehensive strategy to deliver on the value proposition.

The committee on researching, defining, establishing, and recommending outcome scores to ASES decided that there should be 2 packages: a basic package for the community at large and a more robust research package. To accomplish this, the subgroup agreed on criteria by which to rank the available shoulder and elbow outcome tools. The outcomes subgroup members reviewed applicable papers and reports with particular attention paid to the psychometrics of shoulder and elbow scoring systems including reliability, validity, and responsiveness of each tool. Outcome tools with better psychometrics, especially regarding precision, such as the minimal clinically important difference, were ranked higher, given the desire to eventually use them to determine value.[1,10,16,18,20,24]

The committee, following many phone calls and meetings, established the following guidelines for selection:

1. Patient-reported outcomes
2. Validated scores
3. Good psychometrics
4. Ease of use for patient (ie, brief)
5. Ease of scoring and understanding for physician
6. Standardized use nationally
7. Consideration of cost

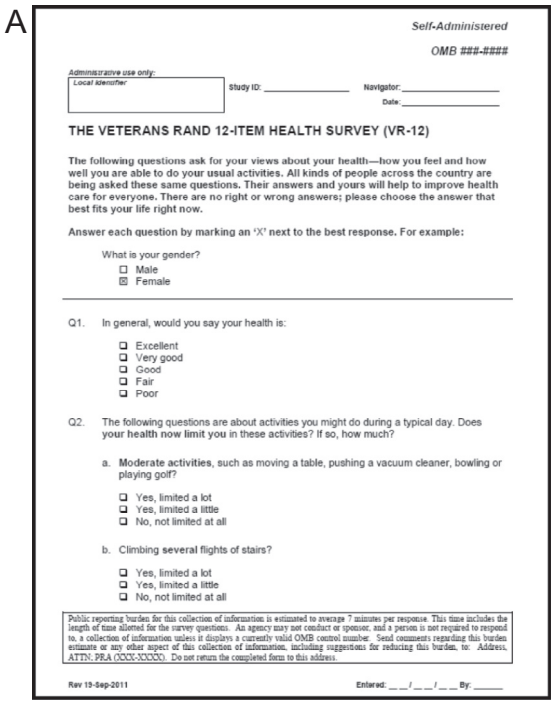
Shoulder

There are approximately 25 shoulder scores used worldwide. These were narrowed to a few following our guidelines for selection. The group agreed a generic quality-of-life score would be required; a joint-specific score would be required; and for research, more sophisticated scores would be needed. The committee debated whether we should include more than 1 score for these different areas or settle on just 1 score.

Generic quality of life

Generic health-related quality-of-life measures such as the Short Form 36 (SF-36), Short Form 12 (SF-12), EuroQol-5 Dimension (EQ-5D), Veterans Rand 12 (VR-12), and Patient-Reported Outcomes Measurement Information System (PROMIS) 10 are important for establishing baseline health status and function beyond comorbidities and have been shown to influence orthopedic outcomes. These measures are also the primary outcome tools that have been used for value calculations, such as quality-adjusted life-years, providing the ability to compare total shoulder arthroplasty with total hip arthroplasty or breast cancer treatment.[2,13] This is thought to be meaningful to payers such as CMS to aid in the allocation of dollars to treatments that offer the most improvement in patient health status over time. In addition, these scores can be used at baseline for purposes of risk adjustment and stratification as emerging payment models are developed.

The SF-36, SF-12, EQ-5D, PROMIS 10, and VR-12 (Fig 1, *A*) were all discussed, and our committee settled on the VR-12 as the generic quality-of-life score. The primary advantages of this tool are that it has US population norms, is a part of the Medicare Health Outcomes Survey and is directly comparable with the SF-12, and is in the public domain and therefore available at no cost. The VR-12 consists of 12 questions with a Likert response system. The SF-36, SF-12, and EQ-5D all carry financial implications. The PROMIS 10 is available in the public domain at no cost, but its use is not widespread at this point. The EQ-5D, although widely used



Self-Administered

OMB ###-####

Administrative use only:

Local Identifier

Study ID: _____ Navigator: _____

Date: _____

THE VETERANS RAND 12-ITEM HEALTH SURVEY (VR-12)

The following questions ask for your views about your health—how you feel and how well you are able to do your usual activities. All kinds of people across the country are being asked these same questions. Their answers and yours will help to improve health care for everyone. There are no right or wrong answers; please choose the answer that best fits your life right now.

Answer each question by marking an 'X' next to the best response. For example:

What is your gender?

☐ Male

☒ Female

Q1. In general, would you say your health is:

❑ Excellent
❑ Very good
❑ Good
❑ Fair
❑ Poor

Q2. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

a. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf?

❑ Yes, limited a lot
❑ Yes, limited a little
❑ No, not limited at all

b. Climbing several flights of stairs?

❑ Yes, limited a lot
❑ Yes, limited a little
❑ No, not limited at all

Public reporting burden for this collection of information is estimated to average 7 minutes per response. This time includes the length of time allotted for the survey questions. An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: Address, ATTN: PRA (XXX-XXXX). Do not return the completed form to this address.

Rev 19-Sep-2011

Entered: __ __ / __ __ / __ __ By: _____

B

ASES Shoulder Evaluation- Pre-Treatment

Patient Pain & Instability Self-Evaluation Section

1. How bad is your pain today? (0 is no pain, 10 is the most painful) _____

Patient Activity of Daily Living Self-Evaluation Section

		Unable to do	Very difficult to do	Somewhat difficult	Not difficult
1. Put on a coat.	Left Shoulder				
	Right Shoulder				
2. Sleep on painful side	Left Shoulder				
	Right Shoulder				
3. Wash back/do up bra in back	Left Shoulder				
	Right Shoulder				
4. Manage toileting	Left Shoulder				
	Right Shoulder				
5. Comb hair	Left Shoulder				
	Right Shoulder				
6. Reach a high shelf	Left Shoulder				
	Right Shoulder				
7. Left 10 lbs. above shoulder	Left Shoulder				
	Right Shoulder				
8. Throw a ball overhand	Left Shoulder				
	Right Shoulder				
9. Do usual work	Left Shoulder				
	Right Shoulder				
10. Do usual sport	Left Shoulder				
	Right Shoulder				

C

SANE

What percentage of normal is your shoulder?

0% _____100%

D

WESTERN ONTARIO

ROTATOR CUFF INDEX

(WORC)©

A disease-specific quality of life measurement tool for patients with rotator cuff disease

Copyright © 1998 (#474673) A. Kirkley MD, S. Griffin, CSS, C. Alvarez, MD

Figure 1 (**A**) Veterans Rand 12 (*VR-12*) score. (**B**) American Shoulder and Elbow Surgeons (*ASES*) score. (**C**) Single Assessment Numeric Evaluation (*SANE*) score. (**D**) Western Ontario Rotator Cuff (*WORC*) score.

in Europe, is not prevalent in the United States and does not have a score that is easily understood clinically.

Patient-Reported Outcomes Measurement Information System (PROMIS)

The committee agreed that PROMIS is a rapidly evolving area with a large amount of National Institutes of Health funding requiring attention. PROMIS presently has a generic quality-of-life score, the PROMIS 10. In orthopedics, work is being done on the foot and ankle with some publications, and some of our hand friends are working on the PROMIS score; however, it has not been generally applicable to other musculoskeletal problems. As part of the PROMIS program, there is computerized adaptive testing. This method allows one to narrow down a large bank of questions to a small number of questions to apply to a certain disease or joint following a computer program. The American Academy of Orthopaedic Surgeons Performance Committee understood that this would be potentially useful in the future but has not been used much in the past other than for the foot and ankle. Thus, the ASES Committee recommended the necessity of keeping an open mind as we move forward regarding PROMIS, but it is not ready for implementation for the shoulder and elbow at this time.

Joint-specific score

We analyzed all the shoulder joint–specific scores and settled on the ASES score and Oxford Shoulder Score. The ASES score consists of 11 questions with a 100-point total scale, with 50 points for function and 50 points for pain.[10,14,18,24] The pain question is 1 question: "What is your pain today?" A high score is best. The ASES score (Fig 1, *B*) has demonstrated good psychometrics and is the standard in North America. The Oxford Shoulder Score consists of 12 questions with a total score of 60, with a higher score being better.[3,8] Likewise, it has good psychometrics, but it is used almost exclusively in the United Kingdom and is used to some degree elsewhere in Europe, not in North America. We eliminated many scores because they did not fit into our guidelines. For example, the Constant score, used at least historically in Europe, is not a totally patient-reported outcome; it requires clinical input to measure strength.

On the basis of this analysis, the committee recommended the ASES score as the shoulder-specific score to the ASES Executive Committee, particularly for North America. For those who choose to use the Oxford Shoulder Score, such as in Europe, the committee agreed it was acceptable.

There was extensive discussion regarding the Single Assessment Numeric Evaluation (SANE) score (Fig 1, *C*). The SANE score simply asks, "What percentage of normal is your shoulder?" A rating on a scale from 0% to 100% is given, with a higher score being better. Emerging evidence suggests it provides a simple, comparative score for knee function (Lysholm and International Knee Disability Classification [subjective]) and shoulder function (ASES score).[4,17,21-23] We have recently demonstrated that the SANE score is a reliable measure and provides similar scores and responsiveness compared with the ASES score across a range of patients undergoing nonoperative treatment, rotator cuff repair, or total shoulder replacement at the Steadman Hawkins Clinic within the Greenville Health System.[19] The committee agreed on the simplicity of the SANE score given its low burden on the clinical and administrative staff for implementation.

In the final analysis, the basic package recommended for a shoulder score is as follows:

1. Generic quality of life: VR-12
2. Joint specific: ASES score
3. SANE score

For research purposes or more sophisticated analysis, the committee recommended the Western Ontario Rotator Cuff (WORC) score for rotator cuff disease (Fig 1, *D*). This score consists of 21 questions, with 0 being the best and 100 being the worst. It has excellent psychometrics. For instability, we recommended the Western Ontario Stability Index, a similar score to the WORC score, and finally, the Western Ontario Osteoarthritis Score. We also included the Pennsylvania Shoulder Score (PENN score) as it demonstrates good psychometrics with improved ceiling and floor effects. The PENN score has 25 questions, including 20 functional questions worth 60 points, 3 pain questions worth 30 points, and 1 satisfaction question worth 10 points. This was an excellent score to be included for those who wish to pursue research. In addition, the ASES score can be calculated from the PENN score, limiting the patient and administrative burden.

The final recommendation for the research package for the shoulder score is as follows:

1. Generic quality-of-life score: VR-12 score
2. Joint-specific score: ASES score
3. SANE score
4. Disease-specific score (WORC, Western Ontario Stability Index, Western Ontario Osteoarthritis Score)
5. PENN score

Elbow

For the elbow, we followed the same process used for the shoulder. After analysis and discussion, the Quick Disabilities of the Arm, Shoulder and Hand score was recommended for the elbow as it shows good psychometrics and is widely used worldwide.[18] As a disability score, a higher score indicates worse patient function. The Oxford Elbow Score was considered as it demonstrates good psychometrics; however, given its limited use outside Europe, it was not recommended.[6,7]

In the final analysis, our recommendation on a basic package for the elbow is as follows:

1. Generic quality-of-life score: VR-12
2. Quick Disabilities of the Arm, Shoulder and Hand questionnaire, which consists of 11 questions and is a disability score, so a higher score is worse; has great psychometrics; and is used worldwide
3. SANE score

For those who wish to pursue research regarding the elbow, the Mayo Elbow Performance Score—which consists of 12 questions, 6 of which require some physician input—has been reported with patient input over the phone.[5] Given the additional administrative and clinical burden, it is recommended for research purposes.

Implementation

It was not specifically the charge of our committee to deal with implementation. Implementation in our practices remains a significant challenge. However, implementing the aforementioned scores may be done by paper, scanning, or a computer and should be web based. There are software programs presently available, such as Socrates (Ortholink Pty Ltd, Sydney, NSW, Australia), Oberd (Oberd, Raleigh, NC, USA), RedCap (Vanderbilt University, Nashville, TN, USA), and EPIC (Epic Systems Corporation, Verona, WI, USA), that can be used. In our system at the Steadman Hawkins Clinic within the Greenville Health System, outcome measures are embedded in EPIC through MyChart (Epic Systems Corporation) and include those described herein. As a basic package, collecting VR-12, ASES, and SANE scores remains a challenge. Even if we all obtain just the SANE score, consisting of 1 question, it would be worthwhile.

The committee agreed that we need to engage other stakeholders such as the Arthroscopy Association of North America, American Orthopaedic Society for Sports Medicine, American Physical Therapy Association, American Society of Shoulder and Elbow Therapists, and National Athletic Trainers' Association. We have had several stakeholder meetings, including one in Chicago on Friday, February 20, 2015, and one at a recent American Academy of Orthopaedic Surgeons meeting in Orlando, with representatives from the Arthroscopy Association of North America, ASES, American Society for Surgery of the Hand, and American Orthopaedic Society for Sports Medicine to suggest and agree on scores.

Finally, our legislators have eliminated the sustainable growth rate formula, and the Medicare Access and CHIP (Children's Health Insurance Program) Reauthorization Act of 2015 has been put in place.[15] This new quality-based program consists of a merit-based incentive payment system, and this replaced the Physician Quality Reporting System beginning in January 2017. In addition, there will be alternative payment models that will be pursued in the future. All of these address value and require outcome performance measures.

Cost, compliance, and risk stratification are 3 examples of significant remaining challenges that must be addressed before real value-based reimbursement programs can be effectively implemented. Defining which outcome measures to use is only the beginning. We now must move to evaluate patient-reported outcome scores as performance measures. This requires pretreatment and post-treatment measures over a diverse and large number of patients across specific diseases. These measures should then be risk adjusted and shown to differentiate providers to become performance measures.

Disclaimer

The 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.

References

1. 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):S174-88. http://dx.doi.org/10.1002/acr.20630
2. Bachman D, Nyland J, Krupp R. Reverse-total shoulder arthroplasty cost-effectiveness: a quality-adjusted life years comparison with total hip arthroplasty. World J Orthop 2016;7:123-7. http://dx.doi.org/10.5312/wjo.v7.i2.123
3. Christiansen DH, Frost P, Falla D, Haahr JP, Frich LH, Svendsen SW. Responsiveness and minimal clinically important change: a comparison between 2 shoulder outcome measures. J Orthop Sports Phys Ther 2015;45:620-5. http://dx.doi.org/10.2519/jospt.2015.5760
4. Cunningham G, Lädermann 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;31:1688-92. http://dx.doi.org/10.1016/j.arthro.2015.03.010
5. Cusick MC, Bonnaig NS, Azar FM, Mauck BM, Smith RA, Throckmorton TW. Accuracy and reliability of the Mayo Elbow Performance Score. J Hand Surg Am 2014;39:1146-50. http://dx.doi.org/10.1016/j.jhsa.2014.01.041
6. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. Comparative responsiveness and minimal change for the Oxford Elbow Score following surgery. Qual Life Res 2008;17:1257-67. http://dx.doi.org/10.1007/s11136-008-9409-3
7. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-73. http://dx.doi.org/10.1302/0301-620x.90b4.20290
8. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. J Bone Joint Surg Br 1996;78:593-600.

9. Kaplan RS, Porter ME. How to solve the cost crisis in health care. Harv Bus Rev 2011;89:46-52, 54, 56-61 passim.
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;87:2006-11. http://dx.doi.org/10.2106/JBJS.C.01624
11. Porter ME. Value-based health care delivery. Ann Surg 2008;248:503-9. http://dx.doi.org/10.1097/SLA.0b013e31818a43af
12. Porter ME, Lee TH. From volume to value in health care: the work begins. JAMA 2016;316:1047-8. http://dx.doi.org/10.1001/jama.2016.11698
13. Renfree KJ, Hattrup SJ, Chang YH. Cost utility analysis of reverse total shoulder arthroplasty. J Shoulder Elbow Surg 2013;22:1656-61. http://dx.doi.org/10.1016/j.jse.2013.08.002
14. Richards RR, An KN, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, et al. A standardized method for the assessment of shoulder function. J Shoulder Elbow Surg 1994;3:347-52.
15. 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:e121-30. http://dx.doi.org/10.5435/jaaos-d-17-00151
16. Schmidt S, Ferrer M, González M, González N, Valderas JM, Alonso J, et al. Evaluation of shoulder-specific patient-reported outcome measures: a systematic and standardized comparison of available evidence. J Shoulder Elbow Surg 2014;23:434-44. http://dx.doi.org/10.1016/j.jse.2013.09.029
17. 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;40:2487-91. http://dx.doi.org/10.1177/0363546512458576
18. Smith MV, Calfee RP, Baumgarten KM, Brophy RH, Wright RW. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. J Bone Joint Surg Am 2012;94:277-85. http://dx.doi.org/10.2106/JBJS.J.01744
19. Thigpen CA, Shanley E, Tokish JM, Kissenberth MJ, Tolan SJ, Hawkins RJ. Validity and responsiveness of the Single Alpha-Numeric Evaluation (SANE) for shoulder patients. Orthop J Sports Med 2017;5(Suppl 6):2325967117S00256.
20. van Kampen DA, Willems WJ, van Beers LW, Castelein RM, Scholtes VA, Terwee CB. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). J Orthop Surg 2013;8:40. http://dx.doi.org/10.1186/1749-799x-8-40
21. 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;27:214-21.
22. 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):184-92.
23. 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;5:523-9. http://dx.doi.org/10.1177/1941738113499300
24. Wright RW, Baumgarten KM. Shoulder outcomes measures. J Am Acad Orthop Surg 2010;18:436-44.