BRIEF ARTICLE

Evaluation of utility in shoulder pathology: Correlating the American Shoulder and Elbow Surgeons and Constant scores to the EuroQoL

Blaine T Bafus, Richard E Hughes, Bruce S Miller, James E Carpenter

Blaine T Bafus, Department of Orthopaedic Surgery, Cleveland Clinic Foundation, Cleveland, OH 44115-2325, United States
Richard E Hughes, Bruce S Miller, James E Carpenter, Department of Orthopaedic Surgery, University of Michigan, Ann Arbor, MI 48109-2200, United States
Correspondence to: Richard E Hughes, PhD, Laboratory for Optimization and Computation in Orthopaedic Surgery, University of Michigan, 2003 Biomedical Science Research Building, 109 Zina Pitcher Pl, Ann Arbor, MI 48109-2200, United States. rehughes@med.umich.edu
Telephone: +1-734-6156410   Fax: +1-734-6470003

Abstract

AIM: To study whether health utility scores can be derived from shoulder-specific scores.

METHODS: Authors investigated two questions: (1) do the American Shoulder and Elbow Surgeons (ASES) score and the Constant score correlate with the EuroQoL (EQ-5D), a measure of health utility? (2) can the ASES and Constant scores be obtained from a complete study sample without bias? Thirty subjects with various shoulder diagnoses completed ASES, Constant, and EQ-5D instruments. Pearson correlations were calculated to assess the associations between EQ-5D score and ASES and Constant scores.

RESULTS: The correlation between EQ-5D score and ASES score was 0.60 ($P < 0.001$); it was 0.54 for EQ-5D and Constant scores ($P < 0.003$). A multiple regression model containing ASES score, Constant score, age, and gender failed to adequately predict EQ-5D. Moreover, 25% of patients meeting the inclusion criteria did not complete the ASES questionnaire because they did not feel that specific questions, such as "do usual sport -list" and "throw ball overhand," applied to them.

CONCLUSION: Authors' results do not support the use of the ASES and Constant scores in predicting EuroQol health utility values. However, the Constant score was more suitable for this patient population because all patients were able to complete it.

Peer reviewers: V Prem Kumar, Professor, Department of Orthopaedic Surgery Yong Loo Lin School of Medicine, National University of Singapore NUHS Tower Block, Level 11, 1E Kent Ridge Road, Singapore 119228, Singapore; Jason Scott Scibek, PhD, ATC, Assistant Professor, Department of Athletic Training Co-Director, Kristen L McMaster Motion Analysis Laboratory, Duquesne University, 600 Forbes Avenue, 118 HLSB Pittsburgh, PA 15282, United States

Bafus BT, Hughes RE, Miller BS, Carpenter JE. Evaluation of utility in shoulder pathology: Correlating the American Shoulder and Elbow Surgeons and Constant scores to the EuroQoL. *World J Orthop* 2012; 3(3): 20-24 Available from: URL: http://www.wjgnet.com/2218-5836/full/v3/i3/20.htm DOI: http://dx.doi.org/10.5312/wjo.v3.i3.20

INTRODUCTION

Over the past several years, outcome reporting has had enhanced importance as patients and payers demand evidence supporting medical interventions. In 1993, the United States Public Health Service established the Quality-Adjusted Life Year (QALY) as the universal measure of health effectiveness[1]. The QALY is a unit measurement accounting for the patient's perceived health-related quality of life. The Panel on Cost-Effectiveness in Health and Medicine also recommended the generic health-state instruments, the Health Utilities Index (HUI) and the EuroQoL (EQ-5D), as suitable for obtaining the patient preference for a health state[1]. These measures can be used to construct utility values for use in cost-utility models in medical decision making. In healthcare, utility is the preference that patients have for a particular health status. Recently, an investigation analyzing utility and cost-effectiveness in rotator cuff repair established that the EQ-5D is more sensitive than the HUI in rotator cuff pathology[2].

The EQ-5D is an internationally validated general measure of health-related quality of life composed of a one-page questionnaire and a one-page visual analogue scale. The EQ-5D questionnaire examines five dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression[3]. Each of the dimensions examined has three levels, reflecting increasing health impact, and associated five-digit descriptors (ranging from 11 111 for perfect health to 33 333 for worst state imagined), yielding up to 243 different health states[4]. Further work by the US Agency for Healthcare Research and Quality (AHRQ) established a US population-based preference weighting system for the EQ-5D used to derive utility[5].

Most investigations of shoulder pathology measure outcomes using a disease-specific outcome instrument, such as the American Shoulder and Elbow Surgeons (ASES) score or the Constant score, and do not include a direct measurement of utility[6,7]. In an effort to derive utility from previously published studies that used a general health assessment without a direct measurement of utility, algorithms have been developed converting two general health assessments, the SF-36 and SF-12, to the HUI2 and the HUI3, respectively[8,9]. To date, no studies have attempted to derive a utility from the commonly used disease-specific shoulder scores. The primary purpose of this study was to formulate an algorithm to convert two of the most commonly used shoulder scores, the ASES and Constant scores, to an EQ-5D utility score to allow for future cost-utility analyses of the existing shoulder literature. A secondary objective was to determine if the ASES and Constant scores could be obtained from every patient in our sample, since an instrument that is completed by only a subset of the sample may provide biased data.

MATERIALS AND METHODS

Subjects

This study was approved by the authors' Institutional Review Board. A power analysis was performed, indicating a need for 13 patients to obtain 80% power to detect a correlation of 0.8 or better when controlling for type I error at 0.05 and using a 2-tailed test. We sought to enroll 50 subjects for increased statistical power. Subjects were recruited from a university orthopaedic sports medicine clinic based on a referral for shoulder pathology and were enrolled by order of convenience. Inclusion criteria were age 18 years or older, history of shoulder pain, no history of high-energy injury to the upper extremity within one year, no history of surgery on either shoulder, and an ability to communicate effectively with the examiner and give written informed consent.

Data collection

Data were collected at the sports medicine clinic of the University of Michigan between November 2007 and April 2008. All data were collected by a single examiner. The EQ-5D, ASES, and Constant instruments were assembled into packets in random order as determined by a random number generator. The EQ-5D had six questions. Five of them are three-level ordinal scales. The sixth is a visual analog scale with anchors "best imaginable health state" and "worst imaginable health state." The ASES instrument had both patient and physician components. The constant instrument asks the patient questions about pain in the shoulders during normal activities and whether the shoulders interfere with work, recreation/sport, and sleep. It also asks how high the hand can be raised (from belt/waist to above the head). The physician component evaluated abduction, forward elevation, internal rotation, external rotation, and abduction strength. The ASES instrument asked the patient about pain, instability, and activities of daily living. Questions regarding activity of daily living include rating the ability to perform ten activities on a four-point scale ("unable to do, very difficult to do, somewhat difficult, and not difficult"): "put on a coat, sleep on your painful or affected side, wash back/do up bra in back, manage toileting, comb hair, reach a high shelf, lift 10 lbs. above shoulder, throw a ball overhand, do usual work – list, and do usual sport-list." The physician portion of the ASES requires assessment of range of motion, signs, strength, and instability.

Subjects were asked to do their best with questions they felt did not apply to their situation. The physical examination component of the Constant score was performed as described by Constant and Murley[6]. Strength was objectively assessed at 90 degrees abduction in the scapular plane using the IsoForce Control (IFC) dynamometer (Medical Device Solutions AG, Oberburg, Switzerland). The IFC records a minimum force of 10 N and yields peak and average force produced over 5 s, with 10 recordings per second. Newtons were converted to pounds by dividing by 4.448 for the Constant score. The EQ-5D was scored using the AHRQ scoring algorithm for Excel (Microsoft, Redmond, Washington)[4]. The ASES and Constant instruments were scored according to their original descriptions[6,7]. We used the

original Constant scoring[6] method because the revised system was published during the course of our study[10]. Subjects with incomplete instruments were excluded from all data analysis.

Statistical analysis
Data were analyzed with MINITAB v.15.1.0.0 (MINITAB, State College, Pennsylvania). Pearson product-moment correlation coefficients were computed to assess association between ASES, Constant, and EQ-5D scores. A multiple regression model was constructed to predict EQ-5D from the ASES score, Constant score, and subject age and gender. Normality was assessed using probability plots of each score.

RESULTS

Demographics
Forty-four patients were approached to participate in the study. Four did not meet the inclusion criteria, giving a total of 40 patients enrolled in the study. Ten were excluded for incomplete surveys. The remaining 30 subjects included 12 males and 18 females. The average age was 48.8 years (range, 25-71 years). The subjects had a broad range of diagnoses; however, one-half suffered from rotator cuff pathology (Table 1). Full-thickness rotator cuff tears were not distinguished from other tendinopathy for this study.

Descriptive statistics
Means for the ASES, constant, and EQ-5D scores were 55.1 (16.5 SD), 60.8 (19.0), and 0.765 (0.108), respectively. Scores for each instrument were tabulated and grouped by diagnosis with the mean, standard deviation, and range (Table 2). For diagnoses represented by only one subject, no standard deviation or range is presented. Three subjects who were diagnosed by the attending surgeon with two diagnoses are presented in the data sets of both diagnoses, as we were unable to ascertain which diagnosis had a greater impact in the data for those individuals.

Model development
Pearson correlations were calculated between the ASES, constant, and EQ-5D total scores. The ASES and EQ-5D correlation was $r = 0.60$, $P < 0.001$ (Figure 1). The Constant and EQ-5D correlation was $r = 0.54$, $P < 0.003$ (Figure 2). Multiple regression modeling yielded a coefficient of multiple determination (R^2) of 0.44, meaning that a model containing ASES, Constant, and age and gender only predicts 44% of the variance. To assess the predictive ability of the regression model, the prediction interval was computed for both men and women at the mean ASES score, Constant score, and age. The prediction intervals were (0.60, 0.96) and (0.59, 0.94) for men and women, respectively. These represent 35% of the full range of the EQ-5D scale.

Completeness of data
Ten patients (25%) did not complete the ASES ques-

Table 1 Patient demographics

Patient	Gender	Age (yr)	Diagnosis
1	Female	48	Rotator cuff
2	Female	56	Rotator cuff, labral tear
3	Male	66	Arthritis
4	Female	30	Labral tear
5	Male	48	Rotator cuff
6	Female	52	Rotator cuff, labral tear
7	Female	56	Rotator cuff
8	Female	52	Adhesive capsulitis
9	Male	38	Impingement
10	Female	25	Instability
11	Female	50	Rotator cuff
12	Female	69	Rotator cuff
13	Female	35	Impingement
14	Female	44	SC subluxation
15	Female	71	Rotator cuff
16	Male	30	Labral tear
17	Female	51	Rotator cuff
18	Female	48	Impingement
19	Male	68	Rotator cuff
20	Male	52	Rotator cuff, biceps
21	Female	73	Arthritis
22	Male	32	Rotator cuff
23	Male	38	Arthritis
24	Female	25	Instability
25	Female	58	Rotator cuff
26	Male	36	Unknown
27	Male	38	Arthritis
28	Female	63	Rotator cuff
29	Male	52	Rotator cuff
30	Male	59	Arthritis

Rotator cuff: Rotator cuff pathology; Biceps: Biceps tendon pathology; SC: Sternoclavicular joint pathology.

tionnaire. Specifically, they left either "do usual sport - list" or "throw ball overhand" (or both) blank. When questioned by the investigator, many of these patients indicated that at their age they did not consider themselves as "playing a sport" and they did not throw balls overhand. This was particularly true of patients with rotator cuff pathology, as they tended to be older.

DISCUSSION
We were to unable to develop a model for converting ASES and Constant scores to EQ-5D utility scores. We found statistically significant, but only modestly strong, correlations between the EQ-5D and ASES ($r = 0.60$) and Constant ($r = 0.54$) scores. Although a multiple regression model could account for 44% of the variance, its prediction intervals were too large for this model to be useful (35% of the full EQ-5D range). We also found that 10 of the 40 patients meeting the inclusion criteria (25%) did not complete the ASES questionnaire because they did not feel that the questions regarding "do usual sport - list" and "throw ball overhand" applied to them. This suggests the Constant score is more suitable for this patient population.

It is unfortunate that we were unable to develop a method to convert ASES and Constant scores to utility values because of the great importance of cost-utility

	n	ASES mean (SD)	ASES range	Constant mean (SD)	Constant range	EQ-5D mean (SD)	EQ-5D range
Adhesive capsulitis	1	58.5[1]	[1]	58.0[1]	[1]	0.827[1]	[1]
Arthritis	5	59.0 (8.3)	49.2-70.8	56.8 (19.2)	32.2-76.0	0.720 (0.086)	0.597-0.827
Biceps tendon[2]	1	30.8[1]	[1]	51.7[1]	[1]	0.816[1]	[1]
Impingement	3	40.5 (6.3)	33.3-45.0	67.0 (10.0)	55.7-75.0	0.705 (0.038)	0.678-0.748
Instability	2	51.7 (28.3)	71.7-31.7	56.4 (39.0)	28.8-84.0	0.797 (0.042)	0.767-0.827
Labral tear[2]	4	74.0 (16.0)	54.2-93.3	77.0 (8.8)	64.7-83.8	0.864 (0.090)	0.813-1.000
Rotator cuff	15	54.7 (18.8)	28.7-93.3	57.9 (20.8)	26.0-85.0	0.783 (0.133)	0.416-1.000
SC subluxation	1	46.7[1]	[1]	69.6[1]	[1]	0.678[1]	[1]
Unknown	1	78.7[1]	[1]	83.0[1]	[1]	0.816[1]	[1]

Table 2 Outcome instrument data presented by diagnosis

[1]Insufficient data to calculate. [2]Some or all of the subjects within this diagnosis also had rotator cuff pathology. ASES: American Shoulder and Elbow Surgeons; SC: Sternoclavicular joint; Biceps: Biceps tendon pathology; Rotator cuff: rotator cuff pathology.
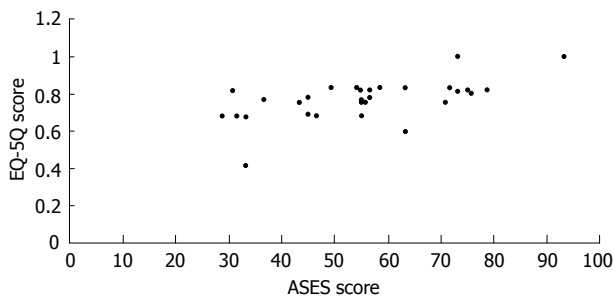
Figure 1 Scatterplot of EuroQol (EQ-5D) score *vs* American shoulder and elbow surgeons (ASES) score.
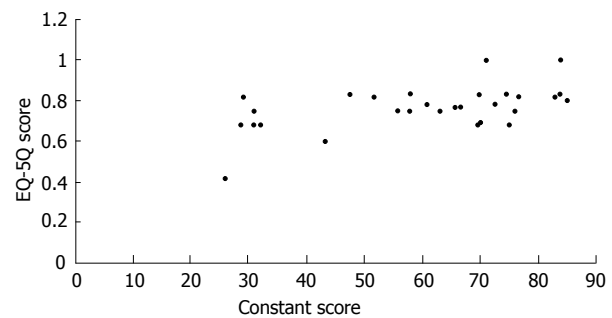
Figure 2 Scatterplot of EuroQol (EQ-5D) score *vs* Constant score.

modeling in modern health care economics. Developing resource allocation strategies in healthcare that rely solely on cost can deprive the public from having access to technologies that significantly improve their health at modest costs. Health care economics has developed cost-utility analysis methods based on decision theory[11] that incorporate both improvements in the quality of life ("utility" which can be measured by EQ-5D scores) and cost of care. For shoulder surgeons to demonstrate their treatments are cost-effective, utility must be quantified. Since so much of the clinical shoulder literature reports ASES and Constant scores, it would be very useful to have a method to convert these scores to utility values.

It is not surprising that there is a weak correlation between the shoulder instruments and the EQ-5D. A previous investigation identified a Pearson correlation of 0.50 between the Constant and ASES instruments, a low correlation, but indicative of some intuitive similarity between the instruments as they both relate to shoulder pathology[12]. Similarly, the EQ-5D asks subjects about problems walking, problems with self care, problems performing usual activities, the presence of pain and discomfort, and if they have any anxiety or depression. It is easy to extrapolate that shoulder pathology could impact some or all of these domains, as we observed in this investigation.

The high percentage of patients refusing to answer two sports-related questions indicates the potential for biased estimates of mean ASES scores in a study population. Since the two questions that posed the greatest challenge involved sports, the ASES instrument may

be best suited to studies limited to athletes. However, a high rate of incomplete data forms presents the strong possibility of bias when conducting studies of broad populations or older patients. This is likely to be a problem in studies of rotator cuff tear patients, who tend to be older. Studies of younger patients being treated for impingement syndrome or instability may have a higher rate of completing all ASES questions.

While multiple diagnoses are represented here, the large percentage of patients with rotator cuff pathology allows for a limited comparison of our instrument scores with the published results for this population. In fact, the high proportion of patients with cuff pathology may bias our mean scores relative to what other occurs in patients being treated in different orthopaedic practices. A recent investigation that examined outcomes after arthroscopic rotator cuff repair and that utilized both the Constant and ASES instruments cited mean preoperative values of 54 and 44, respectively[13]. This compares to our Constant score mean of 57.9 and ASES mean of 54.7. The only published investigation utilizing the EQ-5D in patients undergoing rotator cuff repair noted a preoperative mean score of 0.563[2]; however, our mean was 0.784. While our constant scores are similar to the published literature, our ASES and EQ-5D scores are not. It is possible that this discrepancy can be accounted for by our exclusion of nonathletic individuals as a result of the ASES bias. It may be that the people who were unable to answer the sports and throwing questions on the ASES also have more modest expectations of arm

function, which produces a higher overall satisfaction with their health state.

Our investigation has several limitations, foremost being the ASES bias against the nonathletic individual. Twelve and one-half percent of enrolled subjects were disqualified for an inability to complete the question regarding the impact of their disease on personal sporting behavior. When asked, most indicated they simply did not participate in sports. No attempt was made by the examiner to coerce answers; therefore, these subjects were later disqualified for incomplete questionnaires. Other investigators have noted similar problems with the Shoulder Severity Index regarding driving in a population that did not drive[14]. This problem with the ASES may indicate a weakness with the instrument.

A limitation of the Constant score is the reliability of measuring shoulder abduction strength[10]. We did not assess the inter- or intra-rater reliability of the IFC used in our study, so this limitation applies to our study.

This study attempted to correlate the ASES and Constant scores with the EQ-5D to allow for future cost-utility analyses of the existing shoulder data. We were unable to derive a useful algorithm to support our hypothesis. The limitations of this study reflect the limitations of our discipline and illustrate the need for continued improvements in the shoulder instruments we currently use if they are going to be applicable to the general population.

COMMENTS

Background
Healthcare economic studies of shoulder treatments require estimates of health-related quality of life, which is measured using questionnaires. One common questionnarie is the EuroQol (EQ-5D). However, the most dominant methods for assessing shoulder function are the Constant score and American Shoulder and Elbow Surgeons (ASES) score.

Research frontiers
Developing a way to convert Constant and ASES scores to EQ-5D values would facilitate economic evaluation shoulder disorder treatments.

Innovations and breakthroughs
The study was unable to develop a mathematical formula to convert Constant and ASES scores to EQ-5D scores. The data suggest that such a conversion may not be possible. It was also found that some patients did not complete the ASES questionnaire because of its sports-related questions.

Applications
The significance of the study is that future studies of shoulder injuries should include the EuroQol questionnaire rather than rely on the Constant and ASES tools.

Terminology
EQ-5D: EuroQol; ASES: American Shoulder and Elbow Surgeons.

Peer review
This study attempted to correlate two standard shoulder scores with one quality of life score. The study population was biased towards rotator cuff tear patients; a population of younger patients with a high proportion of impingement syndrome cases may produce different results.

REFERENCES

1   **Weinstein MC**, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-effectiveness in Health and Medicine. *JAMA* 1996; **276**: 1253-1258
2   **Vitale MA**, Vitale MG, Zivin JG, Braman JP, Bigliani LU, Flatow EL. Rotator cuff repair: an analysis of utility scores and cost-effectiveness. *J Shoulder Elbow Surg* 2007; **16**: 181-187
3   **Brooks R**. EuroQol: the current state of play. *Health Policy* 1996; **37**: 53-72
4   Agency for Healthcare Research and Quality. Calculating the U.S. population-based EQ-5D index score, August 2005. Available from: URL: http://www.ahrq.gov/rice/EQ5D-score.htm
5   Agency for Healthcare Resarch and Quality. U.S. valuation of the EuroQol EQ-5D health states, December 2005. Available from: URL: http://www.ahrq.gov/rice/EQ5Dproj.htm
6   **Constant CR**, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987; : 160-164
7   **Richards RR**, An K-N, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, Iannotti JP, Mow VC, Sidles JA, Zuckerman JD. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 1994; **3**: 347-352
8   **Nichol MB**, Sengupta N, Globe DR. Evaluating quality-adjusted life years: estimation of the health utility index (HUI2) from the SF-36. *Med Decis Making* 2001; **21**: 105-112
9   **Sengupta N**, Nichol MB, Wu J, Globe D. Mapping the SF-12 to the HUI3 and VAS in a managed care population. *Med Care* 2004; **42**: 927-937
10 **Constant CR**, Gerber C, Emery RJ, Søjbjerg JO, Gohlke F, Boileau P. A review of the Constant score: modifications and guidelines for its use. *J Shoulder Elbow Surg* 2008; **17**: 355-361
11 **Petitti DB**. Meta-analysis, decision analysis, and cost-effectiveness analysis. New York: Oxford University Press, 2000
12 **Placzek JD**, Lukens SC, Badalanmenti S, Roubal PJ, Freeman DC, Walleman KM, Parrot A, Wiater JM. Shoulder outcome measures: a comparison of 6 functional tests. *Am J Sports Med* 2004; **32**: 1270-1277
13 **Lee E**, Bishop JY, Braman JP, Langford J, Gelber J, Flatow EL. Outcomes after arthroscopic rotator cuff repairs. *J Shoulder Elbow Surg* 2007; **16**: 1-5
14 **Beaton DE**, Richards RR. Measuring function of the shoulder. A cross-sectional comparison of five questionnaires. *J Bone Joint Surg Am* 1996; **78**: 882-890

S- Editor Yang XC   **L- Editor** A   **E- Editor** Yang XC