

Validity and reliability of the SPORTS score for shoulder instability

DAVIDE BLONNA¹, ENRICO BELLATO², FRANCESCO CARANZANO¹, DAVIDE E. BONASIA²,
ANTONGIULIO MARMOTTI¹, ROBERTO ROSSI¹, FILIPPO CASTOLDI²

¹ Maurizioano "Umberto I" Hospital, Department of Orthopaedic and Traumatology, University of Turin, Medical School, Italy

² AO CTO, Maria Adelaide, "Città della salute e della scienza", Orthopaedic Department, University of Turin Medical School, Italy

Abstract

Purpose: athletes affected by shoulder instability cannot be judged solely according to the criteria used for non-athletes. In order to improve the assessment of shoulder instability surgery outcomes, the SPORTS score was tested in a cohort of athletes.

Methods: ninety-eight athletes at an average follow-up of 4.6 years (range 1-9.2) after open or arthroscopic surgery for recurrent anterior shoulder instability were included in this study. The patients were asked to complete the SPORTS score questionnaire twice, with an interval of 2-3 weeks between the two assessments. The Bland-Altman method and the intra-class correlation coefficient were used to measure reliability. Criterion validity was assessed by calculating the Spearman correlation coefficient between the SPORTS score and the Western Ontario Shoulder Instability Index (WOSI) score, the Rowe score, the Oxford Shoulder Instability Score (OSIS), and the Subjective Shoulder Value (SSV).

Results: the SPORTS score showed excellent test-retest reliability. The systematic error between the first and the second assessment was 0.3 points (95% upper limit of agreement = 2.3 points). The criterion validity was found to be strong for the SPORTS score, which correlated best with the SSV and the "sport, recreation, and work" component of the WOSI score. The SPORTS score had an acceptable floor effect (8%). The ceiling effect was 46%, which was better than the ceiling effects seen with the Rowe, OSIS and WOSI scores.

Corresponding Author:

Enrico Bellato, MD
AO CTO, Maria Adelaide, "Città della salute e della scienza",
Orthopaedic Department, University of Turin Medical School
Via Gianfranco Zuretti, 29, 10126 Turin, Italy
E-mail: bellatoenrico@gmail.com

Conclusions: this study suggests that the SPORTS score is a valid score in the assessment of athletes after surgery for shoulder instability and that it adds important information to the currently available scores.

Level of evidence: Level III, diagnostic study of non-consecutive patients.

Key Words: shoulder instability, reliability, score, sport, validity.

Introduction

A correct assessment of outcomes is mandatory after any surgical treatment. Self-assessment questionnaires or scores, shown to be valid and reliable, have been widely used in recent years (1-5). A main advantage of these scores is that they help clinicians and researchers to gather information without directly examining patients, thus saving patients and doctors both time and money.

However, self-assessment questionnaires, even though they are used worldwide, have limitations. Indeed, they tend to be developed and tested either in unspecific cohorts of patients or in healthy populations and they may not be supported by published clinical data (6-9). These factors can predispose these instruments to bias, since the validity of a score depends on the population in which it is applied. In short, a score found to be valid in the assessment of one specific group of patients cannot automatically be assumed to be valid when it is used in a different group of patients. As regards shoulder pathology, this has already been demonstrated. Conboy et al. (10) reported that the Constant-Murley score, which is commonly used to assess shoulder degenerative diseases, should not be used in cases of shoulder instability since it is not valid

in this subgroup of patients.

Athletes (another subgroup of patients) cannot be judged solely according to the criteria used for non-athletes. In athletes the main aim of treatment is usually to allow them to return to their sporting activities, while other areas of functioning, such as the ability to perform activities of daily living, are generally less affected (11-13).

The SPORTS score is a self-assessment score built to measure the ability of athletes to resume their pre-injury sporting activity (14). It has recently been proven to be valid and reliable in athletes after anterior or cruciate ligament reconstruction (15).

The SPORTS score may potentially also be extremely valuable in athletes affected by shoulder pathologies, however its validity has not been tested in this specific subgroup of patients.

The aim of this study was to test the reliability and validity of the SPORTS score in patients after surgery for shoulder instability.

Methods

This multicenter study, conducted in two university hospitals, concerned 110 patients previously submitted to surgery for isolated shoulder instability. The patients were contacted by phone and screened for fulfilment of the following three inclusion criteria: 1) consent to participate in the study, 2) a minimum follow-up of 12 months following surgery for shoulder instability, 3) sporting activity practiced before the onset of shoulder symptoms. Patients who had undergone concomitant surgical procedures to address SLAP lesions or cuff tears were excluded. Of the 110 patients screened, 98 met the inclusion criteria and remained in the study. No patient presented bilateral shoulder instability problems. Ninety-eight was deemed an appropriate number of patients given that the sample-size calculations indicate that 42 patients are required for a reliability study (16).

The study protocol adopted was the same in the two hospitals, but the examination of the patients and the collection of the data were performed by two different researchers.

All the patients had a scheduled appointment at one of

the two hospitals. Three weeks before the appointment, they were sent an envelope containing the following assessment instruments: the SPORTS score, the Oxford Instability Shoulder Score (OISS) (17), the Western Ontario Shoulder Instability Index (WOSI) (18, 19), and the Subjective Shoulder Value (SSV) (20). The patients were instructed to complete the questionnaires on the day, or the day after, they received the envelope. Two weeks before the appointment they were re-contacted by phone in order to ensure that they had received the questionnaires and had completed them.

During the subsequent outpatient clinic appointments, the patients were interviewed and asked to complete the SPORTS score questionnaire again for test-retest analysis (the questionnaires were completed by the patients alone, before being examined by the researchers). This time the patients were given a visually different version of the instrument (the five questions of the SPORTS score test were arranged in a different order): this was done to reduce potential recall bias. During the same appointment the patients were examined and the 1978 version of the Rowe score was applied (21).

Description of the scores

The SPORTS score is a patient-completed measure built around three separate concepts: 1) the ability to resume a sport at the pre-injury level of effort and training, 2) the ability to reach the same level of performance, and 3) the ability to achieve 1) and 2) without or despite pain. The instrument is scored on a 10-point scale with 0 points allocated to the patient who does not resume the same sport and 10 points to the patient who is able to perform, without pain, the same sport, at same level of effort and performance as before the onset of his/her impairment (14). Patients who, after surgery, abandon their sports for other reasons than their shoulder problems are asked to indicate the top SPORTS score recorded after surgery and before they abandoned their sport (Tab. 1).

The Oxford Shoulder Instability Score (OSIS) is a patient-completed measure comprising 12 questions on activities of daily living that are particularly relevant to patients exhibiting shoulder instability. The total score ranges from 0 (most severe symptoms) to 48 (least severe symptoms) (17).

The Western Ontario Shoulder Instability Index

Table 1. Subjective Patient Outcome for Return to Sports (SPORTS) score¹⁴.

Category	Definition	Relative value
Unlimited effort Unlimited performance No pain	Perform same sport at same level of effort and performance as before onset of impairment and with no pain	10
Unlimited effort Unlimited performance Some pain	Perform same sport at same level of effort and performance as before onset of impairment, but with pain	9
Unlimited effort Limited performance	Perform same sport at same level of effort, but reduced performance level compared to before onset of impairment	6
Limited effort Limited performance	Perform same sport, but at reduced levels of effort and performance compared to before onset of impairment	3
Disabled	Unable to return to same sport	0

(WOSI) consists of 21 items. The patient is asked to grade his/her shoulder function on a horizontal visual analog scale from 0 to 100 mm. The questions are divided into four sections: 1) physical symptoms and pain, 2) sport, recreation and work, 3) lifestyle and social function, and 4) emotional well-being. Each question carries between 0 and 100 points and the total score is thus a number between 0 and 2,100 points (where 0 represents no deficit and 2,100 the worst deficit) (18). Alternatively, the score can be presented as a percentage of a normal healthy shoulder (19).

The Subjective Shoulder Value (SSV) is a single-item self-completed measure in which patients are asked to grade their shoulder as a percentage of an entirely normal shoulder, which would score 100% (20).

The 1978 version of the Rowe score covers three domains: stability, function and motion. Motion is evaluated by objective evaluation. The other two items are scored according to subjective assessment. The total score ranges from 0 (worst) to 100 (best) (21).

Italian versions of the SPORTS score and WOSI were already available (15, 19). The SSV, Rowe and OSIS were independently translated into Italian by two bilingual orthopedic surgeons who are fluent in both English and Italian. In each case, the two translations were virtually identical and were thus deemed valid. A third translator (an American native-English speaker) translated the Italian versions of SSV, Rowe and OSIS back into English and no discrepancies were found with the original English versions.

Test-retest reliability

The SPORTS score questionnaire completed at home was compared with the same questionnaire completed in the outpatient clinic. The data were analyzed using a Bland-Altman analysis which defines the “limits of agreement” (22). This method is based on the mean and standard deviations of the differences between the ratings provided by the same subject.

A standard one-way analysis of variance intra-class correlation coefficient (ICC) was also performed (23). The ICC scores were as follows: <0.4, poor reliability; 0.4-0.75, fair to good reliability; >0.75, excellent reliability.

Validity

The criterion validity of the SPORTS score was assessed by calculating the Spearman correlation coefficient between the SPORTS score and 1) the Rowe score, 2) the OSIS score, 3) the WOSI score, and 4) the SSV. Correlation coefficients >0.5, 0.5-0.35, and <0.35 have been defined in the literature to be strong, moderate and weak respectively (24).

The floor and ceiling effects (content validity) were shown graphically by reporting the relative percentages of patients and the scores they obtained on the different questionnaires. For this purpose, the patients' SPORTS score item scores were divided into subgroups according to the upper limits of agreement (representing the minimal detectable differences) (15). Similarly, but using the data available in the literature, the Rowe score was divided into subgroups of 20

points (11), the OSIS into subgroups of six points (17) and the WOSI into subgroups of 10 points (19). For the SSV score we were not able to find information in the literature regarding the limits of agreement. We therefore performed a pilot study to assess the upper limits of agreement for the SSV in 25 patients affected by shoulder pathologies. The limit of agreement was found to be 11.6 points. The outcomes of the SSV were approximated and divided into subgroups of 10 points. Ceiling and floor effects of <30% were considered acceptable (25).

Results

Of the 110 patients screened, 102 patients met the inclusion criteria. Of these 102 patients, four patients did not attend the scheduled follow-up, leaving us with 98 patients available for analysis. The average age at surgery was 33 years (range 19-63) and the average follow-up duration was 4.6 years (range 1-9.2). Eighty-seven patients were male and 11 were female. Forty-eight patients had undergone an arthroscopic Bankart repair and 50 had undergone an open Latarjet procedure. At follow-up the average SPORTS score was 7 points (range 0-10). Five patients who had undergone an arthroscopic Bankart repair had a recurrent dislocation after surgery.

Reliability

The SPORTS score showed excellent test-retest reliability (ICC= 0.95; 95% confidence interval, 0.93-0.97). The systematic error between the first and the second assessment of the SPORTS score was 0.3 points with a 95% upper limit of agreement of 2.3 points (Fig. 1). The SPORTS score completed in the outpatient clinic was usually higher than the SPORTS score completed at home.

Of the 98 patients, 85 recorded the same score at the second assessment, while 13 patients recorded a different score. Of these 13 patients, three changed from score 3 to 6, six changed from score 9 to 10 or from 10 to 9, one patient changed from score 6 to 9, two from score 6 to 10, and one from 3 to 9. Considering the top scores (9 and 10), the upper limit of agreement

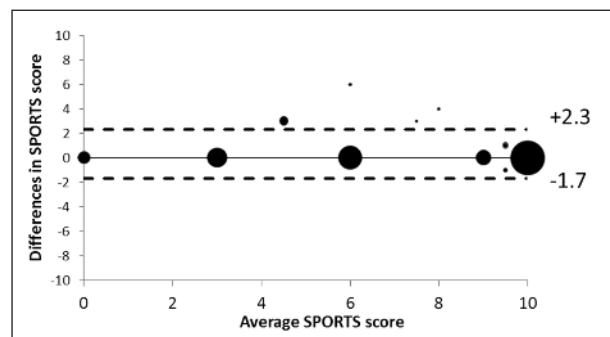


Fig. 1. Test-retest reliability of the SPORTS score.

of the SPORTS score was 0.7 points.

Criterion validity

A strong correlation was found between the SPORTS score and all the scores used as reference scores, with the exception of the “lifestyle and social function” component of the WOSI score (Tab. 2). The best correlation was found between the SPORTS score and the SSV score and between the SPORTS score and the “sport, recreation, and work” component of the WOSI score.

Content validity

The SPORTS score (Fig. 2) had an acceptable floor effect (8%).

The ceiling effect was high (46%), similar to what we observed for the Rowe, OSIS and WOSI scores. The SSV showed a better distribution of the outcome com-

Table 2. Criterion validity of the SPORTS score.

	Scores	Spearman's coefficient	p
SPORTS score	SSV*	0.6	<0.001
	WOSI B*	0.57	<0.001
	Oxford*	0.56	<0.001
	WOSI TOT*	0.55	<0.001
	ROWE	0.41	<0.001
	WOSI D*	0.52	0.001
	WOSI A*	0.51	<0.001
	WOSI C°	0.4	0.005

*: strong correlation

°: moderate correlation

WOSI A: physical symptoms and pain; WOSI B: sport, recreation, and work; WOSI C: lifestyle and social function; WOSI D: emotional well-being.

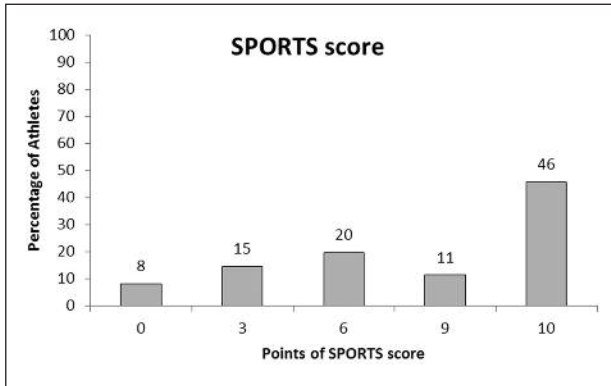


Fig. 2. SPORTS score results.

by knee instability it showed a limit of agreement of 1.8 points, good criterion validity and a better distribution of the outcomes compared with the reference scores (15). In the present study, the SPORTS score, applied in a population of athletes after shoulder instability surgery, was confirmed to have excellent reliability showing 0.3 points of systematic error with a 95% upper limit of agreement of 2.3 points. The fact that the upper limit of agreement was less than 3 points meant that the scores were not overlapping. An upper limit of agreement of less than 1 point for the higher scores confirmed that the scores 9 and 10 were recognized as separate outcomes.

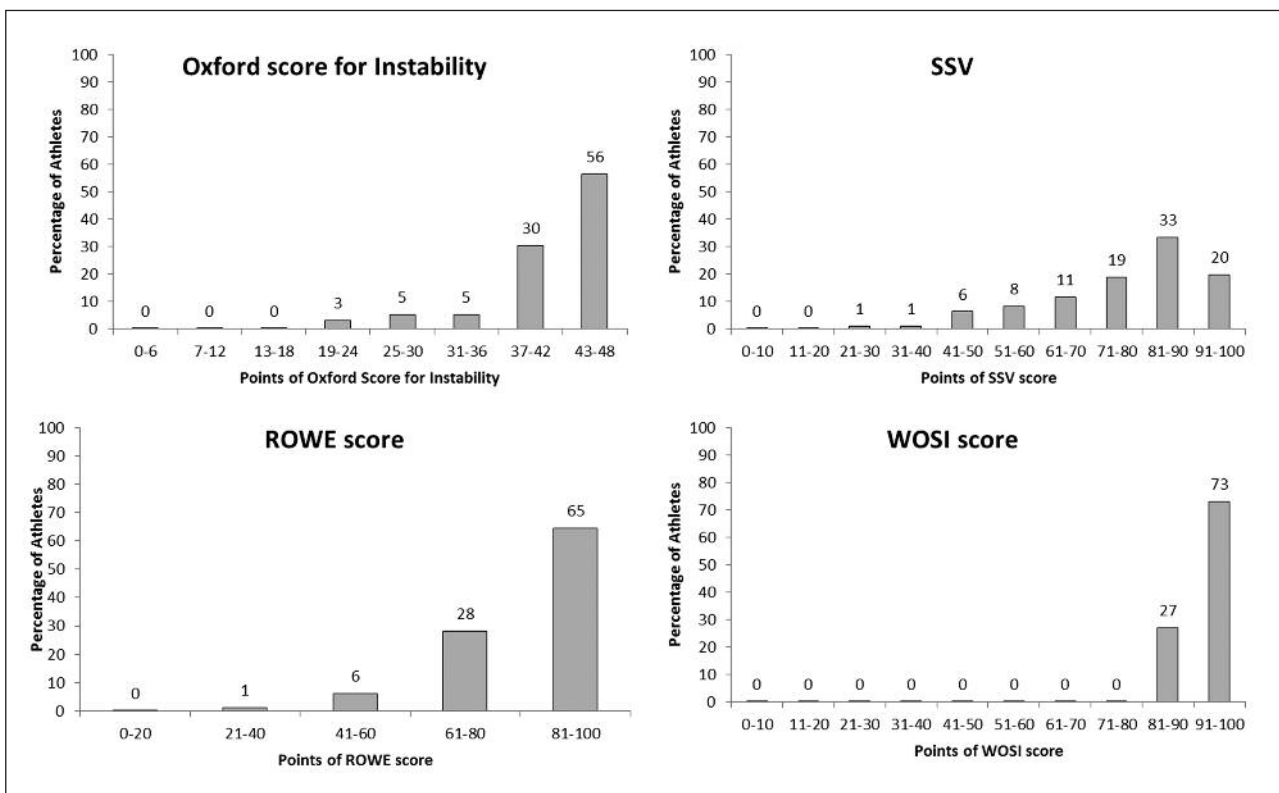


Fig. 3. Distribution of the outcomes of the OSIS, SSV, Rowe and WOSI scores.

pared with the other instruments (Fig. 3).

Discussion

This study was designed to test the validity of the SPORTS score in a cohort of patients after surgery for shoulder instability. In a recent study the SPORTS score was demonstrated to be a valid score during follow-up after ACL reconstruction. In athletes affected

As mentioned, the instrument showed good criterion validity. The SPORTS score, in fact, showed strong correlation with the most widely used and valid shoulder instability assessment scores. As expected one of the best correlations was found between the SPORTS score and the “sport, recreation, and work” component of the WOSI score. Interestingly, the best correlation was

found between the SPORTS score and the SSV, a finding that can be interpreted as a strong correlation between return to sport and the subjective feeling regarding the condition of the shoulder. The patients who felt that their shoulder was 100% compared with the contralateral normal shoulder had a better chance of returning to their full sporting activity without pain. As regards the content validity, the SPORTS score had a ceiling effect of 46%, meaning that almost half of the patients scored 10 points (i.e. performed, without pain, the same sport at the same level of effort and performance as before the onset of their shoulder impairment). This kind of unbalanced distribution of outcomes is generally considered a limitation in the use of a score. This is true, even though it could be interpreted clinically as a positive result, since it means that almost half of the patients achieved the best possible outcome. This distribution was, however, better than what we observed for the OSIS, Rowe and WOSI scores.

The disadvantage of a high ceiling effect becomes clearer when comparing treatments that provide good outcomes. In this scenario it can be difficult to detect statistically significant differences.

Grumet et al. (26) performed a systematic review comparing arthroscopic stabilization as first-time surgeries versus surgeries for recurrent shoulder instability. They concluded that there were no apparent differences in outcomes between the first-time instability and the recurrent instability. They also concluded that additional randomized controlled trials were needed to specifically compare the ability to return to a pre-injury level of activity (26).

Petrera et al. (27) performed a meta-analysis of open *versus* arthroscopic Bankart repair using suture anchors. They were not able to find significant differences between the two groups and concluded that larger and more homogeneous prospective studies were needed (27).

One way to overcome the limitation due to the high ceiling effect is to use the rate of recurrent instability as a primary outcome. Unfortunately this choice has two main drawbacks.

The first is that modern treatment techniques are usually associated with a low rate of recurrent instability. If two techniques are compared with an expected rate of recurrent instability of 5 and 10% respectively, at least

432 patients need to be included in each group in order to have an 80% chance of detecting significant differences ($\alpha = 0.05$, $\beta = 0.20$). This means that if the rate of recurrent instability is used as a main outcome, most of the studies available in the literature are underpowered. Second, the rate of recurrent instability is strongly related to the type of sport performed. In other words, a 20% recurrent instability rate in rugby players could be considered a clinically better result compared with a 10% recurrence rate in a sedentary population. Therefore the rate of recurrent instability should not be used alone as a main outcome indicator. This point was recently highlighted by Castagna et al. (28) who reported the outcomes of arthroscopic stabilization of the shoulder in adolescent athletes participating in overhead or contact sports. They found a 21% recurrence rate, but no evidence of a difference in clinical scores in the patients who had a recurrent dislocation (28).

A recent study by Stein et al. (12) focused on sport-specific shoulder impairments after arthroscopic Bankart repair. These authors concluded that sport-specific impairments are not detectable by established clinical score systems (12).

Furthermore, Netto et al. (29), comparing arthroscopic and open techniques for the treatment of Bankart lesions, found no differences when using established instability assessment scores (Rowe and UCLA scores). Remarkably, they did detect score differences when using the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire (an instrument not specifically designed for shoulder instability). Similarly we used the SSV, a non-specific score initially designed for degenerative conditions of the shoulder (20) and, interestingly, found it to show the best distribution of outcomes with no ceiling effect.

In this scenario, it seems reasonable to conclude that there is a need for new tools that may allow a more complete assessment of outcomes after surgery for shoulder instability. This study has shown that the SPORTS score is a valid score in the assessment of athletes after surgery for shoulder instability. However, it cannot be used in isolation. Since it reports the outcome as a function of the return to sport, an accurate evaluation of the type and level of sport practiced before the occurrence of the shoulder problem is clearly mandatory.

The main limitations of this study are that the responsiveness and Cronbach alpha of the SPORTS score were not measured. Responsiveness will be assessed as part of our ongoing work while the Cronbach alpha cannot be used for one-item questionnaires. Another limitation is that the Italian versions of some of the scores that we used are still to be validated. However, great care was taken to translate the instruments in a reliable way. Furthermore, none of them needed cross-cultural adaptation.

In conclusion, this study suggests that the SPORTS score is a valid score in the assessment of athletes after surgery for shoulder instability. It is a useful tool that adds important information to the currently available scores.

References

1. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. *J Clin Epidemiol*. 1998;51:1025-1036.
2. Rouleau DM, Faber K, MacDermid JC. Systematic review of patient-administered shoulder functional scores on instability. *J Shoulder Elbow Surg*. 2010;19:1121-1128.
3. Mohtadi N. Development and validation of the quality of life outcome measure (questionnaire) for chronic anterior cruciate ligament deficiency. *Am J Sports Med*. 1998;26:350-359.
4. Thorborg K, Roos EM, Bartels EM, et al. Validity, reliability and responsiveness of patient-reported outcome questionnaires when assessing hip and groin disability: a systematic review. *Br J Sports Med*. 2010;44:1186-1196.
5. Ware JE, Jr, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project. *J Clin Epidemiol*. 1998;51: 903-912.
6. Richards RR, An KN, Bigliani LU, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg*. 1994;3:347-352.
7. Roach KE, Budiman-Mak E, Songsirdej N, et al. Development of a shoulder pain and disability index. *Arthritis Care Res*. 1991;4:143-149.
8. Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res*. 1987;214:160-164.
9. Lippitt SB, Harryman DT II, Matsen FA III. A practical tool for evaluating function: The Simple Shoulder Test. In: Matsen FA, Fu FH, Hawkins RJ (Eds.) *The Shoulder: A balance of Mobility and Stability*. Rosemont, American Academy of Orthopaedic Surgeons. 1992; 501-518.
10. Conboy VB, Morris RW, Kiss J, et al. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br*. 1996;78:229-232.
11. Skare Ø, Schröder CP, Mowinckel P, et al. Reliability, agreement and validity of the 1988 version of the Rowe Score. *J Shoulder Elbow Surg*. 2011; 20:1041-1049.
12. Stein T, Linke RD, Buckup J, et al. Shoulder sport-specific impairments after arthroscopic Bankart repair: a prospective longitudinal assessment. *Am J Sports Med*. 2011;39: 2404-2414.
13. Voigt C, Schulz AP, Lill H. Arthroscopic treatment of multidirectional glenohumeral instability in young overhead athletes. *Open Orthop J*. 2009; 3:107-114.
14. Blonna D, Lee GC, O'Driscoll SW. Arthroscopic restoration of terminal elbow extension in high-level athletes. *Am J Sports Med*. 2010; 38:2509-2515.
15. Blonna D, Castoldi F, Delicio D, et al. Validity and reliability of the SPORTS score. *Knee Surg Sports Traumatol Arthrosc*. 2012;20:356-360.
16. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med*. 1987;6:441-448.
17. Dawson J, Fitzpatrick R, Carr A. The assessment of shoulder instability. The development and validation of a questionnaire. *J Bone Joint Surg Br*. 1999;81:420-426.
18. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI). *Am J Sports Med*. 1998;26:764-772.
19. Cacchio A, Paoloni M, Griffin SH, et al. Cross-cultural adaptation and measurement properties of an Italian version of the Western Ontario Shoulder Instability Index (WOSI). *J Orthop Sports Phys Ther*. 2012;42:559-567.
20. Gilbert MK, Gerber C. Comparison of the subjective shoulder value and the Constant score. *J Shoulder Elbow Surg*. 2007;16:717-721.
21. Rowe CR, Patel D, Southmayd WW. The Bankart procedure: a long-term end-result study. *J Bone Joint Surg Am*. 1978;60:1-16.
22. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
23. Haas M. Statistical methodology for reliability studies. *J Manipulative Physiol Ther*. 1991;14:119-132.
24. Fayers PM, Machin, D. *Quality of Life: The assessment, analysis and interpretation of patient-reported outcomes*, 2nd edition. John Wiley & Sons, Ltd, Chichester. 2007.
25. Kane R. *Understanding health care outcomes research*. Gaithersburg, MD, Aspen. 1997.
26. Grumet RC, Bach BR, Jr, Provencher MT. Arthroscopic stabilization for first-time versus recurrent shoulder instability. *Arthroscopy*. 2010;26:239-248.
27. Petrera M, Patella V, Patella S, Theodoropoulos J. A meta-analysis of open versus arthroscopic Bankart repair using suture anchors. *Knee Surg Sports Traumatol Arthrosc*. 2010;18:1742-1747.
28. Castagna A, Delle Rose G, Borroni M, et al. Arthroscopic stabilization of the shoulder in adolescent athletes participating in overhead or contact sports. *Arthroscopy*. 2012;28:309-315.
29. Archetti Netto N, Tamaoki MJ, Lenza M, et al. Treatment of Bankart lesions in traumatic anterior instability of the shoulder: a randomized controlled trial comparing arthroscopy and open techniques. *Arthroscopy*. 2012;28: 900-908.