Normalization of the Constant score

Leonid I. Katolik, MD,[a] Anthony A. Romeo, MD,[b] Brian J. Cole, MD,[b] Nikhil N. Verma, MD,[b] Jennifer K. Hayden, MSN,[b] and Bernard R. Bach, MD,[b] *Seattle, WA, and Chicago, IL*

The strength of the normal shoulder may differ by gender and deteriorate with age. Thus, the Constant score may also decrease in absolute value while still reflecting a normal score. To account for age- and gender-related differences, normal results for this scale must be determined across a population of patients without shoulder disease. Patients presenting for evaluation of nonshoulder conditions participated. A subjective questionnaire was completed. Range of motion and strength were measured. This analysis includes the data of 441 patients. The mean Constant score for men was significantly greater than that for women in each age group ($P < .05$). Significant age-related differences were noted in each group ($P < .05$). Normative values for the Constant score based on age and gender were determined. The adjusted score represents the gender- and age-matched function of the shoulder and is useful in the evaluation of shoulder outcomes. (J Shoulder Elbow Surg 2005;14:279-285.)

The Constant score is a widely used shoulder-specific scoring system. In 1992 the European Shoulder and Elbow Society mandated the use of the Constant score in all peer-reviewed papers, making it the most widely used shoulder evaluation instrument in Europe.[20] First described by Constant and Murley[8] in 1986, it proposes a scoring system directed exclusively toward a numeric description of the quality of function of the shoulder. Instead of relying on tests of specific functional movements (eg, brushing one's hair, removing an object from a shelf, and throwing a ball), it uses subjective and objective measures to determine whether a certain functional movement is possible (eg, forward elevation, external rotation, and internal rotation of the shoulder). It is separated conceptually from the diagnosis of the shoulder, thereby making it applicable regardless of diagnosis.

As an outcomes tool, the Constant score includes an analysis of pain, shoulder motion, strength, and function. From a perfect score of 100, it reserves 35 points for patient-reported subjective assessment, including the presence of pain and the ability to perform basic activities of daily living, and 65 points for objective measurement. For the latter, 40 points are allocated to range of motion and 25 points are allocated to strength. The relative weight of subjective and objective findings is based on statistical analysis correlating subjectively perceived disability with the results of a combined objective and subjective evaluation.[7,23]

Because the strength of the normal shoulder may differ by gender and deteriorate with age, the Constant score will also decrease, although the score may still be normal for the patient's age and gender. To account for age- and gender-related differences, normal results for this scale must be determined across a large population of patients without shoulder disease. Unfortunately, to date, there are no peer-reviewed, large population studies to provide normalization data for the Constant score despite the fact that numerous publications suggest they are using an adjusted or normalized score.[2,10,11,13,15,17-19,22,27]

A careful distinction in terminology must be made between validation and normalization. The former seeks not only to determine test-retest reliability of an outcomes measure but also to assess the correlation between a variety of outcomes instruments. The latter seeks to provide a standard, derived from a large population of individuals without joint-specific complaints, against which measurements are compared and adjusted. The purpose of this investigation is to normalize the results of the Constant score based on age and gender in patients without shoulder symptoms.

MATERIALS AND METHODS

All patients underwent an informed consent process approved by the Institutional Review Board and Human Subjects Committee of Rush University Medical Center (Chicago, IL). Four hundred eighty patients presenting to a sports medicine clinic for evaluation of nonshoulder conditions participated in clinical testing that would allow the

From the Department of Orthopaedic Surgery, Section of Hand and Microvascular Surgery,[a] University of Washington, Seattle, and Department of Orthopaedic Surgery, Section of Shoulder and Elbow Surgery,[b] Rush University Medical Center, Chicago.

Reprint requests: Brian J. Cole, MD, 1725 W Harrison, Suite 1063, Chicago, IL 60612 (E-mail: *bcole@rushortho.com*).

1058-2746/2005/$30.00

doi:10.1016/j.jse.2004.10.009

Table I Subjective assessment of shoulder (35 total points possible)

Function	Points
Ability to work	0-4
Ability to engage in recreational activities	0-4
Ability to sleep	0-2
Ability to work at specific level	
Waist	2
Chest	4
Neck	6
Head	8
Above head	10
Pain	0-15

determination of the Constant score for the dominant shoulder. All participants were carefully screened by the examiner to determine the existence of current or prior shoulder injury, prior nonoperative or operative interventions involving the shoulder or elbow, and prior interventions for cervical or thoracic pathology (eg, axillary lymph node dissection for carcinoma of the breast, prior sternotomy, or rib fracture). Affirmative responses resulted in exclusion from study participation. Scores were determined for the dominant shoulder in all groups.

Patients completed a 1-page questionnaire that assessed subjective pain and the ability to perform activities of daily living. This questionnaire incorporated the subjective items from Constant and Murley's functional assessment.[8] Pain was clearly defined as the "worst pain that the patient experiences during activities of daily living"; 15 points denoted no pain, 10 points denoted mild pain, 5 points denoted moderate pain, and 0 points denoted severe pain interfering with regular activities. The ability to execute and perform activities of daily living was assigned 20 points, with 4 points allocated to work, 4 points to recreational activities, and 2 points to sleep. A patient who can perform all work or recreational activities without restriction is assigned 4 points. If a patient has given up 50% of activities, 2 points are assigned. A patient whose sleep is uninterrupted receives 2 points. If sleep is grossly disturbed, no points are given. For the ability to work at a specific level, points were assigned as detailed in Table I.

In addition to answering the subjective questions, range of motion and strength were objectively assessed. All measurements were performed as specified by Constant and Murley.[8] Range of motion was measured with a goniometer between the upper arm and the upper part of the thorax (Figure 1). Points were assigned as detailed in Table II. As specified by Constant and Murley, measured flexion and abduction reflect the active range without pain. External rotation is indirectly measured as a functional assessment of external rotation. This allows for the exclusion of a theoretically present range of motion that does not translate functionally. Similarly, internal rotation was tested in combination with shoulder extension and adduction, allowing for assessment of a functional rather than a theoretic movement.

Strength was assessed by use of the Isobex Dynamometer (Cursor AG, Bern, Switzerland). This is a microprocessor-driven device whose measurement is triggered by a minimum force of 1 kg. It disregards the first second of force application, where a rapid linear increase in force is noted typically, and averages 10 readings per second for the following 3 seconds of force application to produce a strength reading. All calculations were then converted to pounds. All measurements were made in the scapular plane of abduction. A maximum of 25 points is awarded for the ability to hold 25 lb or more at 90° in the plane of the scapula. The number of points correlated to the number of pounds held by the patient. Strength and motion testing was performed in all individuals by use of the dominant extremity, as numerous studies have confirmed that strength in forward flexion does not vary significantly from side to side in both a sedentary and athletic population.[5,9,25,26]

Data were recorded on forms created and scanned by use of Teleform v.6 software (Cardiff Software, Inc, San Diego, CA). Data were analyzed by use of SPSS v.10 software (SPSS Inc, Chicago, IL). Independent sample *t* test between genders was performed for each age group, and $P < .05$ was considered significant. One-way analysis of variance between age groups for each gender was determined. Any significance between groups was assessed with the Tukey HSD (honestly significantly different) test, and $P < .05$ was considered significant.

RESULTS

Four hundred eighty patients were evaluated. Five were aged less than 18 years, and thirty-nine were not included in the analysis because of data that were either missing or not able to be read by the scanner. Our results are based on complete data from 441 patients (Table III). The mean Constant score for men was significantly greater than that for women in each age group ($P < .05$) (Figure 2). For men, patients aged older than 70 years had a significantly lower Constant score than those aged 18 to 29 years, 30 to 39 years, 40 to 49 years, and 50 to 59 years ($P < .05$) but not in those in the age group from 60 to 69 years. For women, patients aged less than 50 years had a significantly greater Constant score than those aged older than 50 years ($P < .05$).

For men, measured strength degraded with age. This difference was most notable in those aged older than 70 years. In this group, strength scores were significantly lower ($P < .05$) than those in all patients aged younger than 60 years (Figure 3). Furthermore, there was a significant difference in strength scores when we compared patients aged older than 50 years with patients aged younger than 50 years ($P < .05$). For men, there was no significant difference among age groups in subjective score or range-of-motion scores, and these scores tended not to degrade with age.

For women, raw strength scores were significantly lower than for those of men in the same age group ($P < .05$). Strength was significantly lower in women aged older than 50 years compared with those aged younger than 50 years. There was no uniformly significant difference in subjective scores, and these

Figure 1 A, A goniometer is used to measure shoulder range of motion. Three measurements are taken and the average recorded. **B**, Isobex dynamometer is used to assess strength in forward flexion with the arm in the scapular plane.

Table II Objective shoulder assessment (65 total points possible)

Activity	Points
Flexion and abduction (scored separately)	
>150°	10
121° = 150°	8
91° = 120°	6
61° = 90°	4
31° = 60°	2
Combined active external rotation	
Hand behind head, elbow forward	2
Hand behind head, elbow back	2
Hand on top of head, elbow forward	2
Hand on top of head, elbow back	2
Full elevation from top of head	2
Combined active internal rotation of hand	
Interscapular region	10
Inferior tip of scapula	8
Twelfth rib	6
Lumbosacral junction	4
Buttock	2
Lateral thigh	0
Strength	1/lb

Table III Distribution of patients

	Men	Women
Age (y)		
18-29	51	45
30-39	53	40
40-49	44	48
50-59	43	51
60-69	23	16
≥70	13	14
Total	227	214

scores did not degrade with age. When range-of-motion scores were evaluated, a statistically significant difference was noted in the comparison of subjects aged less than 30 years with subjects aged between 50 and 59 years.

DISCUSSION

A shoulder scoring system serves as a simplified means for the evaluation of a variety of shoulder conditions. It is intended to represent the outcome of intervention without sensitivity to observer bias or bias based on the parameters measured. In addition, it should allow the effective communication of the results of medical and surgical intervention in a patient group. It would, therefore, allow comparison of dif-

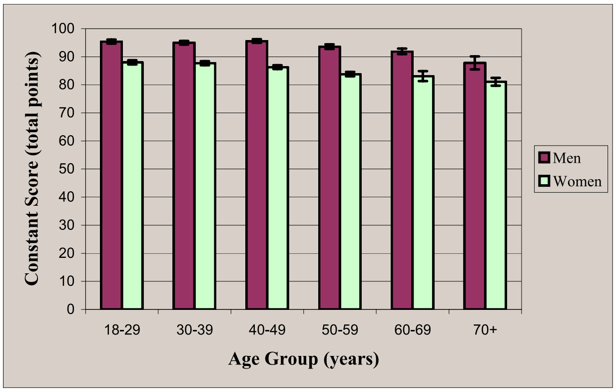

Figure 2 Constant score by age group.

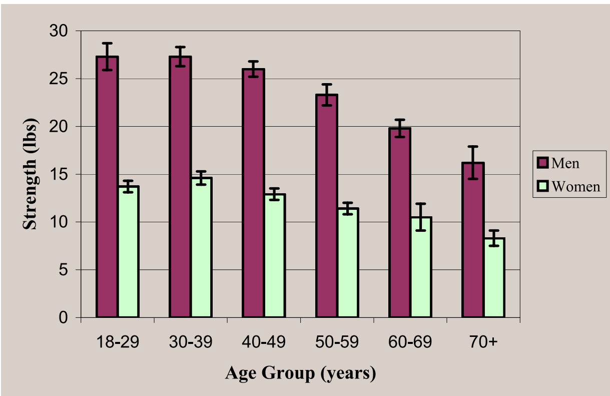

Figure 3 Comparison of strength by gender.

ferent therapeutic modalities by defining functional improvement or deterioration based on disease and treatment. It would also help to standardize comparisons between patient groups in multicenter studies. The ideal scoring system would be simple and readily applicable to clinical practice. It should be easy to administer, its methods should be well defined, and it should provide a useful description of the function of the shoulder, regardless of disease process or intervention. In addition, the system should be weighted toward functional outcome with the patients' perspective prioritized. For example, traditional physician-based parameters such as motion and strength do not always adequately represent patient-perceived value of an intervention. Finally, a shoulder scoring system should be reproducible among practitioners. Ultimately, in a society increasingly pressed to lay a monetary value on treatment, value may be calculated as the benefit divided by cost, where benefit is expressed as a change in score.

The Constant score enjoys a high degree of reproducibility among observers, with an intraobserver error of 3%.[8] It is the first shoulder scoring system to allow the effects of aging on shoulder function to be studied. As such, it is vital to be able to normalize calculated scores based on age and gender. Adjusted scores would allow the determination of whether a specific procedure is most beneficial for a certain age group or particular gender.

To our knowledge, this is the first normalization of the Constant score in normal subjects in the North American population. Our study evaluated a representative population of a large metropolitan area to determine normative values based on age and gender. The calculated score for any patient may then be adjusted based on a gender- and age-specific normative value to yield an adjusted score. The adjusted score would, therefore, accurately represent the function of the shoulder in comparison to patients of a similar age and gender. Comparison of shoulder outcomes would, therefore, enjoy a greater degree of validity.

Strength remains a major determinant of the total score. However, the measurement method of the strength component has not been standardized.[2] The method originally described by Constant and Murley[8] used a spring balance held at arms length. The maximum force that the patient could resist against the downward pull by the examiner was then measured. Constant did not specify the plane of elevation of the shoulder for testing or the duration of each measurement. Furthermore, Conboy et al[6] found that accurate measurement of power by use of a spring scale was difficult to determine consistently between and among observers. They also questioned whether strength assessment in a single arc of motion adequately represented a patient's full functional potential.

These inconsistencies were addressed by Gerber and Arneberg[12] in 1992, who questioned the validity of using a spring balance as a measuring device. As a result, the Isobex Dynamometer was developed, and a normal range of values for elevation in the scapular plane was defined. Several scientific studies have assessed the reliability of the Isobex Dynamometer for test-retest and intraobserver reliability.[3,21] Furthermore, the modified spring balance, as used by Constant, was found to provide similar values to those of the Isobex.[2] Strength was tested with the humerus in the plane of the scapula for maximum biomechanical advantage. This position maintains maximal glenohumeral conformity, an optimum length-tension relationship for the humeral abductors, and a relaxed inferior capsule.[14] Given the strength advantage that this position of testing affords, our study tests the strength of its subjects in the same position, and perhaps all future studies that use the Constant score as an outcomes tool should use this position.

The strength values measured by Gerber and Arneberg[12] are compared with our strength scores in

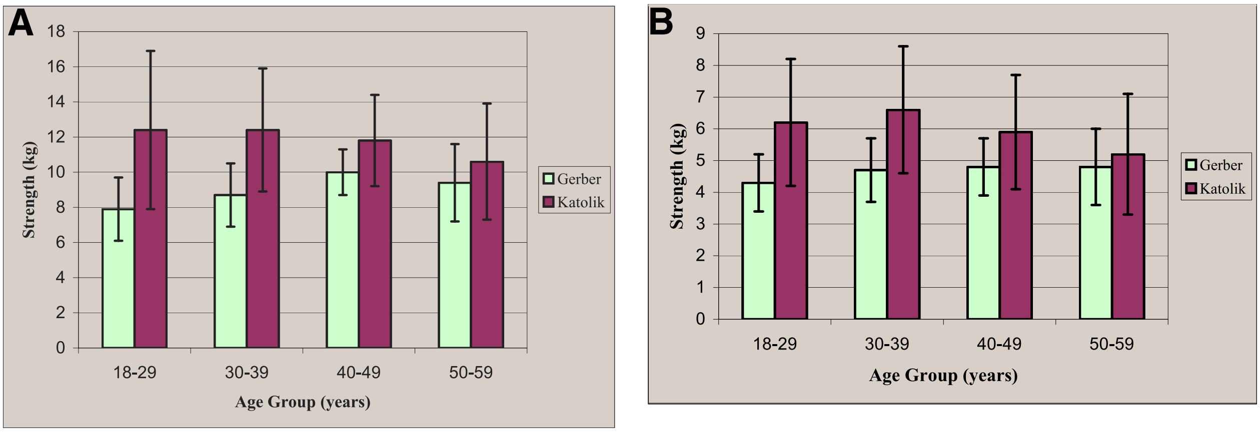

Figure 4 Comparison of raw strength scores in men (**A**) and women (**B**) in the study of Gerber and Arneberg[12] and the current study.

Figure 4. Two observations may be made. The first is that for both genders, our population showed a deterioration of strength with age as described above. Our findings that age and sex affect strength measurement in the shoulder have been confirmed by other reports on upper extremity strength in a variety of other studies.[1,4,16,24] This contrasts to strength values measured by Gerber and Arneberg, who showed an increase with each decade. Second, our strength values were greater for each age group. This difference may be accounted for by the small population used in the study of Gerber and Arneberg. It may also reflect the stringent selection criteria of our study, in that patients were excluded if they acknowledged any past or present difficulties with either shoulder or any history that may indirectly impinge on scapulothoracic or glenohumeral function (eg, axillary node dissection for carcinoma of the breast or rib fracture). This absolute difference in raw strength between our study and the study of Gerber and Arneberg decreased as age decreased.

Furthermore, normal Constant scores as first described by Constant in 1986 are compared with the normal values determined in our study (Table IV and Figure 5). The differences in total scores are most marked for older age groups. For both genders, our scores are initially lower than those originally described by Constant. Our scores, however, do not degrade at the same rate as those described by Constant. These differences may reflect the measurement system used by Constant or a lack of selectivity of normal patients.

Although many studies in the shoulder literature imply the use of normal data to provide an adjusted Constant score, the quality and value of these data are questionable without peer review of the methods and proposed reference values. The results of our study may be used to define the adjusted or normative values when using the Constant score to define shoulder outcomes. We propose the use of the following formula to calculate a normalized Constant score: Normalized score = (Raw score/Normal score) × 100. The normal score for the denominator of the equation can be determined from the values in Table IV. As an illustration of the use of this formula and the age- and gender-matched normal scores from Table IV, a 63-year-old woman with a raw Constant score of 65 after shoulder hemiarthroplasty would have a normalized score of 78. If one were to use the values originally calculated by Constant, the score would be 93, perhaps reflecting marked numeric success in the reality of a fair functional outcome.

Table IV Normal Constant scores

	Men		Women	
Age (y)	Current study	Constant[7]	Current study	Constant[7]
18–29	95	98	88	97
30-39	95	93	87	90
40-49	96	92	86	80
50-59	94	90	84	73
60-69	92	83	83	70
≥70	88	75	81	69

These results indicate that, if one is to convey the results of treatment accurately using the adjusted Constant score, one needs to define the normal results of the Constant score in a population that is similar to the population of patients studied. Using historically proposed normative data as originally described by Constant in 1986 for our patient population would have significantly increased our normalized or adjusted Constant score, suggesting better outcomes. Furthermore, studies that have used Constant scores normal-

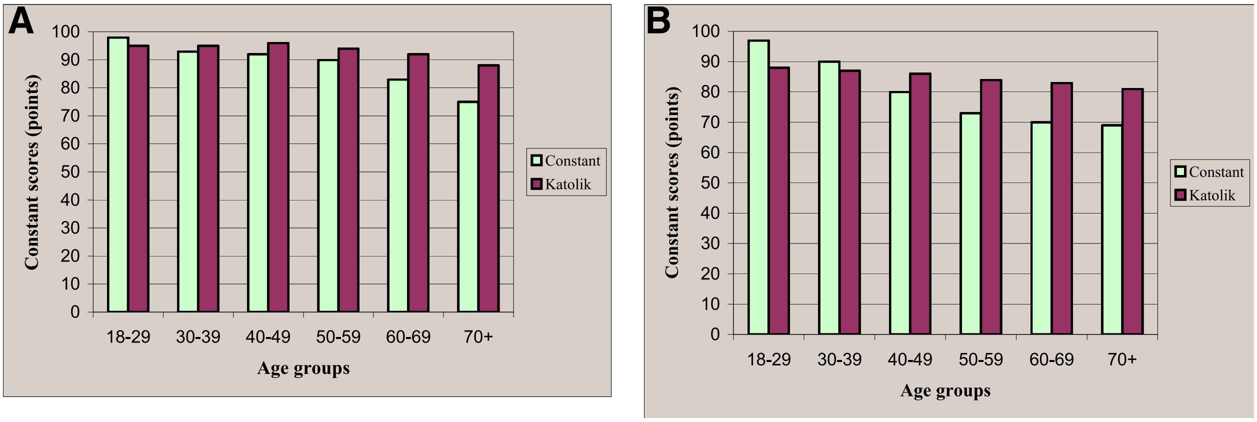

Figure 5 Comparison of normal Constant scores in men (**A**) and women (**B**).

ized on the basis of Constant's original work may have been affected by poorly matched normal data.

The potential for selection bias is a limitation of this study and should be addressed. First, the fact that all patients in this study were accrued from a sports medicine clinic may imply that physical fitness enthusiasts comprised the majority of the study populace drawn from such a setting. The vagaries of modern orthopaedic practice, however, are such that the very term sports medicine is perhaps more reflective of the training received by the attending physicians than the overall fitness or athletic ability of the patient population. Indeed, roughly 40% of the clinic population is referred through the workers' compensation network. Furthermore, limiting this study to patients who did not have any history of shoulder disease or injury may have led to inherently stronger subjects, particularly in the older age groups. Finally, subjects were not evenly divided among all age groups. Nevertheless, we believe that the subjects in this study were representative of the general population and that the degradation of values with age, as well as the differences between sexes, is representative of the population at large.

Our study provides normal data for a large metropolitan population and a means for calculating an adjusted Constant score from a raw score. The adjusted score accurately represents the function of the shoulder in comparison to patients of a similar age and gender and is useful in the evaluation of shoulder outcomes. The reporting of outcomes by use of an adjusted score derived from our age- and gender-stratified data would thus allow a comparison of results from varied sites, if the methods of measurement and scoring are closely adhered to. Given the observed difference between our data and those originally described by Constant, it may be argued that each surgeon reporting the outcomes of interventions for a variety of shoulder pathologies should be required to compile personal data for the normalization of scores. Though ultimately perhaps more accurate, the gathering of such data may prove to be tremendously unwieldy, and a comparison of outcomes between groups of different surgeons would be difficult to perform. The utilization of normal data from a large metropolitan population without shoulder symptoms to generate adjusted age- and gender-matched Constant scores should serve as an excellent basis for the reporting and comparison of outcomes data, to facilitate communication between investigators and to permit and encourage multicenter studies.

We wish to acknowledge Shane Nho, MD, for his outstanding dedication to this project, most specifically the data collection. His efforts in patient recruitment are greatly appreciated, and without his contribution, we might still be consenting patients.

REFERENCES

1. Backman E, Johansson V, Hager B, Sjoblom P, Henriksson KG. Isometric muscle strength and muscular endurance in normal persons aged between 17 and 70 years. Scand J Rehabil Med 1995;27:109-17.
2. Bankes M, Crossman J, Emery J. A standard method of shoulder strength measurement for the Constant score with a spring balance. J Shoulder Elbow Surg 1998;7:116-21.
3. Bankes MJ, Crossman JE, Emery RJ. A standard method of shoulder strength measurement for the Constant score with a spring balance. J Shoulder Elbow Surg 1998;7:116-21.
4. Cahalan TD, Johnson ME, Chao EY. Shoulder strength analysis using the Cybex II isokinetic dynamometer. Clin Orthop 1991; 271:249-57.
5. Codine P, Bernard PL, Pocholle M, Benaim C, Brun V. Influence of sports discipline on shoulder rotator cuff balance. Med Sci Sports Exerc 1997;29:1400-5.
6. Conboy VB, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant-Murley shoulder assessment. J Bone Joint Surg Br 1996; 78:229-32.
7. Constant C. Age related recovery of shoulder function after injury (Thesis). Cork: University College, Ireland; 1986.
8. Constant C, Murley A. A clinical method of functional assessment of the shoulder. Clin Orthop 1987;214:160-4.

9. Ellenbecker T, Roetert EP. Age specific isokinetic glenohumeral internal and external rotation strength in elite junior tennis players. J Sci Med Sport 2003;6:63-70.
10. Fink B, Sallen V, Guderian H, Tillmann K, Ruther W. Resection interposition arthroplasty of the shoulder affected by inflammatory arthritis. J Shoulder Elbow Surg 2001;10:365-71.
11. Gartsman GM, Roddey TS, Hammerman SM. Arthroscopic treatment of multidirectional glenohumeral instability: 2- to 5-year follow-up. Arthroscopy 2001;17:236-43.
12. Gerber C, Arneberg O. Measurement of abductor strength with an electronical device (Isobex). J Shoulder Elbow Surg 1992; 2(suppl 1):S6(Abstract).
13. Gerber C, Fuchs B, Hodler J. The results of repair of massive tears of the rotator cuff. J Bone Joint Surg Am 2000;82:505-15.
14. Greenfield B, Donatelli R, Wooden M, Wilkes J. Isokinetic evaluation of shoulder rotational strength between the plane of the scapula and the frontal plane. Am J Sports Med 1990;18: 124-8.
15. Hossain S, Roy N, Ayeko C, Elsworth CF, Jacobs LG. Shoulder and elbow function following Marchetti-Vicenzi humeral nail fixation. Acta Orthop Belg 2003;69:137-41.
16. Hughes RE, Johnson ME, O'Driscoll SW, An KN. Age-related changes in normal isometric shoulder strength. Am J Sports Med 1999;27:651-7.
17. Jacobs R, Debeer P, De Smet L. Treatment of rotator cuff arthropathy with a reversed Delta shoulder prosthesis. Acta Orthop Belg 2001;67:344-7.
18. Karlsson J, Magnusson L, Ejerhed L, et al. Comparison of open and arthroscopic stabilization for recurrent shoulder dislocation in patients with a Bankart lesion. Am J Sports Med 2001;29:538-42.
19. Kay SP, Dragoo JL, Lee R. Long-term results of arthroscopic resection of the distal clavicle with concomitant subacromial decompression. Arthroscopy 2003;19:805-9.
20. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. Arthroscopy 2003;19: 1109-20.
21. Leggin BG, Neuman RM, Iannotti JP, Williams GR, Thompson EC. Intrarater and interrater reliability of three isometric dynamometers in assessing shoulder strength. J Shoulder Elbow Surg 1996; 5:18-24.
22. Levy O, Wilson M, Williams H, et al. Thermal capsular shrinkage for shoulder instability. Mid-term longitudinal outcome study. J Bone Joint Surg Br 2001;83:640-5.
23. Marechal E. Ruptures degeneratives de la coiffes de rotateurs de l'epaule: evaluation functionnelle: resultats du traitement chirurgical. Faculte de Medecine. Lyon, France: Universite Claude Bernard; 1990 (Thesis).
24. Murray MP, Gore DR, Gardner GM, Mollinger LA. Shoulder motion and muscle strength of normal men and women in two age groups. Clin Orthop 1985;192:268-73.
25. Newsham KR, Keith CS, Saunders JE, Goffinett AS. Isokinetic profile of baseball pitchers' internal/external rotation 180, 300, 450 degrees.s-1. Med Sci Sports Exerc 1998;30: 1489-95.
26. Shklar A, Dvir Z. Isokinetic strength relationships in shoulder muscles. Clin Biomech (Bristol, Avon) 1995;10:369-73.
27. Yanmis I, Tunay S, Komurcu M, et al. Outcomes of acute arthroscopic repair and conservative treatment following first traumatic dislocation of the shoulder joint in young patients. Ann Acad Med Singapore 2003;32:824-7.