The necessity of strength evaluation in assessment of clinical outcome after shoulder surgery: follow-up data from patients with complex proximal humerus fractures treated by locking plate fixation

Y. Wu, P. Shang, L. Che, T. Ye, L. Wang, S. Qiu

From the Department of Orthopedics Shanghai Institute of Traumatology and Orthopedics Ruijin, Shanghai, China

In order to minimize the bias of Constant score we modified the allocation of strength subscore. One hundred and two patients with 3- and 4-part proximal humerus fractures were treated using locking plate fixation and followed up for > 1 year. The clinical outcomes were assessed by DASH score abbreviated Constant score (AbbCS strength item excluded) modified Constant score (ModCS with 12-pound strength) and original Constant score (CS with 25-pound strength). The satisfaction rate was determined for each scoring instrument. Compared to CS the satisfaction rate was significantly higher in DASH score AbbCS and ModCS (all \( p < 0.001 \)) but the latter 3 groups did not show significant difference. ROC analysis showed that a > 7-pound shoulder strength was present in patients with satisfied outcome assessed by DASH score. In conclusion strength evaluation is necessary for the assessment of shoulder function but the over-allocated strength should be modified in Constant score.

Keywords: modified Constant Score; DASH score, proximal humerus fractures; locking plate fixation; satisfaction rate.

INTRODUCTION

There are a number of outcome assessment instruments for patients with shoulder problems, including the American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form (ASES), the Disabilities of the Arm, Shoulder, and Hand (DASH) outcome measure, the University of California Los Angeles (UCLA) Shoulder Rating Scale, Western Ontario Rotator Cuff index (WORC) and the Constant Score, et al (14,21,24). The combined use of Disabilities of the Arm, Shoulder and Hand (DASH) and Constant scoring instruments

No benefits or funds were received in support of this study. The authors report no conflict of interests.
has been frequently applied to the assessment of clinical outcome following shoulder surgeries (12, 15, 22). DASH score is based on patient report but Constant score is mainly derived from physician examination (65 of 100 points) (1, 7, 24). There is a large discordance in outcome assessment between patient-based DASH score and clinician-based Constant score, because the Constant score is susceptible to bias and may not represent the view of the patient (5, 12).

Over-allocation of strength subscore has been considered to increase the risk of bias (16). In the Constant score, shoulder strength has been arbitrary allocated with 25 points (1 point = 1 pound), which accounts for 25% of the total score (7). In fact, it is impossible for many people, especially women and elderly, to lift 25 pounds with a straight arm, so they may lose many points in the Constant assessment (13, 18). However, these people are often satisfactory to their shoulder function. To the best of our knowledge, few people with normal joint movement complain about their shoulder function because of decreased strength. Thus, the inclusion of 25 pound strength would cause a falsely low Constant score to compromise the outcome (16, 18).

In order to minimize the bias, it seems necessary to shrink the strength domain in the Constant score. Patel et al (19) created an abbreviated Constant score, in which the strength was excluded. We postulate that complete removal of strength parameter is improper because strength is a very important component for shoulder function. In this case, it is necessary to figure out a proper maximal value for shoulder strength. Yian et al (25) reported that the mean strength point was 14-17 for men and 8-10 for women; both sexes aged 50-80 years and with no shoulder symptoms.

The purpose of this study was to determine whether the strength allocated in the Constant score would result in bias in the outcome assessment after shoulder surgery. We compared DASH score to original Constant score, an abbreviated Constant score with strength item excluded and a modified Constant score containing 12 points for the maximal strength value in the evaluation of shoulder function in patients with complex proximal humerus fractures treated by locking plate fixation.

MATERIALS AND METHODS

Patients

We retrospectively reviewed patients with locking plate fixation of 3- and 4-part proximal humerus fractures between May 2007 and December 2010. The patients were excluded if they were associated with glenohumeral dislocation, a pathologic fracture or a previous fracture at the ipsilateral humerus or glenoid, or expired during the review period. The fractures were classified according to the Neer classification (17) based on the initial radiographs and computed tomography (CT) scans, on which the fracture parts were greater than either 1 cm of displacement or 45° of angulation. One hundred and two patients, aged 20-90 years, were followed up for > 1 year. The study was approved by the Institutional Review Board of Shanghai Ruijin Hospital.

Surgical technique

The surgical technique has been reported elsewhere (10, 20). In brief, the fracture site was exposed through a deltopectoral approach. Open reduction of the main fracture and tuberosity fragments was achieved under an image intensifier. The reduction was temporarily fixed with K-wires. Once a satisfactory anatomic reduction was achieved, the K-wires were replaced by the locking plate, and fixation was performed under an image intensifier to verify the reduction of bone fragments, plate position, and screw length. Either a periarticular locking plate (Zimmer, Warsaw, Indiana, USA) or a Philos plate (Synthes, Oberdorf, Switzerland) was used to fixate the fracture depending on surgeon’s preference. The affected arm was placed in a sling for 4 weeks. Pendulum and passive elevation/abduction exercises were started 2 days postoperatively. Controlled active mobilization with abduction and flexion beyond 90° was started 4 weeks after surgery. The patients were allowed a free active range of motion 6 weeks after surgery.

Follow-up

The patients underwent a detailed follow-up examination at an average of 30 months (range: 13-55 months). Radiological examination, with anteroposterior in external rotation and axillary views, was used to evaluate the adequacy of reduction, bone union, implant loosening, condition of glenohumeral osteoarthritis and avascular necrosis of the humeral head. Functional outcomes were assessed with the following methods.
Outcome assessment

**DASH score**

DASH is a patient-reported outcome measure that contains 30 items, which are related to activities of daily living, social activities, work activities, symptoms, sleeping, and confidence (1,2). DASH score ranges from 0 to 100, and a higher score reflects greater disability.

**Constant score (CS)**

The Constant score ranges from 0 to 100 points and is composed of 4 domains: pain (15 points), activities of daily living (ADL, 20 points), range of motion (ROM, 40 points), and strength (25 points) (7). A higher score is indicative of better function. The strength was measured by a spring balance as the patient standing with the arm in 90° of lateral elevation in the scapular plane, with the elbow extended and the forearm pronated (3,7,13). The highest strength recorded was 25 pounds (7).

**Modified Constant score**

The measurement approach was the same as that described above. The modification is that the allocated strength was reduced from 25 to 12 pounds. Accordingly, the best score was 87 points in this assessment.

**Abbreviated Constant score**

This assessment was recommended by Patel et al (19). The strength subscore was completely excluded, leading to the best score becoming 75 points.

**Standardization technique**

In order to facilitate data comparison, all the scores were converted into a scale from 0 to 100. The standardized score for each scoring system was calculated using the following formulae:

- Reversed DASH score (Rev.DASH) \( = 100 - \text{original score} \)
- Constant score (CS) \( = \text{original score} \)
- Abbreviated Constant score (Abb.CS) \( = \text{CS} \div 75 \times 100 \)
- Modified Constant score (Mod.CS) \( = \text{CS} \div 87 \times 100 \)

For each standardized score, 100 represent a normal shoulder and 0 a shoulder with maximum disability. All standardized scores were graded as excellent (> 85-100), good (> 70-85), moderate (> 55-70), and poor (0-55) categories. The excellent and good categories were defined as a satisfactory outcome, whereas the moderate and poor categories were defined as an unsatisfactory outcome. In other words, 70 is a cut-off point to discriminate between the satisfactory and unsatisfactory outcomes.

**Statistical analysis**

Continuous variables were expressed as mean and standard deviation (SD). For continuous variables, the mean values were compared using Student t test or one-way ANOVA. Mann-Whitney test or Kruskal-Wallis test was used if the variable was not normally distributed. Categorical variables were compared using Fisher exact (< 100 cases) or Chi-Square test (> 100 cases). Correlations between variables were analyzed with Pearson correlation coefficient and multiple regression analysis. The ability of strength to discriminate between patients with satisfactory (Rev.DASH > 70) and unsatisfactory (Rev. DASH < 70) outcome was examined by construction of receiver operating characteristic (ROC) curves. The values for sensitivity, specificity, area under the curve (AUC) and cut-off point were obtained from ROC analysis. Level of significance for all statistics was set to p < 0.05.

**RESULTS**

Of 102 patients aged 20-90 years, 79 had 3-part fracture and 24 had 4-part fracture. Sixty nine patients were women and 34 were men. The average follow-up period was 30.3 months. There were no significant differences in age, fracture type, and follow-up period between men and women. Nineteen patients sustained complications shown in X-ray including 9 subacromial impingements 3 screw cut-outs 2 mal-unions 2 tuberosity resorptions 2 ischemic necrosis of the humeral head and 1 screw breakage.

<table>
<thead>
<tr>
<th>Table I. The data for different scores</th>
<th>Standardized Scores</th>
<th>range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reversed DASH score (Rev.DASH)</td>
<td>81.3 (10.6)</td>
<td>40 - 94</td>
</tr>
<tr>
<td>Modified Constant score (Mod.CS)</td>
<td>76.5 (12.6)</td>
<td>36 - 100</td>
</tr>
<tr>
<td>Abbreviated Constant score (Abb.CS)</td>
<td>75.6 (11.9)</td>
<td>40 - 100</td>
</tr>
<tr>
<td>Constant score (CS)</td>
<td>68.1 (12.2)</td>
<td>33 - 86</td>
</tr>
<tr>
<td>Pain</td>
<td>11.7 (2.6)</td>
<td>3 - 15</td>
</tr>
<tr>
<td>Activities of Daily Life (ADL)</td>
<td>16.4 (3.14)</td>
<td>4 - 20</td>
</tr>
<tr>
<td>Range of Motion (ROM)</td>
<td>28.8 (5.68)</td>
<td>14 - 40</td>
</tr>
<tr>
<td>Strength for 25 Pounds</td>
<td>11.3 (4.85)</td>
<td>1 - 25</td>
</tr>
<tr>
<td>Strength for 12 Pounds</td>
<td>9.66 (3.05)</td>
<td>1 - 12</td>
</tr>
</tbody>
</table>

Acta Orthopaedica Belgica, Vol. 82 - 2 - 2016
that the satisfaction rate from CS was significantly lower than that from Rev.DASH (p < 0.001) (Table III). However, there was no significant difference in satisfaction rate between Rev.DASH and Abb.CS and Mod.CS (Table III). Both men and women showed similar results (Table III). ROC analyses showed that the cut-off point and area under the curve (AUC) for strength subscore were 7 and 95.9% (95% CI: 90.0-98.8%, p < 0.001) (Fig. 3), suggesting that 7 pound strength may satisfy patients with their shoulder function.

### DISCUSSION

Both DASH and Constant scores are frequently used in the evaluation of functional outcome in patients with surgical management of shoulder injuries and diseases (1,2,24). In this study, we used DASH and Constant scoring systems to determine the functional outcomes of the shoulder in patients with complex proximal humerus fractures treated by locking plate fixation. DASH score is a subjective assessment in which the functional results are reported by patients, whereas Constant score is an objective assessment in which the functional results are determined by clinician (2,11,21,24). Our results demonstrated that the outcomes evaluated by DASH score were remarkably different than that by Constant score. The standardized values showed
that the mean Constant score was about 13 points lower than DASH score, and 43% of patients with satisfied outcome in DASH assessment become unsatisfied in Constant assessment. Walch et al (23) evaluated the functional outcome by Constant score and patient subjective assessments in 307 patients who had received arthroscopic biceps tenotomy for rotator cuff tears. They found that 87% of patients were satisfied or very satisfied with the result, whereas the mean Constant score remained in the poor category (< 70 points). It suggests that constant score may sometimes result in biased outcome during follow-up of patients with treatment of shoulder injury and disease.

Over-allocation of maximum shoulder strength in Constant score has been considered as the reason of producing biased results (16,18,19). Since only a few people, even those with healthy shoulders, can lift 25 pounds with straight arm (13), most subjects are unable to get 100 points in a Constant assessment (13,25). It is well-known that the shoulder strength is lower in female and decreases with aging (3,6,25). However, very few women and elderly people about their shoulder function because of decreased strength. Indeed, the strength of lifting a 25 pound object is often not required in daily life (6).

Patel et al (19) noticed the disadvantage of strength allocation in the Constant scoring system, so they developed an abbreviated Constant score in which the strength item was excluded. Othman et al (18) found that the correlation between Oxford score (a subjective assessment) and Constant score were significantly improved when the strength subscore was deleted. However, the strength, after all, is an important component of shoulder function. Therefore, it looks unreasonable to exclude the strength item in the assessment of shoulder function. Recently we found that in Chinese people aged > 50 years and

Fig. 2. — Comparison of the similarity of reversed DASH score (Rev.DASH) to modified Constant score (Mod.CS), abbreviated Constant score (Abb.CS) and Constant score (CS).

Fig. 3. — ROC curves for discrimination between patients with satisfactory (Rev.DASH > 70) and unsatisfactory (Rev. DASH < 70) outcome by shoulder strength. The sensitivity, specificity, area under the curve (AUC) and cut-off point for strength subscore were 86.2, 93.5, 95.9% and 7, respectively.
with weak strength. The result of ROC analysis suggested that a > 7-pound shoulder strength was needed by patients who were satisfied with their shoulder function assessed by DASH score (Rev.DASH > 70). In modified Constant score, we allocated 12 pounds for maximum strength, which is enough to accomplish the functions referred to in DASH score. Compared to original Constant score, both modified and abbreviated Constant scores can significantly decrease the disagreement with DASH score. In our study, the rate of satisfaction after locking plate fixation of complex proximal humerus fractures were 85.2% for DASH score, 78.4% for modified Constant score and 75.5% for abbreviated Constant score, but only 42.2% for original Constant score. The satisfaction rate derived from original Constant score was significantly lower than that from DASH, modified Constant and abbreviated Constant scores in the same group of patients. However, no significant differences were seen among DASH, modified Constant and abbreviated Constant scores. Synthesizing these results, we agree with the view that in original Constant score the strength was over-allocated, which may cause biased results (16).

Dawson et al (8,9) reported that the patient-based assessment, such as Oxford and DASH scores, could better reflect patients’ own perception of their life quality than Constant score after shoulder

Table III. — The satisfaction rate for functional outcome determined by different scores

<table>
<thead>
<tr>
<th></th>
<th>Satisfied N (%)</th>
<th>Unsatisfied N (%)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reversed DASH score (Rev.DASH)</td>
<td>87 (85.3)</td>
<td>15 (14.7)</td>
<td></td>
</tr>
<tr>
<td>Modified Constant Score (Mod CS)</td>
<td>80 (78.4)</td>
<td>22 (21.6)</td>
<td>0.276</td>
</tr>
<tr>
<td>Abbreviated Constant Score (Abb CS)</td>
<td>77 (75.5)</td>
<td>28 (24.5)</td>
<td>0.112</td>
</tr>
<tr>
<td>Constant Score (CS)</td>
<td>43 (42.2)</td>
<td>59 (57.8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Women</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reversed DASH score (Rev.DASH)</td>
<td>55 (79.7)</td>
<td>14 (20.3)</td>
<td></td>
</tr>
<tr>
<td>Modified Constant Score (Mod CS)</td>
<td>52 (75.4)</td>
<td>17 (24.6)</td>
<td>0.663</td>
</tr>
<tr>
<td>Abbreviated Constant Score (Abb CS)</td>
<td>51 (73.9)</td>
<td>18 (26.1)</td>
<td>0.546</td>
</tr>
<tr>
<td>Constant Score (CS)</td>
<td>28 (40.6)</td>
<td>41 (59.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Men</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reversed DASH score (Rev.DASH)</td>
<td>32 (67.0)</td>
<td>1 (3.0)</td>
<td></td>
</tr>
<tr>
<td>Modified Constant Score (Mod CS)</td>
<td>28 (54.9)</td>
<td>5 (15.2)</td>
<td>0.197</td>
</tr>
<tr>
<td>Abbreviated Constant Score (Abb CS)</td>
<td>28 (78.6)</td>
<td>7 (21.2)</td>
<td>0.054</td>
</tr>
<tr>
<td>Constant Score (CS)</td>
<td>16 (45.5)</td>
<td>18 (54.5)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Compared Mod.CS, Abb.CS and CS with Rev.DASH separately

without shoulder symptoms the mean shoulder strength was 15 pounds in men and 10 pounds in women (data not shown). Based on these data, we developed a modified constant score with a maximum strength of 12 pounds, which is enough to handle the daily affairs of life.

Superficially, the subjective scoring systems, such as DASH score, do not include the measurement of strength, but strength is taken into account in some questions, for example “place an object on a shelf above your head”, “carry a heavy object (over 10 pounds)” and “weakness in your arm, shoulder and hand” (2). Our results showed a strong positive correlation between DASH score and Constant score, and multiple regression analysis indicated that strength subscore was independently correlated with DASH score. Bey et al (4) reported that there were significant associations between shoulder strength and the patients’ subjective assessment. These lines of evidence suggest that the strength is subconsciously concerned by patients undergoing subjective assessment. Normal subjects, especially women and the elderly, are satisfied to the strength that can meet the functional performance in daily activities and occupations. However, the subjective satisfaction can be reduced if the strength is too weak. For example, some activities listed in DASH scoring system cannot be accomplished by patients

Acta Orthopædica Belgica, Vol. 82 - 2 - 2016
surgery. Comparing to DASH score, we found that Constant score was apt to provide poorer postoperative outcome. The results suggest that the bias of Constant score may be attributed to the over-allocation of strength subscore. However, the modified Constant score can significantly reduce the discrepancy of strength subscore. However, the amount of reduction of maximal strength value merits further investigation.

Acknowledgements

This study was supported by Science and Technology Commission of Shanghai Municipality (Grant no.: 124119a8500).

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: S174-188.


controlled trial in 100 patients with 1-year follow-up. 