Comparison of Subjective and Objective Outcomes After Rotator Cuff Repair


Purpose: To determine whether subjective (pain by visual analog scale) or objective (strength by dynamometer) measures correlate with disease-specific measures and quality-of-life metrics in arthroscopic rotator cuff repair.

Methods: The study population included patients who underwent primary arthroscopic rotator cuff repair at a single institution between 2006 and 2009. Within these parameters, data from 166 patients was obtained. Data were collected prospectively and reviewed retrospectively. Preoperative and 1-year postoperative data were compared. Correlation was determined in (1) disease-specific metrics including American Shoulder and Elbow Surgeons (ASES) score, Constant score (CS), and Simple Shoulder Test (SST) score and (2) quality of life measured by the Short Form 12.

Results: Preoperative strength and pain are closely associated with postoperative changes ($P < 1 \times 10^{-5}$, with β coefficients of 0.8 to 1.0). Change in ASES score was most closely associated with change in pain and change in CS with change in strength ($R^2 = 0.82$ and $R^2 = 0.54$, respectively). Only the SST score was found to be statistically linked to changes in both strength and pain ($P < .05$).

Conclusions: Patients, despite sex and age, with good preoperative strength and high preoperative pain will benefit most from arthroscopic rotator cuff repair. The CS best captures changes in strength, and the ASES score best captures changes in pain. Only changes in the SST score show a statistically significant link with changes in both strength and pain.

Level of Evidence: Level IV.

The primary goals of rotator cuff repair are to reduce pain and increase function. The manner in which these goals are considered and documented is not uniform because a myriad of assessment tools exist. To date, the reliability and validity of several assessment tools within the context of rotator cuff repair have been studied. Current literature includes evaluation of one of these particular tools or direct comparison of multiple tools. Continuous evaluation is necessary considering the growing role of quality reporting and transparency in health care.

This study focused on 4 validated and commonly used measures: the Constant score (CS), the American Shoulder and Elbow Surgeons (ASES) index, the Simple Shoulder Test (SST), and the Short Form 12 (SF-12). The CS, also known as the Constant-Murley score, is a 100-point score in which 35% is populated by subjective patient-reported metrics such as pain and activities of daily living. The remaining 65% consists of clinical parameters composed of range of motion and strength testing. In an introductory study, Constant and Murley validated their score by assessing 100 abnormal shoulders and observing a 3% interobserver error (range, 0% to 8%). The CS was later assessed by Conboy et al. and was found to have low systematic error but also low reliability.

The ASES score, which is also scored out of 100 points, assesses patient pain and function with 50% emphasis on each. Ten questions regarding activities of daily living are used for function analysis, whereas one question, namely the visual analog scale (VAS) score, is used for pain assessment. The ASES shoulder score, which is derived entirely from patient self-evaluation, is most often cited. Although it was not validated at its inception, the ASES shoulder score has since been extensively studied and validated.
The SST was developed by the shoulder service at the University of Washington. The SST consists of 12 dichotomous questions that focus on shoulder function and, like the ASES shoulder score, are based solely on patient self-assessment. The SST has recently been independently validated. The SF-12 is a commonly cited patient-completed questionnaire that is used primarily to measure quality of life. In the context of the shoulder and the rotator cuff in particular, the SF-12 and similar metrics are often used in conjunction with disease-specific measures to quantify the impact of disease and treatment on quality of life.

To our knowledge, there are no published studies that investigate the degree to which change in strength and change in pain correlate with validated outcome metrics in rotator cuff disease. The purpose of this study was to determine whether subjective (pain by VAS) or objective (strength by dynamometer) measures correlate with disease-specific measures and quality-of-life metrics in arthroscopic rotator cuff repair. The secondary objective was to ascertain the disease and demographic characteristics that drive changes in strength and pain. We hypothesized that quality-of-life metrics (SF-12) would correlate more closely with a decrease in pain rather than an improvement in strength. Conversely, validated functional outcome metrics (ASES score, CS, and SST score) would more closely correlate with strength rather than pain.

Methods

Patients

The study population included patients receiving primary arthroscopic rotator cuff repair at a single institution between 2006 and 2009. This cohort included patients who may have concomitantly undergone acromioplasty, distal clavicle excision, or subacromial decompression. Exclusion criteria included patients who underwent reoperation before 1-year follow-up and those from whom we were unable to obtain complete follow-up information. Within these parameters, data from 166 patients were obtained. Data were collected prospectively and reviewed retrospectively. Preoperative and 1-year postoperative data were compared.

Dependent Variables

Pain and strength were the primary outcome variables. Strength was measured in an isometric fashion with a digital dynamometer (Isobex; Medical Device Solutions, Burgdorf, Germany) by independent observers. To minimize bias introduced by different examiners, strength was expressed as the percentage of change from preoperative to postoperative and then compared with the contralateral arm. Data included isometric measurements of forward flexion (FF) in the frontal plane at 90° of flexion, as well as external rotation (ER) with the arm adducted and the elbow at 90° of flexion. These measurements were obtained 3 times with a minimum of 1 minute of rest, and the mean of the 3 measurements was used for each patient. Pain was measured according to the VAS.

Independent Variables

Two primary groups of metrics were examined: (1) validated, disease-specific measures including the ASES score, CS, and SST score and (2) general quality-of-life measurements (SF-12 score). In addition, the degree of change in pain and strength was compared with patient- and disease-specific characteristics including sex and age.

Statistical Analysis

First, we assessed the descriptive statistics of the variables. The mean, median, standard deviation, interquartile range, and distribution were examined for the experimental conditions. The raw preoperative and postoperative values, in addition to the difference between those values, were then examined.

We assessed the degree of agreement between the change in pain and change in strength themselves and also between pain and strength in relation to the previously mentioned measures using Spearman rank correlation. Correlation coefficients of greater than 0.50, 0.35 to 0.50, and less than 0.35 were considered strong, moderate, and weak, respectively. The distribution of scores, means, and standard deviations, as well as median pain and strength scores, were also examined. The degree of agreement between strength and pain was analyzed with the intraclass correlation coefficient and a Bland-Altman plot.

To evaluate the relation between pain and strength measures with the metrics of interest (i.e., functional outcome scores and quality of life), we applied multiple regression modeling. For the continuous measures of functioning and quality of life, we applied ordinary least squares regression models. First, in separate models, we evaluated whether (1) changes in strength and (2) changes in pain were associated with the metrics of interest. In each of these models, a term representing the baseline measure of strength or pain was included. We also tested whether quadratic terms representing the score changes contributed significantly to the fit of the models. Both measures of change in strength and change in pain, along with baseline values, were then included to evaluate whether both change scores were independently associated with the patient metrics.

To evaluate whether relations between changes in pain and changes in strength differ according to baseline strength or baseline pain, statistical interactions...
were tested. First, we evaluated interaction terms when modeled as continuous covariates. Patients were then categorized into groups representing lower and higher pain and lower and higher strength at baseline. We assessed 4 groups for each variable using quartiles to delineate the groups. Using these categories, we evaluated whether changes in pain or strength vary significantly across the baseline categories.

After investigating the relation between changes in pain and strength and the validated patient measures, attention was turned to the relation between pain and strength and patient- and disease-specific predictors. Multiple regression modeling was applied according to the procedure described earlier. All data were entered into an Excel spreadsheet (Microsoft, Redmond, WA) and analyzed with SAS software (SAS Institute, Cary, NC). All statistical tests were 2-tailed, and significance was set at .05.

**Results**

Among the cohort of 166 patients, 109 were men and 57 were women. On the basis of an a priori power analysis with estimated effect size of 0.1, an α of .05, and a power of greater than 0.9, our n-value exceeded the calculated requirement of n = 159. The rotator cuff repair was performed arthroscopically in all cases, with none requiring conversion to an open procedure. The mean age at surgery was 55.2 years (range, 34 to 78 years) for men and 56.3 years (range, 26 to 78 years) for women. The mean follow-up period was 1.69 years. The mean change in pain in the overall sample showed a decrease of 3.25 points on the VAS (median, 3.0 points; SD, 3.0 points), with a maximum of 10.0 points and a minimum of –6.0 points. The mean increase in FF strength in the overall sample was 1.43 kg (median, 1.2 kg; SD, 3.46 kg), with a maximum of 18.5 kg and a minimum of an 11.3-kg decrease in strength. The mean increase in ER strength in the overall sample was 1.42 kg (median, 0.9 kg; SD, 4.09 kg), with a maximum of 16.06 kg and a minimum of a 14.8-kg decrease in strength. The SST score correlated weakly with pain and moderately with strength, and the ASES score correlated strongly with pain and moderately with strength, whereas the CS correlated strongly with strength and preoperative pain scores were highly significant ($P < 1 \times 10^{-5}$) with high β coefficients, as will be further delineated.

Because ER and FF strength tightly trended together, the multiple regression model included only change in FF strength and pain. The SF-12 score, SST score, ASES score, and CS, as well as preoperative strength, preoperative pain, sex, and age, were set as the predictors, and the outcomes of the regression were set as change in strength and change in pain. The complete results of the multiple regression analysis are shown in Table 3.

The age and sex of the patient were not statistically significant variables in any model. Preoperative strength and preoperative pain scores were highly significant ($P < 1 \times 10^{-5}$) with high β coefficients, as will be further delineated.
analysis of respective patient metrics with strength was highest for CS (0.57). The $R^2$ values for SF-12 score, ASES score, and SST score were 0.37, 0.38, and 0.36, respectively. Changes in strength exerted statistically significant independent effects on all 4 patient metrics ($P < 1 \times 10^{-4}$), but the effect size varied. The $\beta$ coefficients of the patient metrics and change in strength were 0.25 for SST score, 0.1 for CS, 0.09 for SF-12 score, and 0.06 for ASES score. The preoperative SF-12 score and CS were statistically significant with negative $\beta$ coefficients (-0.098 and -0.075). The preoperative ASES score and SST score did not influence the models.

Preoperative pain had a $\beta$ coefficient of 0.97 to 1.07 with $P < 1 \times 10^{-7}$. The adjusted $R^2$ value of multiple regression analysis of respective metrics with pain was highest for the ASES score (0.82). The $R^2$ values for SF-12 score, SST score, and CS were 0.64, 0.69, and 0.74, respectively. The $\beta$ coefficients of the patient metrics and change in pain were $-0.25$ for SST score, $-0.07$ for ASES score, $-0.07$ for SF-12 score, and $-0.06$ for CS. All preoperative outcome scores were statistically significant with $\beta$ coefficients of 0.05 to 0.3.

### Discussion

There are several notable findings within this study. First, age and sex have no significant association with change in strength or change in pain. All of the patient outcome metrics tested do have statistically significant associations, the strongest of which is the preoperative strength measurement and the preoperative pain measurement. Interestingly, a high preoperative strength measurement yields a large change in strength. Similarly, a high initial pain score predicts a large change in pain. The clinical relevance herein is that a patient, regardless of sex and age, with good preoperative strength and high preoperative pain may benefit most from arthroscopic rotator cuff repair.

Our study does not agree with 2 previously completed studies showing that female sex is associated with lower functional scores after rotator cuff repair.28 However, patient strength has been shown to be comparable across age groups because older patients (mean age, 60 years) have the same potential for healing and recovery of strength, with near equivalence when the operative and nonoperative shoulders are compared after rotator cuff repair.29 Regarding strength and pain in particular, our findings are consistent with those of MacDermid et al.,30 who found that patients may attain good outcome scores by achieving pain relief and restored range of motion; however, they may be unsatisfied because of a lack of strength restoration. Therefore it can be said that patients who have greater strength preoperatively have greater changes postoperatively, which ultimately improves overall satisfaction and

<table>
<thead>
<tr>
<th>Change in FF</th>
<th>Change in ASES Score</th>
<th>Change in CS</th>
<th>Change in SST Score</th>
<th>Change in SF-12 Score</th>
<th>Change in FF</th>
<th>Change in ER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman correlation</td>
<td>0.41</td>
<td>0.67</td>
<td>0.43</td>
<td>0.25</td>
<td>0.29</td>
<td>0.45</td>
</tr>
<tr>
<td>$P$ value</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>0.021</td>
<td>0.003</td>
<td>$&lt;.0001$</td>
</tr>
<tr>
<td>Change in ER</td>
<td>Spearman correlation</td>
<td>0.29</td>
<td>0.45</td>
<td>0.34</td>
<td>0.13</td>
<td>0.69</td>
</tr>
<tr>
<td>$P$ value</td>
<td>0.003</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>.11</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
</tr>
<tr>
<td>Change in VAS</td>
<td>Spearman correlation</td>
<td>0.69</td>
<td>0.40</td>
<td>0.28</td>
<td>0.23</td>
<td>$&lt;.0001$</td>
</tr>
<tr>
<td>$P$ value</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
<td>$&lt;.0001$</td>
</tr>
</tbody>
</table>

*Correlation of independent variables.

**Note.** Change in strength correlates most strongly with the CS and change in pain with the ASES score. All values were statistically significant, with $P < .05$, meaning that these Spearman correlation coefficients across all groups remained statistically significant. The change in pain correlated very weakly with the change in strength.

### Fig 1.

Bland-Altman plot of change in pain and strength. The x-axis represents the change in pain (VAS) (chg_vas), and the y-axis represents the change in strength (FF) (chg_FF). The plot shows a weak correlation overall. The correlation is strongest nearest the mean; however, it is not uniform across the whole cohort. This plot suggests that decreases in pain and increases in strength are not uniformly observed in the same patients.
patient outcomes, despite the amount of pain relief experienced.

The second key finding of this study is that the CS best captures changes in strength \((R^2 = 0.57)\) and the ASES score best captures changes in pain \((R^2 = 0.82)\) on univariate analyses of the validated outcome metrics. Multiple regression analysis including all patient metrics confirms that change in CS has the most dominant association with change in strength \((P < 1 \times 10^{-5})\) and that changes in ASES score and SST score have the closest link with change in pain. That is, if one is interested specifically in either strength or pain independent of the other, these are the best-suited metrics. If capturing changes in both are desired, this study indicates that the CS, in addition to the ASES or SST score, should be collected.

Upon exclusion of the SF-12, a generic quality-of-life metric, multiple regression analysis including the disease-specific metrics as predictors shows that a change in SST score is also linked with a change in strength \((P = .04)\). Although all metrics were found to have a significant association with change in strength and pain on univariate regression, the SST score was the only patient metric to have a statistically significant association with both a change in strength and a change in pain confirmed on multiple regression analysis. This is a critical conclusion to this study because the SST score is solely patient-reported and thus requires fewer resources than the CS, for instance, which requires clinical examination. This could ultimately facilitate broader application and efficient use of the SST. Finally, it is noteworthy that the SF-12 performed poorly in capturing both changes in strength and changes in pain on multiple regression analysis. This finding confirms the necessity of collecting a disease-specific measure that will reflect the characteristics of the disease as well as a general quality-of-life metric to allow comparisons with other pathologic entities.

Next, the relation between strength and pain is weak, as evidenced by Fig 1, suggesting variable effects of the ability of rotator cuff repair to increase strength and decrease pain. One would expect a successful rotator cuff repair to both increase strength and decrease pain. At the mean, this is correct; the average outcome for this cohort showed an increase in strength and a decrease in pain. However, this study suggests that decreases in pain and increases in strength are not uniformly observed in the same patients. Covariates that lead to increased strength may not alleviate pain. These findings deserve further study.

Finally, FF strength appears to exert greater effects than ER strength on the studied patient metrics. The change in FF strength was strongly correlated with ER strength, but FF exerted greater influence on each metric. In particular, an insignificant correlation was noted with the SF-12 for ER strength, whereas a weak but statistically significant correlation was found with FF strength. This suggests that if one desires to measure strength, only FF may be captured, with a minimal loss of quality of the data, and this cuts the collection time in half.

The role of outcome measurement is growing. The Physician Quality Reporting Initiative, started in 2007, laid the groundwork for outcome reporting and transparency. The Patient Protection and Affordable Care Act expanded on the Physician Quality Reporting Initiative with efforts to increase transparency and make reporting mandatory.

It is the responsibility of providers to develop outcome measures that best reflect treatment effects. In addition to accuracy, an ideal outcome metric must be as short as possible to allow widespread collection. This study suggests that the SST and CS may be the preferred outcome measures when one is evaluating the specific goals of rotator cuff repair. However, ease of administration between the SST and CS should be further examined. It stands to reason that the SST is a resource-sparing metric because it does not require physical examination.

The application of patient preferences is important as well. Data from our institution suggest that patients may value strength restoration over pain alleviation. Most

**Table 3. Multiple Regression Analysis**

<table>
<thead>
<tr>
<th>Variable</th>
<th>(R^2)</th>
<th>Scale of Outcome Measure</th>
<th>(\beta) Coefficient</th>
<th>SE</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in FF strength</td>
<td>0.58</td>
<td>100</td>
<td>0.099</td>
<td>9.8 \times 10^{-3}</td>
<td>&lt;1 \times 10^{-5}</td>
</tr>
<tr>
<td>Change in CS</td>
<td>0.36</td>
<td>12</td>
<td>0.25</td>
<td>6.3 \times 10^{-2}</td>
<td>1.2 \times 10^{-4}</td>
</tr>
<tr>
<td>Change in ASES score</td>
<td>0.38</td>
<td>100</td>
<td>0.01</td>
<td>1.1 \times 10^{-2}</td>
<td>&lt;1 \times 10^{-5}</td>
</tr>
<tr>
<td>Change in SF-12 score</td>
<td>0.37</td>
<td>12</td>
<td>0.05</td>
<td>2.5 \times 10^{-2}</td>
<td>&lt;3.1 \times 10^{-4}</td>
</tr>
</tbody>
</table>

**NOTE.** Statistically significant effects of changes of the 4 metrics on changes in both strength and pain were found. The SF-12 score, SST score, ASES score, and CS, as well as preoperative strength, preoperative pain, sex, and age, were set as the predictors, and the outcomes of the regression were set as change in strength and change in pain.
likely, different patients value different outcomes. For example, an elderly patient may value pain alleviation, whereas a laborer values strength restoration. The treatment for each of these may be different; a double-row repair or biologics to enhance healing may be indicated in the patient valuing strength, whereas a low threshold for biceps treatment or a distal clavicle excision may be warranted for a patient valuing pain relief.3,34 A treatment to decrease pain may lead to a successful outcome for the elderly patient whereas the laborer will be dissatisfied. Better delineating which interventions decrease pain and which interventions increase strength should lead to more efficient care.

Limitations
There are several weaknesses in this study. First, we did not examine all available disease-specific measures. The University of California, Los Angeles shoulder score; Disabilities of the Arm, Shoulder and Hand score; and Rowe score, for example, were not available. Second, although the data were collected prospectively, the research question was formulated retrospectively. In addition, patients were not included in a consecutive manner because patient loss to follow-up. A superior study design would have been a more formal prospective cohort. Although all examiners were trained to perform testing in a homogeneous manner, interobserver and intraobserver sampling error is unavailable because the examiner identification was not recorded with patient data. Lastly, we did not examine range-of-motion changes, which might have had an impact on outcome in addition to pain and strength.

Conclusions
Patients, despite sex and age, with good preoperative strength and high preoperative pain will benefit most from arthroscopic rotator cuff repair. The CS best captures changes in strength, and the ASES score best captures changes in pain. Only changes in the SST score show a statistically significant link with changes in both strength and pain.

References
13. Angst F, Schwezyer HK, Aeschlimann A, Simmen BR, Goldhann 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 2011;63:174-188.