Minimal important changes in the Constant-Murley score in patients with subacromial pain

Theresa Holmgren, PT, PhD\textsuperscript{a,}\textsuperscript{*}, Birgitta Öberg, PT, PhD\textsuperscript{a}, Lars Adolfsson, MD, PhD\textsuperscript{b}, Hanna Björnsson Hallgren, MD, PhD\textsuperscript{b}, Kajsa Johansson, PT, PhD\textsuperscript{a}

\textsuperscript{a}Department of Medical and Health Sciences, Division of Physiotherapy, Linköping University, Linköping, Sweden
\textsuperscript{b}Department of Orthopedics, University Hospital of Linköping; Department of Experimental Medicine, Linköping University, Linköping, Sweden

Objective: The purpose of this study was to determine the minimal important change (MIC) of improvement in the Constant-Murley score in patients with long-standing subacromial pain and in subgroups of patients with subacromial pain with and without rotator cuff ruptures.

Method: The MIC was estimated by the anchor-based MIC distribution method, which integrates an anchor- and distribution-based approach: the optimal cutoff point of the receiver operating characteristic curve (MIC\textsubscript{ROC}) and the 95\% limit cutoff point (MIC\textsubscript{95\% limit}). The study population consisted of 93 patients included in a randomized clinical trial evaluating the effect of a specific exercise strategy.

Results: The MIC\textsubscript{ROC} was found at a mean change of 17 points in the Constant-Murley score, which corresponds to a sensitivity of 91\% and a specificity of 79\%. The MIC\textsubscript{95\% limit} was found at a mean change of 24 points. In the subgroup analysis, the MIC\textsubscript{ROC} was found at a mean change of 19 points and the MIC\textsubscript{95\% limit} at 18 points in patients with an intact rotator cuff. In patients with rotator cuff ruptures, the MIC\textsubscript{ROC} was found at a mean change of 15 points and the MIC\textsubscript{95\% limit} at 30 points.

Conclusion: The Constant-Murley score is able to detect the MIC in individual patients with long-standing subacromial pain when the rotator cuff is intact. The estimated MIC values could be used as an indication for relevant changes in the Constant-Murley score in clinical practice and guide the clinician in how to interpret the results of specific treatments.

Level of evidence: Basic Science, Validation of Outcome Instrument.

Keywords: Subacromial pain; Constant-Murley score; minimal important change; clinical important change; anchor-based methods; distribution-based methods

The Constant-Murley shoulder assessment score (CM score) is one of the most commonly used scores for assessing shoulder function.\textsuperscript{5} The CM score is often used in clinical trials to evaluate the effects of interventions for shoulder disorders and as a standard in purpose to validate other clinical shoulder scores.\textsuperscript{1,10} Notably, the score is widely used and accepted even though some aspects of its reliability and validity have not been investigated fully.\textsuperscript{3,21} The main
concerns seem to be the lack of standardized guidelines in how to use the score,\textsuperscript{19,21,22} with some contending that the original publication lacked sufficient information and guidance.\textsuperscript{20} A review of the CM score along with modifications and guidelines for its use has been published to increase the standardization.\textsuperscript{4} The CM score has been shown to be excellent for gauging responsiveness to treatment in patients with subacromial pain, and it is also able to distinguish between patients with a slightly better response and those with a much better response by a global assessment score.\textsuperscript{2} A recent review concluded that evidence supports the use of the CM score for specific clinical and research applications but emphasized the need for more structured standardization. The review also stressed that important psychometric properties, such as content validity, minimal detectable change, and minimal important change (MIC), need to be further investigated in the CM score.\textsuperscript{30} To date, only one study could be found investigating the MIC in the CM score in patients undergoing rotator cuff rupture surgery.\textsuperscript{15} The MIC in patients with subacromial pain with or without rotator cuff rupture being conservatively treated has not yet been studied. Because there is widespread acceptance of the use of the CM score, it is critical to further investigate and to develop guidelines for its optimal use. Information about the MIC in patients with specific shoulder diagnoses would be helpful in designing studies to perform sample size calculations and to interpret changes in CM scores in a clinical setting. MIC is defined as “the smallest change in score in the construct to be measured which patients perceive as important.”\textsuperscript{18} In practice, the minimal detectable change is thus defined as the smallest detectable change that can be found beyond measurement errors and is often referred to as a real change or a true change.\textsuperscript{8} In large randomized clinical trials, even small changes may be statistically significant. However, the more pertinent question is whether the changes are clinically relevant for the patients. Interpreting changes in outcome measures can be difficult because the statistical significance is partly a matter of sample size rather than about the importance and clinical relevance of the observed change.

The MIC depends on several factors, such as the method used to calculate it and the patient group being studied.\textsuperscript{7,8} There is no consensus regarding the best method to determine the MIC. In this study, a combined method that integrates a distribution-based and anchor-based approach was used to define the MIC.\textsuperscript{9} The purpose of this study was to determine the MIC for improvement in the CM score in patients with long-standing subacromial pain and in subgroups of patients with subacromial pain with and without rotator cuff ruptures.

Methods

Study population

The study population included patients in a randomized clinical trial evaluating the effect of a specific exercise strategy on long-standing subacromial pain.\textsuperscript{13} The inclusion criteria were as follows: age between 30 and 65 years, diagnosis of subacromial impingement syndrome, and on the waiting list for arthroscopic subacromial decompression. An orthopedic surgeon set the following criteria for surgery: a positive Neer impingement test result (injection of 1 mL of 20 mg/mL triamcinolone mixed with 6 mL of 10 mg/mL mepivacaine), at least 6-month duration of pain, and no or unsatisfactory results after treatment with physical therapy for at least 3 months. Participants were recruited from the Department of Orthopedics, Linköping University Hospital, Sweden, between January 2008 and February 2010. All patients provided written informed consent after being given oral and written information by the orthopedic specialist at the inclusion visit.

At the inclusion visit, the orthopedic specialist also informed the patients that they would be asked at the 3-month follow-up if they still wanted to go through with surgery. If sufficient improvement occurred with exercise treatment, surgery might not be necessary. Random allocation of the patients to either a specific exercise group or a control exercise group was done. Exercises for both groups were provided by a physiotherapist. Data were collected at baseline and at the 3-month follow-up when the treatment period was completed.

Follow-up and outcome measures

Background data such as age, sex, medication, earlier treatment, occupation, and sick leave were collected at the inclusion visit. Ultrasound was performed between inclusion and the 3-month follow-up to investigate the status of the rotator cuff (intact, partial tear, or full-thickness tear). A partial tear was defined as a localized absence of the tendon in 2 orthogonal imaging planes seen as a mixed hyperechogenic and hypoechogenic region on the bursal side, on the joint side, or in the intratendinous region that did not penetrate the entire tendon. A full-thickness tear was defined as nonvisualization of the tendon throughout its thickness.

The primary outcome was change (from baseline to 3 months) in the CM score. The CM score is a shoulder-specific score that evaluates both objective measures, including range of motion and strength, and subjective measures like pain and activities of daily living. The maximum CM score is 100 points and indicates excellent shoulder function. The CM score has 4 subscales: pain (40 points), activities of daily living (20 points), range of motion (15 points), and strength (25 points). Strength in abduction was measured with a hand-held myometer (Nottingham Mecmesin Myometer; Mecmesin Corp, Sterling, VA, USA) in the standing position with the arm in the scapular plane and 90° of elevation, hand and forearm pronated. The measurement should be pain free, and the highest value of 3 measurements was used. Questions and measurements were standardized according to the original description by Constant and Murley and further developed in a later publication.\textsuperscript{4,5}

At the 3-month follow-up, after completion of the exercise program, the orthopedic surgeon asked the patients if they still wanted surgery. Patients who had a successful outcome with exercise treatment and no longer wanted surgery still had the option of surgery viable at any time until the 1-year follow-up. At this 3-month follow-up also, the patient’s global impression of change (PGIC) in symptoms due to treatment was registered on a 5-point Likert scale: (1) recovered, (2) large improvement, (3) small...
improvement, (4) unchanged, (5) worse. The following question was asked of the patient: Since beginning the treatment at this clinic, how would you describe the changes (if any) in activity limitations, symptoms, and overall quality of life related to your painful shoulder condition? This scale was dichotomized: patients who indicated that they were unchanged or had a small improvement were classified as not importantly changed, and patients who indicated that they had a large improvement or that they were recovered were classified as importantly improved.

Data analysis

A method that integrates an anchor- and distribution-based approach was used to determine the MIC (the anchor-based MIC distribution method). Two values for MIC are presented: the optimal cutoff point of the receiver operating characteristic curve (ROC) (MICROC) and the 95% limit cutoff point (MIC95% limit).9 The PGIC, dichotomized as not importantly changed (unchanged or small improvement) and importantly improved (large improvement or recovered), was used as the external criterion (anchor). The two treatment groups were ignored and the study population was considered a cohort. The correlation between the changes in the CM score and the PGIC was analyzed by Spearman r to determine if the choice of the anchor was adequate. A correlation coefficient of at least 0.5 has been recommended in earlier studies.8,9 The individual changes in the CM score (from baseline to 3 months) were compared with the PGIC categories for all the included patients.

The distribution of changes in the CM score was plotted according to the anchor (Figs. 1-3). Two cutoff points for the MIC were calculated. The MICROC considered the CM score to be a diagnostic test that distinguished between importantly improved and not importantly changed patients. The anchor function of the PGIC was to discriminate between patients who experienced a clinical important change from those who did not. The ROC cutoff point for each change in the CM score was determined by calculating the sensitivity and specificity. The sensitivity is the proportion of patients correctly classified as importantly improved by the CM score. The specificity is the proportion of patients correctly classified as not importantly changed by the CM score. The value found at the point closest to the upper left corner on the ROC curve (Fig. 4) was defined as the MICROC. This value corresponded to the point at which the sum of the percentages of misclassified patients was the lowest. The MIC95% limit was calculated by the mean change + 1.645*SD change of the not importantly changed group.8 The MICpercentage (percentage of change from baseline to 3 months) was also calculated by the ROC method because the MIC can be affected by baseline values.14

A subgroup analysis was performed by the same methods described before in patients with and without rotator cuff ruptures. Patients were divided into two groups: those with an intact rotator cuff and those with partial- and full-thickness rotator cuff tears. IBM SPSS Statistics (version 19.0; SPSS, Chicago, IL, USA) was used for all analyses.

Results

Our analysis was based on 93 patients because 2 of the 97 patients included in the original RCT study13 lacked complete CM score change or PGIC data. Another 2 patients were excluded because of deterioration in the PGIC; this exclusion is in line with earlier studies.17,18 Background data for the patients included in the analysis are presented in Table I. The mean changes in the CM score for each PGIC category are presented in Table II. Spearman ρ between the changes in the CM scores and the PGIC categories was 0.72. The sensitivity and specificity for changes in the CM scores are presented in Figure 4. The MICROC...
was found at a mean change of 17 points in the CM score, which corresponds to a sensitivity of 91% and a specificity of 79%; this resulted in a finding that 30% of the patients were misclassified. The MIC\textsubscript{95% limit} was found at a mean change of 24 points in the CM score (Fig. 1). The distribution of the not importantly changed and the importantly improved patients, with the indicated MIC\textsubscript{ROC} and the MIC\textsubscript{95% limit} cutoff points, is presented in Figure 1. The ROC cutoff point for MIC\textsubscript{percentage} from baseline to 3 months was 22%, which corresponds to a sensitivity of 71% and a specificity of 90%. This resulted in a finding that 39% of the patients were misclassified.

A subgroup analysis was performed in patients who had rotator cuff ruptures (n = 28) and in patients who had intact rotator cuffs (n = 65). The mean changes in the CM score for the PGIC categories in both subgroups are presented in Table III. Spearman \( \rho \) between the changes in the CM score and the PGIC categories was 0.65 in the group of patients

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Distribution of changes in the Constant-Murley score for patients with rotator cuff ruptures (n = 28) who reported that they were importantly improved (much improved or completely recovered) and those who were not importantly changed (slightly improved or no change). Presented with the optimal ROC and 95% limit cutoff points.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3}
\caption{Distribution of changes in the Constant-Murley score for patients with an intact rotator cuff (n = 65) who reported that they were importantly improved (much improved or completely recovered) and those who were not importantly changed (slightly improved or no change). Presented with the optimal ROC and 95% limit cutoff points.}
\end{figure}
with rotator cuff ruptures and 0.75 for patients with intact rotator cuffs. The MICROC was 15 points for the group with rotator cuff ruptures, which corresponds to a sensitivity of 82% and a specificity of 91%. The group with intact rotator cuffs had an MICROC of 19 points, which corresponds to a sensitivity of 97% and a specificity of 76%. These findings correspond to a misclassification of 27% of patients in both subgroups. The distribution of not importantly changed and importantly improved patients with rotator cuff ruptures is presented in Figure 2 with the indicated MICROC and MIC95% limit cutoff points. The same data are presented in Figure 3 for patients with an intact rotator cuff.

**Discussion**

The CM score is often used in research and clinical practice to evaluate the effect of interventions for shoulder disorders. Therefore, it is important to identify the smallest change in the CM score that patients and clinicians should consider a clinically important change.

This study presents 2 MIC values for the CM score: the MICROC and the MIC95% limit cutoff points. In the current study, these values were quite different, with the MICROC at 17 points and the MIC95% limit at 24 points. The optimal MICROC was set at the point where the overall misclassification (i.e., the sum of the percentages of false-positive and false-negative outcomes) was minimized. The MIC95% limit was defined as the 95% upper limit of the distribution of patients who were not importantly changed according to the anchor. One challenge in a combined method is to determine which cutoff point to use as the MIC. Because false positives and false negatives are weighted equally in using the MICROC, this could be the right choice if there is no reason for one of these two to be prioritized. However, to avoid the risk of classifying patients as importantly improved when the changes in their scores fall within the measurements errors of patients who were not importantly changed, it is preferable to use the MIC95% limit.

MIC values vary by the method used to calculate them, the type of anchor used, the definition of MIC for the anchor, baseline values, and the patient groups being studied. Therefore, some have suggested that the MIC could be presented as a range rather than as a fixed value. On the basis of the results of the current study, this would mean that the MIC for the CM score could be within the range of 17 to 24 points. The change of 17 points might still have clinical importance for the patient, even if it falls within the measurement error of the instrument. It is therefore important to interpret these MIC values in a clinical setting and to use different values according to the specific situation.

Kukkonen et al recently reported a change of 10.4 points in the CM score as a clinically relevant change in patients with partial- and full-thickness tears operated on. In the current study, we present an MIC value of 17 points in the CM score for patients with subacromial pain conservatively treated. The lower MIC value presented by Kukkonen et al is probably due to a more homogeneous patient group (all operated on and all with rotator cuff tears) as well as use of another method with a less sensitive anchor. Because the MIC value will vary according to these factors, it is difficult to compare the results of these studies. To our knowledge, no other studies have presented MIC values in the CM score for patients with shoulder disorders. MIC values have been determined for self-assessed shoulder outcomes such as the Disabilities of the Arm, Shoulder and Hand questionnaire, the Simple Shoulder Test, the Oxford Shoulder Score, and the visual analog scale, all using different patient groups and methods.

The size of the MIC should also be considered in light of the particular treatment. For example, a smaller MIC might be acceptable when the treatment is cheap and safe, whereas a larger MIC is appropriate if the treatment is expensive and might be considered risky. In the current study, the MIC was evaluated for a 3-month exercise program to address shoulder pain. All of the patients expected a moderate to large improvement after treatment, and this might have affected their rating on the PGIC. The fact that all patients were on a waiting list for surgery might also have affected their global assessment. If patients want to have surgery, they may give a worse assessment on this scale to justify their choice to undergo surgery.

In this study, a combined method that integrates a distribution-based and anchor-based approach was used to define the MIC. Various anchor-based and distribution-based methods have been proposed to assess the MIC, but both approaches have limitations when they are used separately. The disadvantage of using distribution-based methods to define the MIC is that they do not, in themselves, provide a good indication of the clinical importance of the observed change. Thus, the main purpose of the
MIC, which is to distinguish clinical importance from statistical significance, is ignored. Another limitation is that these methods are sample specific because the value depends on the variability of the scores in the studied sample. The use of anchor-based methods has also been questioned because these methods fail to take into account the variability of the scores in the sample. Therefore, Crosby et al recommended a combination of distribution-based and anchor-based methods to take advantage of both an external criterion (anchor) and a variability measure. de Vet et al introduced an anchor-based MIC distribution model that integrates both approaches, which is used in this study.

Because the MIC depends on baseline values and on the type of patients being studied, we performed a separate analysis of patients with rotator cuff ruptures and with an intact rotator cuff. This resulted in a lower MICROC of 15 points in patients with rotator cuff ruptures vs an MICROC of 19 points in those with intact rotator cuffs. However, the MIC95% limit was much higher in patients with rotator cuff ruptures (30 points) than in patients with intact rotator cuffs (20 points). This is in line with other studies reporting that the MIC can differ by the patient’s symptom severity. The subgroup of patients with an intact rotator cuff was more homogeneous, as reflected by decreased variability in the CM score, which resulted in similar values for the MICROC (18 points) and the MIC95% limit (19 points). A smallest detectable change greater than the MIC is noted for most measurements, which implies that these instruments are not able to detect the MIC at individual levels. Our results support that the CM score is able to detect MIC at the individual level when the subacromial pain patients are a more homogeneous population.

The MIC depends on the choice of anchor and on the anchor’s definition of important change. In the current study, we used a global assessment scale as the anchor, and unchanged or slightly improved patients were considered not importantly changed. The underlying reasoning was that if the cutoff point was in between importantly improved and not importantly changed, the cutoff point for the CM change would be at the point where the slightly

Table I  Background variables for the 2 groups and the result of the statistical analysis used for group comparison

<table>
<thead>
<tr>
<th>Groups</th>
<th>Specific exercise group (n = 50)</th>
<th>Control exercise group (n = 43)</th>
<th>Statistical analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (female/male)</td>
<td>14/36</td>
<td>21/22</td>
<td>P = .04, Pearson χ²</td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>52 (9)</td>
<td>53 (8)</td>
<td>NS, Student t test</td>
</tr>
<tr>
<td>Duration of pain (months)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>24</td>
<td>12</td>
<td>NS, Mann-Whitney U test</td>
</tr>
<tr>
<td>Range</td>
<td>6-120</td>
<td>6-156</td>
<td></td>
</tr>
<tr>
<td>Dominant side affected</td>
<td>30 (60)</td>
<td>21 (49)</td>
<td>NS, Pearson χ²</td>
</tr>
<tr>
<td>Affected shoulder, right:left</td>
<td>32:18</td>
<td>21:22</td>
<td>NS, Pearson χ²</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heavy load</td>
<td>22 (44)</td>
<td>20 (47)</td>
<td>NS, Pearson χ²</td>
</tr>
<tr>
<td>Light load</td>
<td>28 (56)</td>
<td>23 (53)</td>
<td></td>
</tr>
<tr>
<td>Sick leave at start</td>
<td>9 (18)</td>
<td>9 (21)</td>
<td>NS, Pearson χ²</td>
</tr>
<tr>
<td>Rotator cuff status *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intact</td>
<td>33 (66)</td>
<td>33 (77)</td>
<td>NS, Fisher exact test</td>
</tr>
<tr>
<td>Partial tear</td>
<td>14 (28)</td>
<td>5 (12)</td>
<td></td>
</tr>
<tr>
<td>Full-thickness tear</td>
<td>3 (6)</td>
<td>5 (12)</td>
<td></td>
</tr>
<tr>
<td>HAD (0-21)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety, mean (SD)</td>
<td>3.4 (3.1)</td>
<td>3.5 (2.5)</td>
<td>NS, Student t test</td>
</tr>
<tr>
<td>Depression, mean (SD)</td>
<td>2.2 (2.3)</td>
<td>2.2 (2.1)</td>
<td>NS, Student t test</td>
</tr>
</tbody>
</table>

Values are numbers (percentages) unless stated otherwise. HAD, Hospital Anxiety Depression scale; NS, nonsignificant (statistical level of significance P < .05); SD, standard deviation.

* Ultrasonographic examination.

Table II  Mean change in the Constant-Murley score (standard deviation) for the different categories in the patient’s global impression of change (PGIC) due to treatment

<table>
<thead>
<tr>
<th>Categories of PGIC</th>
<th>Constant-Murley score (range, 0-100 points) (n = 93)</th>
<th>Statistical analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>3 Months</td>
</tr>
<tr>
<td>Completely recovered * (n = 9)</td>
<td>59 (20)</td>
<td>94 (5)</td>
</tr>
<tr>
<td>Much improved * (n = 40)</td>
<td>50 (11)</td>
<td>77 (10)</td>
</tr>
<tr>
<td>Small improvement † (n = 18)</td>
<td>44 (15)</td>
<td>55 (18)</td>
</tr>
<tr>
<td>Unchanged † (n = 26)</td>
<td>38 (13)</td>
<td>40 (18)</td>
</tr>
</tbody>
</table>

* Categories completely recovered and much improved were considered importantly improved.
† Categories small improvement and unchanged were considered not importantly changed.

Because the MIC depends on baseline values and on the type of patients being studied, we performed a separate analysis of patients with rotator cuff ruptures and with an intact rotator cuff. This resulted in a lower MICROC of 15 points in patients with rotator cuff ruptures vs an MICROC of 19 points in those with intact rotator cuffs. However, the MIC95% limit was much higher in patients with rotator cuff ruptures (30 points) than in patients with intact rotator cuffs (20 points). This is in line with other studies reporting that the MIC can differ by the patient’s symptom severity. The subgroup of patients with an intact rotator cuff was more homogeneous, as reflected by decreased variability in the CM score, which resulted in similar values for the MICROC (18 points) and the MIC95% limit (19 points). A smallest detectable change greater than the MIC is noted for most measurements, which implies that these instruments are not able to detect the MIC at individual levels. Our results support that the CM score is able to detect MIC at the individual level when the subacromial pain patients are a more homogeneous population.

The MIC depends on the choice of anchor and on the anchor’s definition of important change. In the current study, we used a global assessment scale as the anchor, and unchanged or slightly improved patients were considered not importantly changed. This is in line with earlier studies using the same method. The underlying reasoning was that if the cutoff point was in between importantly improved and not importantly changed, the cutoff point for the CM change would be at the point where the slightly
improved patients had not improved according to the anchor. All changes greater than that would be considered to be important clinical improvement. One reason for the higher MIC\textsubscript{95\%} limit could be that the slightly improved patients were categorized as not importantly changed, which also has been demonstrated in an earlier study.\textsuperscript{19} Furthermore, global assessment scales have been questioned because these scales are seldom thoroughly investigated in terms of validity and reliability.\textsuperscript{14} It has been debated whether patients can recall their previous status, and these scales have been shown to be influenced more by recent events and by one’s current state than by changes due to treatment.\textsuperscript{7} However, global assessment scales have also been shown to be sensitive to change.\textsuperscript{12} In this, the correlation between the anchor and the change in the CM score was 0.65 to 0.75 in the different subgroups. A value of 0.5 is recommended.\textsuperscript{3}

A statistical difference at the group level depends mainly on the size of the sample and has no or limited relation to the MIC for individual patients. The 95\% limit cutoff point used in our study included the concept of statistical significance because MIC\textsubscript{95\%} limit was found at the 95\% upper limit of the distribution of the patients who were not importantly changed, which corresponds to a mean change + 1.645*SD change. The MIC as defined by the ROC cutoff point may result in MIC values that are not statistically significantly different from the mean change value of patients who do not experience an important change. The use and interpretation of MIC values in clinical trials at the group level were presented previously.\textsuperscript{11} One method is to determine the proportion of patients who show larger changes than the MIC in each treatment group and then to compare these proportions.

**Conclusion**

Here we presented MIC values for the CM score for patients with long-standing subacromial pain. To our knowledge, this is the first time that MIC values in the CM score for patients conservatively treated with subacromial pain have been presented. The CM score is able to detect the MIC in individual patients with subacromial pain when the rotator cuff is intact. In all patients with subacromial pain, the MIC value was dependent on the subgroup being studied as well as on the choice of statistical analysis. The estimated MIC values could be used as an indication of relevant changes in the CM score in clinical practice and guide clinicians in how to interpret the results of specific treatments for patients with subacromial pain. Use of the MIC values as a benchmark for clinically important changes may also improve interpretation of the results of future studies.

**Acknowledgment**

We acknowledge the support and contribution of the Physiotherapy Orthopaedic Department of the University Hospital in Linköping, Sweden. We would also like to thank Henrik Magnusson for all statistical assistance.

**Disclaimer**

The authors, their immediate families, and any research foundation with which they are affiliated did not receive any financial payments or other benefits from any commercial entity related to the subject of this article.

**References**


**Table III** Mean change in the Constant-Murley score (standard deviation) for the different categories in the patient’s global impression of change (PGIC) due to treatment in patients with and without ruptures

<table>
<thead>
<tr>
<th>Categories PGIC</th>
<th>Constant-Murley score (range, 0-100 points) (n = 93)</th>
<th>Baseline</th>
<th>3 Months</th>
<th>Mean change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not importantly changed, no cuff rupture * (n = 28)</td>
<td>43 (15) 48 (20) 5 (9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not importantly changed, cuff rupture * (n = 17)</td>
<td>37 (11) 44 (18) 7 (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importantly improved, no cuff rupture (n = 37)</td>
<td>53 (14) 81 (12) 28 (14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importantly improved, cuff rupture (n = 11)</td>
<td>5 (9) 77 (11) 28 (11)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Not importantly changed were patients in categories small improvement and unchanged.

1 Importantly improved were patients in categories completely recovered and much improved.


