Author's response to reviews

Title: Minimal clinically detectable and important changes for pain in patients with nonspecific neck pain

Authors:

Francisco M Kovacs (kovacs@kovacs.org)
Victor Abraira (victor.abraira@hrc.es)
Ana Royuela (victor.abraira@hrc.es)
Josep Corcoll (jcorcoll@dgpif.caib.es)
Luis Alegre (lalegre@ibsalut.caib.es)
Miquel Tomas (miqueltomas@ibsalut.caib.es)
Maria Antonia Mir (mamir@ibsalut.caib.es)
Alejandra Cano (victor.abraira@hrc.es)
Alfonso Muriel (alfonso.muriel@hrc.es)
Javier Zamora (javier.zamora@hrc.es)
Maria Teresa Gil del Real (mtgildelreal@kovacs.org)
Mario Gestoso (mgestoso@kovacs.org)
Nicole Mufraggi (kovacs@kovacs.org)

Version: 2 Date: 18 October 2007

Author's response to reviews: see over
REPLIES TO REVIEWERS’ COMMENTS

The authors wish to sincerely thank the reviewers for their valuable comments, which have led to the improvement of the manuscript. Point by point reply to reviewers’ comments is as follows:

Reviewer's #1 report
Title: Minimal clinically important change for pain in patients with nonspecific neck pain.
Version: 1 Date: 17 July 2007
Reviewer: Julie Fritz
Reviewer’s report:
General

1.1. Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1.1.1 The authors propose to examine the construct of minimum clinically important change, but one analytic method used is to determine minimally detectable change. These constructs are fundamentally different in their interpretation. The authors appear to consider minimally detectable change as synonymous with minimum clinically important change and anticipate that the results of the analysis should be equivalent to the methods used for calculating minimally important change. The authors should consider revising the manuscript (including the title) to indicate that they are examining both minimally detectable change and minimally important change.

In accordance with the reviewer’s comment, this has been further clarified in the updated version of the manuscript, which now reads:

Title:

MINIMAL CLINICALLY DETECTABLE AND IMPORTANT CHANGES FOR PAIN IN PATIENTS WITH NONSPECIFIC NECK PAIN.

Abstract, Objectives:

To estimate the Minimal Changes in neck and referred pain intensity that are clinically detectable….

Introduction, paragraph 1:

Minimal clinically detectable change is defined as the minimal change in the score measuring a symptom that the patient is able to perceive. Minimal clinically important change (MCIC) is defined as the minimal variation in the
score that measures a Minimal clinically detectable change is defined as the minimal change in the score measuring symptom that is meaningful for patients. Since a change needs to be perceived before being potentially clinically important, minimal clinically detectable change is one of the values searched for in studies aiming to assess MCIC. Therefore, MCIC is used as the generic abbreviation throughout this article.

1.1.2 The authors do not adequately describe the recruitment process and likely generalizability of the results. How many potentially-eligible patients were considered for inclusion over the recruitment time period? What were the reasons for exclusion?

In accordance with the reviewer’s comment, this has been further clarified in the updated version of the manuscript, which now reads:

Methods, paragraphs 5 and 6:

Inclusion criteria were: seeking health care at any of the primary care centers belonging to the Ib-Salut, for neck pain (NP) lasting 14 or more days, either with or without pain referred to the arm or arms (AP), reporting a pain severity ≥ 3 points on a pain intensity numerical rating scale (PI-NRS), and undergoing a neuroreflexotherapy (NRT) intervention.

Exclusion criteria were: data suggesting potential underlying diseases, (the current neck pain episode being the first one in a patient under age 20 or onset over 55 in which appropriate diagnostic procedures had not yet been performed before referral to the study, non-mechanical pain, widespread neurology (disseminated neurological findings), fever, weight loss, systemically unwell, a history of: significant trauma, systemic steroids, osteoporosis, cancer, drug abuse, HIV), widespread (>1 nerve root) or progressive motor weakness in the arms, and patients’ refusal to sign the study’s informed consent.

Results paragraph 1:

In this study 678 patients were recruited and 20 (2.9%) were excluded because of refusal to sign the informed consent (15 cases), fever (3 cases) and patients feeling systematically unwell (2 cases). Therefore, 658 were included, and none were lost to follow-up.

1.1.3 Background (first paragraph, first sentence) - The MCIC relates to changes in scores that may measure pain, disability, impairment etc., The MCIC is not related to "variation of symptoms" as stated in the manuscript.

In accordance with the reviewer’s comment, wording has been improved in the updated version of the manuscript, which now reads (Introduction, paragraph 1):

Minimal clinically detectable change is defined as the minimal change in the score measuring a symptom that the patient is able to perceive. Minimal clinically important change (MCIC) is defined as the minimal variation in the score that measures a symptom that is meaningful for patients. Since a change needs to be perceived before being potentially clinically important, minimal clinically detectable change is often one of the values searched for in studies aiming to assess MCIC. Therefore, MCIC is used as the generic abbreviation throughout this article.
1.1.4 Background (second paragraph, first sentence) LBP patients do not have MCIC values. Instruments used with LBP patients have MCIC values (e.g., Oswestry, ROland Morris, numeric pain rating, etc.) Please change the wording of this sentence.

In accordance with the reviewer’s comment, the wording has been changed in the updated version of the manuscript, which now reads (Introduction, paragraph 2):

Different approaches can be used to determine MCIC. One is to estimate the mean change in score in patients who actually report to have improved (referred to as “mean change score”, or MCS). Another approach is to determine a cutoff point (referred to as “minimal detectable change”, or MDC) so that a patient with a change smaller or equal to the MDC has a chance of more than 95% that no real change has taken place.

1.1.5 Methods, (study population, second paragraph) The meaning of this paragraph is not clear. How does the protocol of patients with neck pain relate to individuals with subacute/chronic LBP? Do the authors intend to say that all subjects in this study received neuroflexotherapy? What does “post-marketing surveillance” have to do with this study? Please clarify this paragraph?

In accordance with the reviewer’s comment, this has been further clarified in the updated version of the manuscript, which now reads:

Methods, Study Population, paragraph 2 and 3:

In accordance with the current treatment protocol for neck pain in routine practice in the Ib-Salut, subacute and chronic patients were derived to a specialized Unit, where patients with a pain severity ≥ 3 points on a pain intensity numerical rating scale (PI-NRS), underwent a neuroreflexotherapy (NRT) intervention. This is a minimally invasive procedure that has proven to be safe, effective and cost/effective, and has been extensively described in the literature.

Inclusion criteria were: seeking health care at any of the primary care centers belonging to the Ib-Salut, for neck pain (NP) lasting 14 or more days, either with or without pain referred to the arm or arms (AP), reporting a pain severity ≥ 3 points on a pain intensity numerical rating scale (PI-NRS), and undergoing a neuroreflexotherapy (NRT) intervention.

Discussion, paragraphs 8 and 9:

Mean duration of pain when patients entered this study was over 540 days (Table 1). During that period, they had all received many forms of treatment and many still received them during the study. Since data being analyzed in this study derive from post-marketing surveillance of neuroreflexotherapy, all of them received that specific form of treatment. This does not affect the generalizability of results from this study, since MCIC calculation relies on patients’ self-assessment of their own evolution and instruments used to assess evolution of symptoms, no matter what treatments are influencing that evolution. No data suggests that MCIC are sensitive to the treatment being received and,
in fact, they seem to be consistent even across different chronic pain conditions.\textsuperscript{11}

On the contrary, using post-marketing surveillance methods in a National Health Service to assess MCIC has a number of advantages. It makes it possible to assess MCIC values in routine practice conditions, as opposed to using data from randomized controlled trials in which Hawthorne and other unspecific effects might influence patients’ perception of global improvement and, therefore, the results. In addition, post-marketing surveillance makes it possible to recruit large representative samples and to minimize losses to follow-up, therefore giving a better general picture of what MCIC values are likely to be in “normal” clinical conditions.

1.1.6 Methods, (study population, second paragraph) - Please describe the neuroflexotherapy that these patients apparently received.

\textit{In accordance with the reviewer’s comment, the updated version of the manuscript reads (Methods, Study Population, paragraph 2):}

In accordance with the current treatment protocol for neck pain in routine practice in the Ib-Salut, subacute and chronic patients were derived to a specialized Unit, where patients with a pain severity \(\geq 3\) points on a pain intensity numerical rating scale (PI-NRS),\textsuperscript{19} underwent a neuroreflexotherapy (NRT) intervention. This is a minimally invasive procedure that has proven to be safe, effective and cost/effective, and has been extensively described in the literature.\textsuperscript{20-25} Data included in the current study derive from methods used for post-marketing surveillance of this technology in routine clinical practice.\textsuperscript{20}

1.1.7 Methods (external criterion) - Studies examining the responsiveness of various outcome measures have frequently used 7-point or 15-point scales for patient self-assessment. This study uses a 4-point scale. It is unclear if this scale is likely to provide adequate discrimination to permit identification of patients who report meaningful improvement. Can the authors provide any support for the use of this external criterion? Has its use been described elsewhere or has this 4-point scale been validated for this purpose?

\textit{In reply to the reviewer’s question:}

In fact, studies that used 5 or 7 seven categories collapsed them into 3 (improved, unchanged and deteriorated) for the analysis. As seen in the table below, categories used for calculating MDC, MCIC and ROC were very similar to those used in our study

<table>
<thead>
<tr>
<th>Studies</th>
<th>Categories presented to the patient</th>
<th>Categories used in the analysis</th>
<th>MDC</th>
<th>MCIC</th>
<th>ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ostelo*</td>
<td>1- Complete recovery</td>
<td>1- Improved (1+2)</td>
<td>3+4+5</td>
<td>2</td>
<td>1+2 vs 3+4+5</td>
</tr>
<tr>
<td></td>
<td>2- Much improved</td>
<td>2- Unchanged (3+4+5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3- Slightly improved</td>
<td>3- Deteriorated (6+7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-</td>
<td>Complete recovered</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-</td>
<td>Much improved</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-</td>
<td>Slightly improved</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-</td>
<td>No change</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-</td>
<td>Slightly worsened</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-</td>
<td>Much worse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 1- | Improved (1+2) |
| 2- | Unchanged (3+4+5) |
| 3- | Deteriorated (6) |

| Current study | 1- | Asymptomatic |
| 2- | Improved |
| 3- | No change |
| 4- | Worsened |

| 1- | Asymptomatic |
| 2- | Improved |
| 3- | No change |
| 4- | Worsened |

| 2+3 vs 1+4 |


In accordance with the reviewer’s comment, this has been further clarified in the updated version of the manuscript, which now reads (Discussion, second to last paragraph):

In this study, patients’ own classification was rated on a 4-point scale (“completely recovered”, “improved”, “unchanged” or “worsened”), while other studies have used 5 or 7-point scales to that end, in which “improvement” and “worsening” were split into further categories, such as “much improved”, “slightly improved”, “slightly worsened” or “much worse”. However, those categories must usually be collapsed at the analysis phase, so it is up to researchers to decide how to group them. On the contrary, we preferred for patients to rate their own evolution in the categories that were going to be analyzed. This may have led to patients who perceived an improvement as clinically irrelevant selecting the “unchanged” category. Taking into account the objectives of this study, we find that to be suitable.

1.1.8 Methods (Analysis) - In the analysis of optimal cut-point - did the authors actually select the cut-off point where sensitivity and specificity were equal? Typically the cut-off point selected is the one that maximizes the specificity and sensitivity, not necessarily the point at which these values are equal. Please clarify.

This was a mistake in reporting in the previous version of the manuscript. In fact, as the referee points out, we had defined the optimal cutoff point as the
one that maximizes the sum of specificity and sensitivity, and not the one in which sensitivity and specificity were equal, as was incorrectly stated in the previous version of the manuscript.

In accordance with the reviewer’s comment, this has been corrected in the updated version of the manuscript, which now reads (Methods, Analysis, paragraph 4, point 3.):

The optimal cutoff point was estimated by the point that maximizes the sum of specificity and sensitivity.

1.1.9. Methods (Analysis) In the text in this section the authors state that they will analyze all patients and those patients with arm pain as a separate analysis. In the results (for example Table 4) it appears that the authors examined all patients and those with neck pain only as a separate analysis. Please clarify.

In accordance with the reviewer’s comment, this has been further clarified in the updated version of the manuscript, which now reads (Methods, Analysis, paragraph 6):

Data from all recruited patients (both with and without AP) were included in the main analysis, in which MCIC values for neck pain were calculated. In a subgroup analysis, only patients with referred pain at baseline were included, and MCIC values for neck and referred pain were calculated.

1.1.10. Results (first paragraph) - The text in this paragraph reports that 487 patients (74%) had AP. Table 1 appears to indicate that 487 patients (74%) had NP only. Please clarify this discrepancy.

In accordance with the reviewer’s comment, the typo in Table 1 has been corrected, and it now reads:

All included patients N=658
Patients with referred pain N=487

1.1.11. The text in this paragraph reports that Table 3 include those patients with "NP only in patients who also reported AP". This statement is unclear. It appears that Table 3 includes all subjects included in the analysis. Please clarify.

In accordance with the reviewer’s comment, this has been further clarified in the updated version of the manuscript, which now reads:

Results, paragraph 3:

Table 3 shows the MCIC for neck pain estimated in all included patients, while Table 4 shows the MCIC values for both neck and referred pain only in patients who also reported referred pain at baseline. As seen in those tables, MCIC values range between 1.5 and 6.2 PI-NRS points, depending on the method used to estimate them, with ROC leading to the smallest values.

Heading for Table:
Table 3. MCIC, MDC and Optimal Cutoff points values for neck pain in all patients included in the study (with and without referred pain), and differences depending on baseline pain severity and chronicity

Heading for Table:

Table 4. MCIC, MDC and Optimal Cutoff points for neck pain and referred pain only in patients with referred pain at baseline, and differences depending on baseline pain severity and chronicity

1.2. Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1.2.1. Abstract (conclusion) - Stating that the size of the MCIC was consistent for NP or AP and for patients with or without AP seems redundant. Please clarify.

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads (Abstract, Results):

MCIC values, except MDC, were similar for neck and referred pain, and MCIC for neck pain were similar in the entire sample and in the subset of patients reporting referred pain.

1.2.2. Background (first paragraph, last sentence) - I am not certain how defining MCIC values "makes it easier to take into account patients' perceptions...". Please clarify this statement.

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads (Introduction, paragraph 1):

- to help clinicians select among different treatments by allowing them to anticipate patients’ perception of their effect.

1.2.3. Methods (study population, paragraph 3) What is meant by "widespread neurology" as an exclusion criterion?

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads (Methods, Study Population, paragraph 4):

- Exclusion criteria were: data suggesting potential underlying diseases, (the current neck pain episode being the first one in a patient under age 20 or onset over 55 in which appropriate diagnostic procedures had not yet been performed before referral to the study, non-mechanical pain, widespread neurology (disseminated neurological findings), fever, weight loss, ...

1.2.4. Results (fourth paragraph) - What does the abbreviation "RP" stand for? This has not been used previously in the text.

In accordance with this comment and comments from other reviewers, most abbreviations have been deleted, and terms have been homogenized. Therefore “referred pain” is used instead of “arm pain” or its abbreviation (“RP”).
For instance (last paragraph under Results):

However, higher baseline pain severity only led to larger ROC values (both for neck and referred pain)

1.2.5. Patients over age 55 were apparently excluded from the study. This should be acknowledged as limiting the generalizability of the results.

In reply to the reviewer’s comment:

In fact, there was no age limit for being included in this study. The onset of pain in subjects over 55 years of age was only considered to be a “red flag” (please refer to reference #26), so that those subjects underwent diagnostic procedures to rule out potential specific pathologies.

According to the reviewer’s comment, this has been further clarified in the new version of the manuscript which now reads (Methods, Study population, paragraph 4):

Exclusion criteria were:26 ….. (the current neck pain episode being the first one in a patient under age 20 or onset over 55 in which appropriate diagnostic procedures had not yet been performed before referral to the study, …..

1.3. Discretionary Revisions (which the author can choose to ignore)

1.3.1. Methods - I would imagine that the number of inhabitants covered by the Ib-Salut varies over time. The authors may wish to round off the figure (916,453 inhabitants) since the actual number likely changes on a daily basis.

According to the reviewer’s comment, this has been done in the new version of the manuscript which now reads (Methods, Setting, paragraph 1):

The Ib-Salut is a public organization that belongs to the Spanish National Health Service, in which universal, tax funded health care services are provided to every citizen. The Ib-Salut covers all of the inhabitants of the Balearic Islands, which were 916,500 when this study started.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions
Level of interest: An article of limited interest
Quality of written English: Not suitable for publication unless extensively edited
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests:
I declare that I have no competing interests
Reviewer's # 2 report
Title: Minimal clinically important change for pain in patients with nonspecific neck pain.
Version: 1 Date: 22 August 2007
Reviewer: soledad cepeda
Reviewer's report:
General

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
None

2.2. Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

2.2.1. I recommend that the authors avoid the use of so many abbreviations. In any case, the authors should make sure that they provide a complete list of abbreviations. For example, the “RP” and “MC” abbreviations are not spelled out in the manuscript and do not appear on the list of abbreviations on page 13.

In accordance with this comment and comments from other reviewers, most abbreviations have been deleted, and terms have been homogenized. Therefore “referred pain” is used instead of “arm pain” or its abbreviation (“RP”). For instance (last paragraph under Results):

However, higher baseline pain severity only led to larger ROC values (both for neck and referred pain)

2.2.2 I also recommend that the authors use the same wording when referring to the external criteria used to evaluate the meaning of declines in pain intensity in the manuscript and tables. For example, on Table 2 they used “recovered” instead of “completely recovered” and “not changed” instead of “unchanged”.

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads:

Methods, Analysis, paragraph 3, point 2:

….. of patients categorized as “unchanged” by external criterion.

Results, paragraph 2:

….. 210 (31.9%) reported feeling 2.2.2:completely recovered,…..

Results, paragraph 2:

….. “not changed” or “worsened”.

Discussion, paragraph 4:

….. The low number of patients who felt “unchanged”…..

Table 2:
2.3. Discretionary Revisions (which the author can choose to ignore)

2.3.1. In the Discussion section, page 12, the authors describe potential differences in the minimal clinically important changes between patients with acute and chronic pain; there has been previous work on the topic. The clinical meaning of decline in pain intensity has been found to be similar in patients with acute (Cepeda et al, 2003 Pain) or chronic pain (Farrar et al, 2001 Pain), despite that both studies used different methodology. The authors could cite this previous work.

*In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads:*

Abstract, Conclusions:

MCIC values in neck pain patients are similar to those for low back pain and other painful conditions.

Discussion, paragraph 3:

In general, these results are also consistent with those from studies conducted on other painful conditions.\(^{10,11}\) The consistency of those findings could be interpreted as contributing to the validity of the MCIC values deriving from the current study.

Discussion, paragraphs 6 and 7:

Therefore, it could be argued that this could limit the generalizability of results from the current study to acute patients and those with mild pain. However, both in the research and clinical settings, acute patients with mild pain are probably not those in which it is more important to estimate MCIC. In addition, although some studies suggest that MCIC are different across acute and chronic low back pain patients,\(^8\) they are similar across acute and chronic patients with different different pain conditions,\(^{10,11}\) and they are also similar between subacute and chronic low back pain patients.\(^9\)

No data suggests that MCIC are sensitive to the treatment being received and, in fact, they seem to be consistent even across different chronic pain conditions.\(^11\)

Discussion, second to last paragraph:

In this study, patients’ own classification was rated on a 4-point scale (“completely recovered”, “improved”, “unchanged” or “worsened”),\(^28\) while other studies have used 5 or 7-point scales to that end, in which “improvement” and “worsening” were split into further categories, such as “much improved”, “slightly improved”, “slightly worsened” or “much worse”.\(^1-6,11\) However, those categories must usually be collapsed at the analysis phase, so it is up to researchers to decide how to group them. On the contrary, we preferred for patients to rate their own evolution in the categories that were going to be analyzed. This may
have led to patients who perceived an improvement as clinically irrelevant selecting the “unchanged” category. Taking into account the objectives of this study, we find that to be suitable.

What next?: Accept after minor essential revisions
Level of interest: An article of limited interest
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests: I declare that I have no competing interests
Reviewer’s # 3 report
Title: Minimal clinically important change for pain in patients with nonspecific neck pain.
Version: 1 Date: 3 September 2007
Reviewer: Henrica C de Vet
Reviewer’s report:
General
This is a sound paper with sound methodology, but a large number of sloppy descriptions, unclarities, contradictions and inconsistencies, but these can be repaired. I will indicate them as much as possible.

3.1. Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

3.1.1. I suspect an error in the calculations somewhere. This comes into expression at page 9 and in Table 3. Page 9, last paragraph: it is very difficult to explain why using the ROC method, larger differences were found between sub acute and chronic NP patients, than with the MC and MDC method, because from theory one would expect it the other way around (See de Vet et al. 2007). I have also doubts on the soundness of the data and calculations because the difference between the optimal cut off point and the MDC value is so different. In theory, if the sensitivity and specificity are around 95%, the ROC method and the MDC method should provide almost the same value, because the MDC is (indirectly) defined as specificity 95% (for, 95% of patients not importantly improved have a value below the MCIC value), and the optimal cutoff point is where the % of misclassification is minimal, i.e. where sensitivity and specificity are almost equal.

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads (Methods, Analysis, paragraph 2, point 1):

The changes of scores PI-NRS were calculated by subtracting the final values from the baseline values, so that positive scores correspond to improvement.

In addition we wish to state that:

In fact, we have modified the sense of the rest and improved the wording in order to further clarify this issue and the one related to comment 3.1.7.

a) It was misleading to have positive (for MDC) and negative (for MCS and ROC) values. Therefore, we have changed the sense of the subtraction to ensure that positive values always reflect improvement.

b) In addition, although in the previous version of the manuscript we had calculated differences as: final value – baseline value (i.e., negative scores reflected improvement) we had described that saying that we had subtracted “the final values from the baseline values” which could be interpreted otherwise. We hope that the current wording better reflects what we have actually done.

In fact, as shown in Tables 3 and 4, MCIC found in this study are generally consistent with results from the study by de Vet et al (2007). In that study, MCIC values estimated through ROC were 2.5 for acute patients and 3.5 for chronic (while in this study, they were 0.5 for subacute and 1.5 for chronic: same magnitude and direction of changes). In that study, MIC method (conceptually
similar to MDC, used in this study) led to values ranging from 4.5 in acute patients to 4.7 in chronic subjects (while, as seen in Table 4, in the current study the corresponding values for neck pain are 4.1 for both subacute and chronic patients, and 3.7 in subacute patients and 3.6 in chronic subjects for referred pain). The differences in the actual values across studies may be due to the inaccuracy of estimation and to differences in the patients included in those studies (acute vs. chronic, in the study by de Vet, and subacute vs. chronic in the current study).

c) As to the difference of values deriving from the ROC and MDC methods, it must be said that specificity in the ROC curve is not 95%, but around 80%, as is now shown in tables 3 and 4. As seen in those tables and as was to be expected, the smaller the specificity, the larger the distance between ROC and MDC values.

3.1.2 In Table 3 it can be seen that you have almost this situation, but in that case the MCIC determined by the ROC method and the MDC should almost be equal. The difference between 1.5 point (ROC method) and 4.0 is much too large. Please check your data.

As stated in the previous comment and in Tables 3 and 4 (in the current version of the manuscript), we do not achieve the same value for specificity and sensitivity. Our specificity is 83%.

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, in which Tables 3 and 4 show values for sensitivity and specificity of ROC curves:

Table 3. MCIC, MDC and Optimal Cutoff points values for neck pain in all patients included in the study (with and without referred pain), and differences depending on baseline pain severity and chronicity

<table>
<thead>
<tr>
<th>Measurement of change</th>
<th>Value (all 658 patients)</th>
<th>Baseline pain severity (PI-NRS)</th>
<th>Chronicity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lowest tertile (&lt; 6 points)</td>
<td>Highest tertile (&gt; 8 points)</td>
</tr>
<tr>
<td>Neck pain (PI-NRS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDC; Value (95% CI)</td>
<td>4.0 (3.4-5.0)</td>
<td>2.9 (2.1-4.6)</td>
<td>3.6 (2.8-4.9)</td>
</tr>
<tr>
<td>MCS (SD)</td>
<td>4.1 (2.3)</td>
<td>2.6 (1.4)</td>
<td>4.9 (2.3)</td>
</tr>
<tr>
<td>ROC curve Area (IC 95%)</td>
<td>0.92 (0.87-0.96)</td>
<td>0.94 (0.85-1.00)</td>
<td>0.96 (0.91-0.99)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.93</td>
<td>0.89</td>
<td>0.96</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.83</td>
<td>0.93</td>
<td>0.80</td>
</tr>
<tr>
<td>ROC</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Table 4. MCIC, MDC and Optimal Cutoff points for neck pain and referred pain only in patients with referred pain at baseline, and differences depending on baseline pain severity and chronicity

<table>
<thead>
<tr>
<th>Measurement of change</th>
<th>Value (all 487 patients with referred pain at baseline)</th>
<th>Baseline pain severity (PI-NRS)</th>
<th>Chronicity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lowest tertile (&lt; 6 points)</td>
<td>Highest tertile (&gt; 8 points)</td>
</tr>
<tr>
<td>Neck pain</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.1.3. The distinction in subgroups is not always clear. Eg in the abstract, line 18: patients with or without AP, and for NP and for AP. Is MCIC also determined for arm pain, or for neck pain in the subgroup of patients who also have arm pain?

MCIC was determined for both neck and arm pain. In addition, MCIC for neck pain was determined both in patients with and without referred pain.

In accordance with the reviewer’s comment, this has been clarified in the updated version of the manuscript, which now reads

Abstract, Results:

MCIC values, except MDC, were similar for neck and referred pain, and MCIC for neck pain were similar in the entire sample and in the subset of patients reporting referred pain.

Methods, Analysis, paragraph 6:

Data from all recruited patients (both with and without AP) were included in the main analysis, in which MCIC values for neck pain were calculated. In a subgroup analysis, only patients with referred pain at baseline were included, and MCIC values for neck and referred pain were calculated.

3.1.4 Page 10, line 2: ‘... MCIC for local and referred pain are similar’. Does this mean that for the whole population and the subgroup with comcomitant arm pain the MCIC were similar, or is really the MCIC for arm pain assessed?

It meant that MCIC values for neck pain are similar to those for referred pain, as seen in Table 3 (identical MCS and ROC values). In addition, MCIC for neck pain were similar across the entire sample and the subgroup of patients with referred pain (identical values for MCS and ROC, variation of just 0.2 points for MDC).

In accordance with the reviewer’s suggestion, this has been further clarified in the current version of the manuscript, which now reads (Results, paragraph 3):

<table>
<thead>
<tr>
<th>Neck pain (PI-NRS)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MDC, Value (95% CI)</td>
<td>4.2 (3.4-5.5)</td>
<td>3.3 (2.3-5.7)</td>
<td>4.0 (3.0-5.9)</td>
<td>3.1 (1.9-7.5)</td>
<td>4.4 (3.5-6.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCS (SD)</td>
<td>4.1 (2.3)</td>
<td>2.5 (1.2)</td>
<td>4.8 (2.4)</td>
<td>4.1 (2.4)</td>
<td>4.1 (2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROC curve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area (IC 95%)</td>
<td>0.91 (0.86-0.97)</td>
<td>0.92 (0.80-1.00)</td>
<td>0.93 (0.88-0.99)</td>
<td>1.00 (0.99-1.00)</td>
<td>0.89 (0.82-0.96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>0.93</td>
<td>0.88</td>
<td>0.87</td>
<td>0.99</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>0.81</td>
<td>0.91</td>
<td>0.86</td>
<td>1.00</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROC</td>
<td>1.5</td>
<td>1.5</td>
<td>2.5</td>
<td>0.5</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Pain referred to the arm (PI-NRS) |   |   |   |   |   |   |   |   |   |   |   |   |
| MDC, Value (95% CI) | 6.2 (5.0-8.2) | 3.8 (2.6-6.6) | 7.1 (5.4-10.1) | 6.6 (4.3-14.6) | 6.8 (5.4-9.2) |   |   |   |   |   |   |
| MCS (SD)           | 4.1 (2.7) | 2.4 (1.7) | 5.3 (3.0) | 3.7 (3.4) | 3.6 (3.2) |   |   |   |   |   |   |
| ROC curve          |   |   |   |   |   |   |   |   |   |   |   |   |
| Area (IC 95%)      | 0.81 (0.73-0.89) | 0.80 (0.65-0.95) | 0.86 (0.76-0.96) | 0.89 (0.78-1.00) | 0.76 (0.68-0.85) |   |   |   |   |   |   |
| Sensitivity        | 0.87 | 0.79 | 0.87 | 0.90 | 0.81 |   |   |   |   |   |   |
| Specificity        | 0.71 | 0.82 | 0.77 | 0.86 | 0.71 |   |   |   |   |   |   |
| ROC                | 1.5 | 1.5 | 2.5 | 0.5 | 1.5 |   |   |   |   |   |   |
Table 3 shows the MCIC for neck pain estimated in all included patients, while Table 4 shows the MCIC values for both neck and referred pain only in patients who also reported referred pain at baseline. As seen in those tables, MCIC values range between 1.5 and 6.2 PI-NRS points, depending on the method used to estimate them, with ROC leading to the smallest values. MCIC values, except MDC, were similar for neck and referred pain, and MCIC for neck pain was similar for all patients and for the subset who reported referred pain.

3.1.4 Abstract, line 5 it says that no MCIC for NP is available. There has been a publication by Sim et al. in Clin J Pain 2006; 22: 820-826 about the MCIC of the Northwick Park Neck Pain Questionnaire.

In fact, that study was already cited, but our statement referred to neck pain severity as measured by a point intensity numerical rating scale.

In accordance with the reviewer’s suggestion, this has been further clarified in the current version of the manuscript, which now reads:

Abstract, Summary of background data:

No data on MCIC for pain severity are available for neck pain patients.

Introduction, paragraph 4:

In patients with nonspecific neck pain (NP), MCIC have been explored for disability and quality of life, but not specifically for pain severity.

3.1.5 I would suggest to label the optimal cut off point of the Receiver Operating Curves (not operant curve) as ROC, as is usually done, and not OCP.

In accordance with the reviewer’s suggestion, we have corrected “receiver operant curve” and replaced it with “receiving operating curve”. In addition, the label “ROC” is now used. For instance:

Abstract, Methods:

The mean change score (MCS) estimated by the mean value of change score in patients who self-assess as improved, the minimal detectable change (MDC) estimated by means of the standard error of measurement in patients who self-assess as unchanged, and the optimal cutoff point in receiver operating characteristics curves (ROC) were used to estimate…

Introduction, paragraph 2:

Finally, a third approach is to use receiver operating characteristics curves (ROC) to define the cutoff point that best discriminates between

3.1.6 The clarification of the three methods used on page 4 is very sloppy. The mean change method takes the mean of a special category of patients. The MDC method looks only at non improved patients and draws the cutoff point where 95% of the patients have a lower score. Note that this is not the same as: “the threshold below
which most patients deny any improvement”. Because below that point there are a large number of patients who do say they are improved on the anchor (see anchor based MIC distribution by de Vet et al QOL 2007;16:131-42).

In accordance with the reviewer’s suggestion, the manuscript has been modified and it now reads (Introduction, paragraph 2):

Different approaches can be used to determine MCIC. One is to estimate the mean of change in score in patients who actually report to have improved (referred to as “mean change score”, or MCS). Another approach is to determine a cutoff point (referred to as “minimal detectable change”, or MDC) so that a patient with a change smaller or equal to the MDC has a chance of more than 95% that no real change has taken place.

3.1.7 Page 7, last paragraph: the way the substraction is described here ‘improvement’ leads to a positive value. Note that throughout the manuscript the values of the mean change score are presented as positive values (also in Table 2- 4), while the values of the ROC method and MDC are presented as negative values. They all should be positive if the 12 weeks score is subtracted from the baseline score (more severe pain, thus higher value).

The manuscript has been modified in accordance with the reviewer’s suggestion. Please, see the answer to query # 3.1.1.

3.1.8 Page 8, line 6: I doubt whether the 95% CI of the MDC should be calculated using a chi square distribution.

The Chi square distribution was used because MDC is based on a variance.1,2


3.1.9 Page 8, forelast paragraph: here is described for which populations the MCIC for neck pain is determined, which is the total population and for the subgroup of patients who reported concomitant arm pain at baseline. Is that right? Or is the MCIC for arm pain also assessed? This is suggested in the abstract.

MCIC for neck pain was assessed in all patients and MCIC for referred pain in all patients who reported referred pain. In addition, MCIC for neck pain was also assessed in the subgroup of patients who reported referred pain.

In accordance with the reviewer’s suggestion, this has been further clarified in the updated version of the manuscript, which now reads:

Methods, Analysis, paragraph 6:

Data from all recruited patients (both with and without AP) were included in the main analysis, in which MCIC values for neck pain were calculated. In a subgroup analysis, only patients with referred pain at baseline were included, and MCIC values for neck and referred pain were calculated.

Results, paragraph 3:
Table 3 shows the MCIC for neck pain estimated in all included patients, while Table 4 shows the MCIC values for both neck and referred pain only in patients who also reported referred pain at baseline. As seen in those tables, MCIC values range between 1.5 and 6.2 PI-NRS points, depending on the method used to estimate them, with ROC leading to the smallest values. MCIC values, except MDC, were similar for neck and referred pain, and MCIC for neck pain was similar for all patients and for the subset who reported referred pain.

3.1.10 Avoid the term ‘first and later set of analyses’, as this comes not to expression in the data analysis and data presentation later on.

In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Methods, Analysis, paragraph 6):

Data from all recruited patients (both with and without AP) were included in the main analysis, in which MCIC values for neck pain were calculated. In a subgroup analysis, only patients with referred pain at baseline were included, and MCIC values for neck and referred pain were calculated.

3.1.11 Page 8, last paragraph: the patients are divided in subacute and chronic patients. I would restrict the analysis of MCIC to these subgroups. Analysis on tertiles of duration do not have clinical meaning. If you want to do an analysis on the neck pain with very long duration I would suggest to do an analysis of chronic patients with more and with less than one year duration. The reference [8] in this paragraph can be left out.

In accordance with the reviewer’s suggestions, reference 8 has been left out of the paragraph, tertiles of duration have been suppressed and pain duration is now analyzed by comparing subgroups of subacute and chronic patients. Accordingly, the manuscript now reads:

Methods, Analysis, last paragraph:

Values were also estimated for chronic and subacute patients, with the cut-off point for chronicity established at 90 days.1,28

Last paragraph under Results:

….. Both for neck and referred pain, ROC values were consistently smaller in subacute than in chronic patients (Tables 3 and 4), while MDC and MCS were similar in those subgroups.

Discussion, paragraph 1, sentence 2:

They also show that, in general, MCIC are similar across subacute and chronic patients, while the higher the baseline pain severity, the higher the improvement has to be for a patient with neck pain to perceive it as “clinically important” (Tables 3 and 4).
ROC was consistently larger in chronic than in subacute patients, while pain duration did not modify MCS. As related to MDC, values for neck pain were smaller for chronic than for subacute patients in the entire sample, while in the subset of patients who reported referred pain, MDC values were larger for chronic than for subacute patients (both for referred and, especially, for neck pain) (Tables 3 and 4). The low number of patients who felt “unchanged” or “worse” may account for those differences, which may also be due to the fact that...

3.1.12 Page 11, last sentence: It is suggested that a different definition is used for specific purposes and context. Please give examples in which situation you would choose for which method.

In accordance with the reviewer's suggestion, that sentence has been deleted.

3.1.13 Page 12, 2nd paragraph: the need for assessment in MCIC for acute patients is said to be not important in the middle of the paragraph, and ‘should be explored’ in the last sentence. This sounds contradictory.

Obviously, the reviewer is absolutely right. In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Discussion, paragraph 6):

To be included in this study, patients had to be subacute or chronic, and their pain had to be ≥3 PI-NRS points. Therefore, it could be argued that this could limit the generalizability of results from the current study to acute patients and those with mild pain. However, both in the research and clinical settings, acute patients with mild pain are probably not those in which it is more important to estimate MCIC. In addition, although some studies suggest that MCIC are different across acute and chronic low back pain patients, they are similar across acute and chronic patients with different different pain conditions, and they are also similar between subacute and chronic low back pain patients.

3.1.14 Page 12, 3rd paragraph: you might also comment on why the category of patients who reported to be ‘much improved’ was chosen as minimally important change. Some authors choose the category of ‘slightly improved’ as minimally importantly changed.

In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads:

Methods, External criterion, paragraph 1:

At the 12 week follow-up, patients scored the change in their clinical status on the following scale: 1: Completely recovered, 2: improved, 3: unchanged, 4: worsened. Patients rated their clinical status on their own and without assistance.

Discussion, second to last paragraph:

In this study, patients’ own classification was rated on a 4-point scale (“completely recovered”, “improved”, “unchanged” or “worsened”), while other
studies have used 5 or 7-point scales to that end, in which “improvement” and “worsening” were split into further categories, such as “much improved”, “slightly improved”, “slightly worsened” or “much worse”.\textsuperscript{1-8,11} However, those categories must usually be collapsed at the analysis phase, so it is up to researchers to decide how to group them. On the contrary, we preferred for patients to rate their own evolution in the categories that were going to be analyzed. This may have led to patients who perceived an improvement as clinically irrelevant selecting the “unchanged” category. Taking into account the objectives of this study, we find that to be suitable.

---

3.2 \textbf{Minor Essential Revisions} (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

3.2.1 In abstract, line 13 : I think that RP should be AP

\textit{In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Abstract, Methods):}

…..and existence of referred pain were assessed…..

3.2.2 Introduction, line 1: minimal change in the measurement instrument instead of minimal variation in symptoms. Because you are interested in the change and not the variation; and not all characteristics that are measured are symptoms.

\textit{In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Introduction, paragraph 1):}

Minimal clinically detectable change is defined as the minimal change in the score measuring a symptom that the patient is able to perceive. Minimal clinically important change (MCIC) is defined as the minimal variation in the score that measures a symptom that is meaningful for patients.\textsuperscript{1-11} Since a change needs to be perceived before being potentially clinically important, minimal clinically detectable change is often one of the values searched for in studies aiming to assess MCIC.\textsuperscript{1-11} Therefore, MCIC is used as the generic abbreviation throughout this article.

3.2.3 Page 5, last sentence: referred arm? Pain

\textit{In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Methods, Study population, paragraph 1):}

…..without referred pain.

3.2.4 Page 6, line 1: instead of ‘set at 14 days and 90 days’, ‘set at between 14 and 90 days and longer than 90 days, respectively’.

\textit{In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Methods, Study population, paragraph 1):}

Based on the available evidence for LBP patients, time limits to consider a patient as “subacute” and “chronic” were set at between 14 and 90 days, and longer than 90 days, respectively.\textsuperscript{16-18}
3.2.5 Page 6, line 4: should LBP be NP?

*In accordance with the reviewer’s suggestion,* the manuscript has been modified and now reads (Methods, Study population, paragraph 3):

……belonging to the Ib-Salut, for neck pain (NP)……

3.2.6 Page 7, line 2: follow-up

*In accordance with the reviewer’s suggestion,* the manuscript has been modified and now reads:

*Methods, Outcome assessment, paragraph 1:*

Patients were assessed on their first visit (baseline assessment) and three months later (12 week follow-up).

*Analysis, paragraph 1:*

…… for neck and referred pain were estimated for the follow-up …..

3.2.7 Page 7, line 8 and 11: clinical status or health status: please be consistent

*In accordance with the reviewer’s suggestion,* the manuscript has been modified and now reads (Methods, External criterion, paragraph 2):

Since patients graded the evolution of their own clinical status, this classification was considered as the external criterion for a change to be “clinically important”.

3.2.8 Page 8, line 4 from below: I prefer the term ‘subgroup analysis’ instead of ‘sensitivity analysis’.

*In accordance with the reviewer’s suggestion,* the manuscript has been modified and now reads (Methods, Analysis, paragraph 7):

Values were also estimated for chronic and subacute patients, with the cut-off point for chronicity established at 90 days.1,28

3.2.9 Page 9, line 9: 3 months and 12 weeks is used interchangeably. Please, stick to one of these terms.

*In accordance with the reviewer’s suggestion,* the manuscript has been modified and now reads:

*Analysis, paragraph 1:*

…… follow-up period of 3 months.

*Results, paragraph 2:*

…… and changes in scores 3 months later, …..
3.2.10 Page 9, third paragraph: also here it is suggested that MCIC for AP is being assessed.

*In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Results, paragraph 3):*

Table 3 shows the MCIC for neck pain estimated in all included patients, while Table 4 shows the MCIC values for both neck and referred pain only in patients who also reported referred pain at baseline. As seen in those tables, MCIC values range between 1.5 and 6.2 PI-NRS points, depending on the method used to estimate them, with ROC leading to the smallest values. MCIC values, except MDC, were similar for neck and referred pain, and MCIC for neck pain was similar for all patients and for the subset who reported referred pain.

3.2.11 Page 9, last paragraph: the term RP appears (also in other places in the manuscript). I think this should be AP.

*The reviewer is correct. However, in accordance with other reviewers’ suggestions, in order to increase the readability of this manuscript those abbreviations have been deleted, and now the terms “referred pain” appear instead.*

3.2.12 Page 11, line 8: ‘.. was slightly different ...’. Please specify compared to what?

*In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Discussion, paragraph 4):*

In this study, the influence of baseline pain severity and duration of pain was slightly different depending on the method used to estimate the size of MCIC.

3.2.13 Page 11, line 10: a large MC size? Do you mean MCIC size, or MCIC values assessed by the MC method? Also in the rest of this paragraph the terminology is sloppy. (MCS?)

*In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Discussion, paragraph 4):*

In this study, the influence of baseline pain severity and duration of pain was slightly different depending on the method used to estimate the size of MCIC. Higher baseline pain severity consistently led to a larger MCS size. In the entire sample, ROC value for neck pain was the same regardless of baseline pain severity, while MDC and MCS values were higher for patients with a higher baseline pain severity. On the contrary, in the subset of patients with referred pain, both for neck and referred pain a higher baseline pain severity led to larger ROC, MCS and MDC values, although the 95% confidence intervals for MDC overlapped between patients in the highest and lowest tertiles of baseline pain severity (Tables 3 and 4). ROC was consistently larger in chronic than in subacute patients, while pain duration did not modify MCS. As related to MDC, values for neck pain were smaller for chronic than for subacute patients in the entire sample, while in the subset of patients who reported referred pain, MDC
values were larger for chronic than for subacute patients (both for referred and, especially, for neck pain) (Tables 3 and 4).

3.2.14 Page 12, line 9,10: references are in superscript.

In accordance with the reviewer’s suggestion this has now been corrected.

3.3 Discretionary Revisions (which the author can choose to ignore)

What next?: Accept after minor essential revisions
Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.
Declaration of competing interests: I declare that I have no competing interests

Reviewer’s # 4 report
Title: Minimal clinically important change for pain in patients with nonspecific neck pain.
Version: 1 Date: 7 September 2007
Reviewer: Cecilie Roe
Reviewer’s report:
General
It is a well written manuscript addressing an important topic. The introduction is addressing the relevant issues, but several questions remain related to the aims, methods and results which should be taken into consideration in reviewing this manuscript.

4.1 Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

4.1.1 Only one single aim is presented in the introduction of the present study, although several questions are raised in the introduction, and presented in results and conclusion. For example if there are differences between patients with neck pain and arm pain.

In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads:

Abstract, Objectives:

To estimate the Minimal Changes in neck and referred pain intensity that are clinically detectable and clinically important (MCIC) for subacute and chronic neck pain (NP) patients, and to assess whether MCIC values for neck pain severity are different in the subset of patients reporting referred pain.

Introduction, last paragraph:

Therefore, the primary objective of this study was to calculate the MCIC values for neck and referred pain severity in subacute and chronic NP patients treated in routine clinical practice. An additional objective was to explore if MCIC values for neck pain were different in the subset of patients who reported referred pain.
4.1.2 The difference between patients with neck pain and neck and arm pain is an important aspect. However, is that actually analysed? It seems like the comparison is made between the whole sample and those with neck and arm pain. Addressing this questions requires a comparison between subjects with only neck pain and those with both neck and arm pain.

**In accordance with the reviewer’s suggestion,** this has been further clarified in the updated version of the manuscript, which now reads (Methods, Analysis, paragraph 6):

Data from all recruited patients (both with and without AP) were included in the main analysis, in which MCIC values for neck pain were calculated. In a subgroup analysis, only patients with referred pain at baseline were included, and MCIC values for neck and referred pain were calculated.

4.1.3 The main issue in this manuscript is the minimal changes in pain intensity that is of clinical importance. However, pain intensity can be measured with different specifications. These specifications may affect the level, the variations over time as well as the inter-individual differences, all of importance when MCIC should be estimated. Hence, was it present pain or pain reported over a certain time period that was assessed? What was in that case the length of the time period? Was it the normal, highest or lowest pain over the time interval or pain at rest or during normal activity that was reported.

**In accordance with the reviewer’s suggestion,** the manuscript has been modified and now reads (Methods, Outcome assessment, paragraph 2):

At each assessment, patients rated, on their own and without assistance, two separate PI-NRS for current neck and referred pain, ..... 

4.1.4 Lack of validated NP functional measurements is a problem. However, one could question the approach of applying only the subjects’ general feeling of improvement in these analyses. Hence, a thorough discussion about this weakness of the study and how it could affect the results should be included.

**In accordance with the reviewer’s suggestion,** the manuscript has been modified and now reads (Discussion, paragraphs 9 and 10):

Since no validated instruments to assess neck pain-related disability were available for Spanish speaking patients when this study started, pain was the only symptom that could be assessed and MCIC values for that variable were the only ones to be calculated. Based on previous studies in neck and low back pain patients, it is very likely that disability also influences patients’ perception of general improvement, and further studies should explore MCIC for that variable in this cultural environment. However, since MCIC for pain and disability are calculated separately, the impossibility of assessing disability does not affect the validity of the MCIC values for pain deriving from this study.

The use of patients’ classification of their own general clinical evolution as the external criterion is somewhat controversial, since it requires for them to compare their initial and final states. However, that is how patients assess their own evolution in routine clinical practice. It also happens to be the usual
“gold standard” for “patients’ subjective global improvement” in the research setting, and has a high face validity. In routine clinical practice it is of the utmost importance, since it would not make sense to classify a patient as improved or deteriorated against the patient’s own personal assessment. For these reasons this was considered to be an appropriate external criterion for the study.

4.2 Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

4.2.1 There are several abbreviations in this manuscript; RP is used in the abstract as well as in the results (page 9). It is not included in the list of abbreviations and I cannot see that it is otherwise explained.

The reviewer is correct. In accordance with the reviewer’s and other reviewers’ suggestions, in order to increase the readability of this manuscript those abbreviations have been deleted, and now the terms “referred pain” appear instead. For instance:

Abstract, Methods:

….. and existence of referred pain were assessed.

Results, last paragraph:

….. (both for neck and referred pain)…..

4.2.2 MC is used at page 9 and not included in the abbreviation list either.

In accordance with the reviewer’s suggestion, in order to clarify the text most abbreviations have now been deleted, and those that are left have an explanation (please, see query # 4.3.1).

4.2.3 Mean change of PI-NIRS in patients scoring 2 on the external criterion is used to calculate MCS. This is an appropriate way of analysing the data, but I think the authors also should provide the change in those subjects not believing they have improved, in particular because other external criteria are lacking.

In accordance with the reviewer’s suggestion, the mean change in those subjects who denied improvement is shown in Table 2 in the current version of the manuscript, under “Unchanged”.

4.2.4 A sentence about the analysis and statistics should be provided in the methods section of the abstract.

In accordance with the reviewer’s suggestion, the manuscript has been modified and now reads (Abstract, Methods):

The mean change score (MCS) estimated by the mean value of change score in patients who self-assess as improved, the minimal detectable change (MDC) estimated by means of the standard error of measurement in patients who self-assess as unchanged, and the optimal cutoff point in receiver operating characteristics curves (ROC)…..
4.2.5 Page 8, second sentence from the bottom; should it be the highest tertile.

The reviewer is obviously right. **In accordance with the reviewer’s suggestion**, the manuscript has now been changed to read *(Methods, Analysis, last paragraph)*:

..... and high baseline scores highest tertile).

4.2.6 Page 11 sentence beginning with: “That was also the case...". The meaning of this sentence is difficult to catch.

**In accordance with the reviewer’s suggestion**, the manuscript has been modified and now reads *(Discussion, paragraph 4)*:

In the entire sample, ROC value for neck pain was the same regardless of baseline pain severity, while MDC and MCS values were higher for patients with a higher baseline pain severity. On the contrary, in the subset of patients with referred pain, both for neck and referred pain a higher baseline pain severity led to larger ROC, MCS and MDC values, although the 95% confidence intervals for MDC overlapped between patients in the highest and lowest tertiles of baseline pain severity (Tables 3 and 4).

4.2.7 Table 1, should the heading of the second column be patients with AP?

**In accordance with the reviewer’s question**, the heading that column has now been changed to “Patients with referred pain”.

4.2.8 Please provide explanations to the abbreviations and external criterions in the table texts.

**In accordance with the reviewer’s suggestion** this has now been done for all tables.

4.2.9 The references should be provided in front of the "point" of the sentences not after?

**The authors will abide by the Editor’s decision**

4.3 Discretionary Revisions (which the author can choose to ignore)

4.3.1 There are several abbreviations in this manuscript. Could abbreviations for arm and neck pain be skipped in order to improve the readability?

**In accordance with the reviewer’s suggestion**, most abbreviations have now been deleted, and those that are left have an explanation.

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable
**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**
I declare that I have no competing interests.