Author's response to reviews

Title: Subacute and chronic, non-specific back and neck pain: cognitive-behavioural rehabilitation versus primary care. A randomized controlled trial.

Authors:

Odd Lindell (odd.lindell@ki.se)
Sven-Erik Johansson (sven-erik.johansson@ki.se)
Lars-Erik Strender (lars-erik.strender@ki.se)

Version: 2 Date: 9 June 2008

Author's response to reviews: see over
Comments on the report of referee 1, Eva Schonstein, 28 November 2007

The paragraph figures refer to the report of Eva Schonstein. The revised version of the article is provided with row numbers, except in the tables and figures. Within brackets are references to the revised article.

1. The term “Traditional primary care” is replaced by the term “primary care”, which we think is more accurate. We have added a description of primary care in terms of consultations by different care staff and how these were distributed among the three six-month periods (page 15, rows 352-357 and Figures 4 and 5).

2. The description of cognitive-behavioural rehabilitation has been transferred to a table (Table 1, page 29) and complemented with frequency, the duration of each session and of the entire program, and the intensity of the program. We have added an Appendix with a detailed description of how our program appeared for the rehabilitation group. We have included a description of cognitive-behavioural rehabilitation in terms of consultations by different care staff and how these were distributed among the three six-month periods (page 15, rows 345-348 and Figure 4).

3. The primary outcome measure is renamed Return-to-work share to differentiate it from a new secondary outcome measure, Return-to-work chance, which is added to fit such data more satisfactorily (page 8, rows 180-182). As background for having two return-to-work measures we have added two sentences describing return-to-work measurement (page 4, rows 69-72. The analyses concerning Return-to-work chance are described on page 9, rows 196-199; the results are shown on pages 13-14, rows 311-316, and in the new table 5, page 33, and are discussed on page 17, rows 399-401. We found that this new outcome measure made Work ability superfluous, so we excluded that outcome variable. However, the term ‘work ability’ is retained and is now clearly defined (page 7, rows 152-155). Information about who decides on work ability is added (page 6, rows 117-124).

4. The possible deficiencies of our programme are described more in detail. As background, we mention the possible connection between relapse problems and short duration of rehabilitation programmes (page 5, row 92-95), how our programme tried to address these problems (page 7, rows 155-156) and the possible negative consequences of our methodology (pages 18-19, rows 433-445).

5. Complete cost-effectiveness analyses are described, for example by Goosens and Evers (Goossens M, Evers S: Cost-Effectiveness of treatment for neck and low back pain. In Neck and back pain. The scientific evidence of causes, diagnosis, and treatment. Edited by Nachemson A, Jonson E. Philadelphia: Lippincott, Williams & Williams; 2000:399-419). As that kind of complete analysis was beyond the resources of our study, we used the same approach to estimating costs as was used in previous research [10]. The cost of our programme over 18 months was, according to those estimates, comparable with the costs of primary care; beyond 18 months it was possibly lower (page 19, rows 448-453).
6. To increase the focus on sick-listing and health-care visits in the present study, we have excluded information of the existence of a 10-test package, exclusively mentioned the carrying out of “among other tests, a lift test” and given a reference to a detailed description of the 10-test package in a previous study (page 12, rows 262-264).

7. We have excluded the retrospective data for work ability.

Discretionary Revisions:
The valuable Cochrane review by Schonstein et al. is included [43] and commented on (page 17, rows 404-407).

Other major changes, outside the views of Eva Schonstein

The last sentence of the conclusions is modified from: “Conclusive and generalizable results concerning this eventual long-term effect would demand a longer follow-up” to: “More conclusive results concerning this possible long-term effect might require a longer follow-up.” (page 3, rows 52-53, and page 21, rows 490-491).

The outcome variable Total sick-listing is renamed Net days, which is a more accepted term [30].

The title is shortened.

The article is shortened (from 5300 to 4700 words).

Nineteen of the original 49 references are excluded. To accommodate the valuable views of referees, we have included (in addition to Schonstein et al. [43]) some new references, for example, Hlobil et al. [7], Staal et al. [17] and Steenstra et al. [19]. The number of references is now 44.

The figures for Net days (former Total sick-listing) and Visits (Figures 2 a-b and 3 a-b) have been reworked to clarify the changes over time.
Comments on the report of referee 2, Jos H Verbeek, 3 December 2007

The paragraph figures refer to the report of Jos H Verbeek. The revised version of the article is provided with row numbers, except in the tables and figures. Within brackets are references to the revised article.

The two major problems

The outcome measures are now defined and described more satisfactorily (pages 8-9, rows 174-192). The primary outcome measure is renamed Return-to-work share to differentiate it from a new secondary outcome measure, Return-to-work chance, which is based on survival analyses (page 8, rows 180-182). As background to having two return-to-work measures we have added sentences describing return-to-work measurement (page 4, rows 69-72). The analyses concerning Return-to-work chance are described on page 9, rows 196-199; the results are shown on pages 13-14, rows 311-316 and in the new table 5, page 33, and are discussed on page 17, rows 399-401. We found that this new outcome measure made Work ability superfluous, and excluded that outcome variable. However, the term ‘work ability’ is retained and is now clearly defined (page 7, rows 152-155). Information about who decides on work ability is added (page 6, rows 117-124).

The power calculation was determined from the number of patients we had to include to be able to show a significant difference in the primary outcome measure, Return-to-work share. This rationale for the power calculation is now explained more explicitly (page 11, rows 251-255). Return-to-work share, which concerned the whole follow-up period, is shown to be equivalent (= fairly equal) between the groups. For the secondary measures, however, enough patients were included to show some significant differences in favour of our programme. In those measures a time variable was included. We have clarified the significant changes over time by adding the outcome measure Return-to-work chance. Those results indicated that in the long run, i.e. beyond 18 months, the programme might be superior, especially for subacute patients. We have modified the last sentence of the conclusions from: “Conclusive and generalizable results concerning this eventual long-term effect would demand a longer follow-up” to: “More conclusive results concerning this possible long-term effect might require a longer follow-up.” (page 3, rows 52-53, and page 21, rows 490-491). With that re-wording, we think our conclusions are congruent with our results.

Other problems

Baseline characteristics showed a significant difference in widespread pain, which motivates the adjustment of the mixed linear model for that variable.

Nineteen of the original 49 references have been excluded. However, to accommodate the valuable views of referees, we have included some new references, for example, Hlobil et al. [7], Staal et al. [17], Steenstra et al. [19] and Schonstein et al. [43]. The number of references is now 44.

The title is improved and shortened.

Language revision has been conducted by a bureau recommended by BMC Musculoskeletal Disorders (Biomedes).
Other major changes, outside the views of Jos H Verbeek

The term “Traditional primary care” has been replaced by the term “primary care”, which we think is more accurate.

The outcome variable Total sick-listing is renamed Net days, which is a more accepted term [30].

The description of cognitive-behavioural rehabilitation has been transferred to a table (Table 1, page 29) and complemented with frequency, the duration of each session and of the entire program, and the intensity of the program.

We have added an Appendix with a detailed description of how our program appeared for the rehabilitation group. We have included a description of cognitive-behavioural rehabilitation in terms of consultations by different care staff and how these were distributed among the three six-month periods (page 15, rows 345-348 and Figure 4).

We have added a description of primary care in terms of consultations by different care staff and how these were distributed among the three six-month periods (page 15, rows 352-357 and Figures 4 and 5).

The possible deficiencies of our programme are described more in detail. As background, we mention the possible connection between relapse problems and short duration of rehabilitation programmes (page 5, row 92-95), how our programme tried to address these problems (page 7, rows 155-156) and the possible negative consequences of our methodology (pages 18-19, rows 433-445).

The figures for Net days (former Total sick-listing) and Visits (Figures 2 a-b and 3 a-b) have been reworked to clarify the changes over time.

The article has been shortened (from 5300 to 4700 words).
Comments on the report of referee 3, Bart J.B. Staal, 19 December 2007

The paragraph figures refer to the report of Bart J.B. Staal. The revised version of the article is provided with row numbers, except in the tables and figures. Within brackets are references to the revised article.

Abstract
1. We now explicitly mention that the population is sick-listed (page 2, rows 20-21).

2. The introduction of a new secondary outcome measure made Work ability superfluous, so we excluded the latter variable. The new variable is Return-to-work chance (page 2, row 30, defined at page 8, rows 180-182). It was added to accommodate the comments about the use of survival analysis. The main measuring point is at 18 months (page 2, row 33), but it is also analysed at six and 12 months (page 9, row 199). The primary outcome measure (Return-to-work) is renamed Return-to-work share (page 2, rows 29-30) to differentiate it from the new outcome measure.

3. The three six-month periods that the whole follow-up period was composed of. The earlier expression is replaced by “the three component six-month periods” (page 2, row 32).

4. The outcome variable Total sick-listing is renamed Net days, which is a more accepted term [30]. Point estimates (including 95% intervals) at 18 months of Return-to-work share, Net days and Visits are presented on page 32, Table 4. Point estimates of Return-to-work chance at six, 12 and 18 months are presented on page 33, Table 5.

5. The power calculation was determined from the number of patients we had to include to be able to show a significant difference in the primary outcome measure; this is now explained more explicitly (page 11, rows 251-255). Return-to-work share, which concerned the whole follow-up period, was shown to be equivalent between the groups. As regards the secondary measures, however, enough patients were included to show some significant differences in favour of our programme. In those measures a time variable was included. We have clarified those significant changes over time by adding the outcome measure Return-to-work chance and reworked the figures of Net days (former Total sick-listing) and Visits (Figures 2 a-b and 3 a-b). Those results indicated that in the longer run, i.e. beyond 18 months, the programme might be superior, especially for subacute patients. We have modified the last sentence of the conclusions from: “Conclusive and generalizable results concerning this eventual long-term effect would demand a longer follow-up” to: “More conclusive results concerning this possible long-term effect might require a longer follow-up.” (page 3, rows 52-53, and page 21, rows 490-491). With that re-wording, we think our conclusions are congruent with our results.

Background
Sick-listing is a comprehensive term, which is now described more in detail (page 6, rows 117-124). The connection between sick-listing and work ability is elucidated (page 7, rows 152-155).
7. OECD is now written out (page 5, rows 101-102).

8. The definitions of acute, subacute and chronic BNP are now re-formulated (page 4, rows 76-78). The upper limit of acute is changed to 21 days (from 28 days) to be more congruent with the evidence-based Swedish definitions [2]. Concomitantly, the lower limit of subacute BNP is lowered from 29 to 22 days.

Methods

9. The history of the rehabilitation centre has been shortened (page 7, rows 145-147).

10. The description of the cognitive-behavioural rehabilitation has been transferred to a table (page 29, Table 1) and complemented with frequency, the duration of each session and of the entire program, and the intensity of the program. ‘Behaviourist’ is now explained (page 7, row 150). The content of our programme, including its possible deficiencies, is described in greater detail. As background, we mention the possible connection between relapse problems and short rehabilitation programmes (page 5, row 92-95). A sentence describing how our programme tried to address these problems is added (page 7, rows 155-156) along with discussion of the possible negative consequences of our methodology (pages 18-19, rows 433-445). We have added an Appendix with a detailed description of how the programme appeared and a description of the consultations by different care staff and how they were distributed over the six-month period (page 15, rows 345-348, and Figure 4).

11. The term “Traditional primary care” is replaced by “primary care”, which we think is more accurate. The two treatment groups and the interventions that were given are now explicitly defined (page 7, rows 141-142).

Data collection

12. For self-ratings, we used valid and reliable questionnaires where possible. As regards pain score, we have added a reference concerning the visual analogue scale [38]. Health-related quality of life was measured by EQ-5D [39]. Data concerning treatment at the rehabilitation centre were collected from the medical records of the centre. Other health-care data were obtained from follow-up forms, the reliability and validity of which have not yet been established. We have added an explanation and motivation for using those non-validated questionnaires (pages 10-11, rows 235-238).

Outcome measures

13. Sick-listing data for the period prior to the randomization have now been excluded from the analyses of Net days (former Total sick-listing)(Figure 2 a-b).
Power calculation

14.
The paragraphs describing the power calculation and the preliminary study are merged and the net length is shortened. The rationale of the expected difference of at least 22% is explained more explicitly (page 11, rows 248-255).

Inclusion procedure and randomization

15.
We wanted a wide-ranging description of the patients that would include not only their subjective ratings but also objective measures of functioning. That is useful, for example, for evaluating possible differences in the treatment results. According to (e.g.) Lindström et al., the PILE-test is sensitive to changes of sick-listing (Lindström I, Öhlund C, Eek C, Wallin L, Petersen LE, Nachemson A: Mobility, strength and fitness after a graded activity program for patients with subacute low back pain. Spine 1992;17:641-652).

The randomization was performed at the health-centre from which the patient was recruited (pages 11-12, rows 260-265).

Blinding

16.
A health adviser participated in cognitive rehabilitation (Figure 4 a). There were no such health-care staff in primary care (Figure 4 b). It was therefore not possible to blind the analyser to the treatment allocation in the analyses of health-care visits.

Results

17.
The results of the pre-randomization period are excluded from the analyses of Net days (Figures 2 a-b). Most of the information about missing data has been transferred to a table (page 31, Table 3).

18.
Some information about the participant flow has been moved to Figure 1, which is complemented with information concerning the two deceased patients in the rehabilitation group. Those two patients were excluded from the outcome analyses, except the Cox regression analysis (page 10, rows 212-213). One patient in the rehabilitation group preferred to continue in primary care (Figure 1), but according to the intention-to-treat, that patient was included in the outcome analyses of the rehabilitation group. That leaves 61 countable patients in the rehabilitation group concerning the analyses of Return-to-work share, Net days and Visits.

19.
The results section has been shortened. In Table 2 (page 30), some less relevant baseline characteristics are excluded and some of the variables are renamed for greater clarity. The figure for return-to-work is excluded and instead a more comprehensive table of the point estimates of the effect measures at 18 months has been added (pages 32-33, Tables 4-5). The figures for Net days and Visits have been reworked (Figures 2 a-b and 3 a-b). Treatment efforts have been illustrated in Figures 4 and 5.
20. The former Figure 3 has been excluded together with the variable *Work ability* and the related analyses by GEE (General estimated equations).

**Conclusion**

21. See our answer on p. 5. With that rewriting, we think the conclusion now is congruent with the results.

**Other major changes, outside the views of Bart J.B. Staal**

The title is shortened.

The article is shortened (from 5300 to 4700 words).

Several references have been excluded. To accommodate the valuable views of the referees, some new references have been added, for example, Hlobil et al. [7], Staal et al. [17], Anema et al. [8], Steenstra et al. [19], Heymans et al. [45]. The net number of references is slightly lower (44 as compared with 49).

As background to having two return-to-work measures, we have added two sentences describing return-to-work measurement (page 4, rows 69-72).

We have added a closer description of primary care in terms of consultations by different care staff and how these were distributed among the three six-month periods (page 15, rows 352-357 and Figures 4-5).
Comments on the report of referee 4, Ivan Steenstra, 24 December 2007

The paragraph figures refer to the report of Ivan Steenstra. The revised version of the article is provided with row numbers, except in the tables and figures. Within brackets are references to the revised article.

1. The aim of the study in the final paragraph of the introduction paragraph has been rephrased (page 5, rows 104-106). The specific aim is also reproduced in the Abstract (page 2, rows 22-24). The term “traditional primary care” has been replaced by “primary care”, which we think is more accurate.

2. The description of the cognitive-behavioural rehabilitation has been transferred to a table and complemented with frequency, the duration of each session and of the entire program, and the intensity of the program (page 29, Table 1). We have added an Appendix with a detailed description of how the programme appeared for the rehabilitation group. We have included a description of the program in terms of consultations by different care staff and how these were distributed among the three six-month periods (page 15, rows 345-348 and Figure 4).

The disability legislation in Sweden is now described (page 6, rows 117-124).

The outcome measures are defined and described more satisfactorily (pages 8-9, rows 174-192). The primary outcome measure (Return-to-work) is renamed Return-to-work share to differentiate it from a new secondary outcome measure, Return-to-work chance (page 2, rows 29-30, and pages 8-9, rows 175-182). This variable has been added to accommodate the valuable point about Cox regression as a preferred technique. As background to having two return-to-work measures we have added sentences describing return-to-work measurement (page 4, rows 69-72). We found that this new outcome measure made Work ability superfluous, and excluded that variable. However, the term ‘work ability’ is retained and is now clearly defined (page 7, rows 152-155). Information about who decides on work ability (= inversely proportional to sick-listing) is added (page 6, rows 119-126). The outcome variable Total sick-listing is renamed Net days, which is a more accepted term [30].

Sick-listing is the comprehensive term and is now described in more detail (page 6, rows 117-124). The connection of sick-listing to work ability is elucidated (page 7, rows 152-155).

The intervals with “1” as the lower figure have been changed to “0”. For example, 0–6 months replaces 1–6 months (Figure 2 a).

Power calculation:
The rationale for the 22% difference is now explained (page 11, rows 251-255).

Inclusion procedure:
Irrespective of our cognitive-behavioural approach we wanted a wide-ranging description of the patients that would include not only their subjective ratings but also objective measures of functioning. That is useful, for example, for evaluating possible differences in the treatment results. According to Lindström et al., the PILE-test is sensitive to changes in the sick-listing (Lindström I, Öhlund C, Eek C, Wallin L, Petersen LE, Nachemson A: Mobility, strength
and fitness after a graded activity program for patients with subacute low back pain.


Randomisation:
We wanted to compare two interventions and to have possible predictors evenly distributed between the treatment groups. In our preliminary study, age was shown to be an important predictive factor with a cut-off at 45 years and older (page 11, rows 242-247). Age as a predictor is supported by, for example, the review of Hansson and Jensen [1]. In our preliminary analyses, however, we noticed no differences in the results between the different age strata, as we did concerning phase strata, and we therefore chose to leave those analyses outside this study. We have shortened and rephrased the description of the randomization (page 12, rows 264-270).

Analyses and statistics:
‘Widespread pain’ and ‘Pain score’ are now mentioned earlier among the baseline characteristics (page 30, Table 2). Even if several researchers analyze pain score as a continuous variable, we prefer to regard it as the ordinal variable it is. While the age of 40 really is double the age of 20, a pain score of 40/100 does not explicitly mean twice as painful as 20/100.

The *Return-to-work chance* (page 8, rows 180-182) is based on a Cox regression analysis for recurrent events (page 9, rows 196-199). This analysis replaces the generalized estimating equations (GEE), which have been removed from the article. The results concerning *Return-to-work chance* are shown on pages 13-14, rows 311-316, and in the new table 5, page 33, and are discussed on page 17, rows 399-401.

It is necessary to collect enough patient data so that possible differences at baseline, which might explain some of the possible differences in the treatment results, are not missed. Large numbers of baseline characteristics, however, even with randomization *by the book*, entail the risk of getting a few variables that will differ significantly between the treatment groups. Such significant differences must then be handled in the evaluation of the possibly different results among the groups, which we describe on page 9, rows 207-208). As widespread pain differed significantly (page 13, rows 297-302), we adjusted the mixed linear models for that variable (page 14, rows 325-326 and page 15, rows 340-341).

We have now added information about the two patients who died (Figure 1, footnotes 3 and 4, and page 13, rows 304-305).

The calculations were performed on an intention-to-treat basis, i.e. with the randomization as starting points (page 10, row 219). We discuss the consequences of not reaching the planned number of participants (page 20, rows 476-478). However, to avoid misunderstanding of this procedure, we have excluded the term ‘extrapolating’.

Results:
Our program for the rehabilitation group had a mean duration of 328 days (Appendix). We have added a paragraph describing the possible connection between relapse problems and short duration of rehabilitation programmes (page 5, row 92-95). A sentence describing how our programme tried to address these problems has been included (page 7, rows 155-156). A discussion of the possible negative consequences of our methodology, including a description of our modifications of graded activity, has been added (pages 18-19, rows 432-445).

3.
Please see our answer on p. 2, paragraph 3.
4. The reference to CONSORT is now mentioned earlier (page 10, rows 219-220; in the previous manuscript at page 13).

5. The results are now presented in a more concentrated way; the whole paragraph is shortened. In Table 2 (page 30), some less relevant baseline characteristics have been excluded and some of the variables renamed for greater clarity. Some information about participant flow has been transferred to Figure 1. The figure for return-to-work is excluded and instead more comprehensive tables of the point estimates of the effect measures have been added (page 32-33, Tables 4-5). The figures for Net days (former Total sick-listing) and Visits have been reworked (Figures 2 a-b and 3 a-b). Treatment efforts are illustrated in Figures 4 and 5. Most of the information about missing data has been transferred to a table (page 31, Table 3). Discussions of the intensity of other programs in comparison with our program have been added (pages 18-19, rows 414-419 and 436).

6. The title is shortened.

7. The article is shortened (from 5300 to 4700 words). Nineteen of the original 49 references have been excluded. However, to accommodate the valuable views of referees, we have included some new references. The number of references now is 44. The following RCTs are added:
   - Staal et al. [17]: page 4, rows 85-88, and page 17, rows 401-404.
   - Hlobil et al. [7]: page 4, rows 70-72 and 87, pages 17-18, rows 409-410, and pages 18-19, rows 433-436.
   - Heymans et al. [44]: page 18, rows 414-416.
   - Steenstra et al. [19]: page 4, rows 87-88 and pages 18-19, rows 433-436.

On average, 2.4 workplace visits were made (Appendix).

Concerning the handling of recurrences, we now mention the relapse problems (page 5, rows 92-95) and how our programme tried to address these problems (page 7, rows 155-156). We have also added a reference in which this problem is discussed [5].

Other major changes, outside the views of Ivan Steenstra

We have modified the last sentence of the conclusions from: “Conclusive and generalizable results concerning this eventual long-term effect would demand a longer follow-up” to: “More conclusive results concerning this possible long-term effect might require a longer follow-up.” (page 3, rows 52-53, and page 21, rows 490-491). With that re-wording, we think our conclusions are more congruent with our results.