CHAPTER 7

Assessing fidelity to treatment protocol: Evaluation of an intervention aimed at changing illness perceptions of patients with chronic low back pain

Petra C. Siemonsma, Carin D. Schroder, Lieke van Oort, David P. French, Ilse Stuive, Leo D. Roorda, Ant T. Lettinga
ABSTRACT

Objective: Treatment fidelity refers to the degree to which a treatment was delivered as planned, and is a key requirement for drawing valid conclusions about trial results. The primary aim of this study was to report on the fidelity to the treatment protocol of a physical rehabilitation intervention for patients with chronic low back pain. Our secondary aim was to investigate which patient-related or therapist-related factors are associated with fidelity, and explore the relationship of fidelity with intervention outcomes.

Design: Observational design embedded in randomized controlled trial. Therapists’ (n=10) fidelity to structure and content of the treatment protocol was assessed independently by two raters in all available patient files (n=125).

Results: Fidelity to the structure of the protocol was very satisfactory (67-100%). Fidelity to the treatment content was less satisfactory (44-56%). Fidelity seems to be influenced more by therapist than by patient related variables. No statistically significant relationship was found between fidelity ratings (structure and content) and outcome (changes in physical activity).

Conclusions: These results confirmed that the treatment had a close resemblance to the planned treatment. No impact of fidelity on outcome could be discerned. Future research should further investigate which therapist features produce high or low fidelity.
INTRODUCTION
Demonstrating treatment fidelity, i.e. the degree to which a treatment was delivered as planned (Waltz, Addis et al., 1993; Yeaton & Sechrest, 1981), is a key requirement for drawing valid conclusions about trial results (Bellg, Borreli et al., 2004; Moncher & Prinz, 1991; Hardeman, Michie et al., 2008). If therapists do not adhere to the planned treatment in the trial (e.g. by omitting essential elements or by adding new ingredients), trial outcomes will not reflect the effects of the intended treatment, leading to erroneous conclusions being drawn (Bellg et al., 2004; Dumas, Lynch et al., 2001; Moncher et al., 1991; Perepletchikova & Kazdin, 2005). Significant effects on outcome may be misattributed to effective treatment if new treatment ingredients are added. Non-significant effects on outcome may be interpreted as ineffective treatment if essential treatment elements are omitted. Despite this, to date, empirical evidence of the relationship between treatment fidelity and outcome is still equivocal (Perepletchikova et al., 2005).

Bellg and colleagues recommended five areas for assessing and improving treatment fidelity: study design, training providers, delivery of treatment, receipt of treatment, and enactment of treatment skills (Bellg et al., 2004). This study was focused on the fidelity of treatment delivery (Waltz et al., 1993; Perepletchikova, Hilt et al., 2009; Bellg et al., 2004), i.e. the extent to which therapists delivered the interventions as intended (Bellg et al., 2004). This type of fidelity can only be adequately judged in the presence of a clear description of the content and conceptual underpinning of the treatment (Moncher et al., 1991; Dumas et al., 2001; Fuhrer, 2005), usually in the form of a treatment protocol. Therefore we refer to this as fidelity to treatment protocol.

The present study reports on the fidelity to treatment protocol of an intervention that is aimed at increasing physical activity in patients suffering from chronic low back pain (CLBP). The treatment is called Cognitive Treatment of Illness Perceptions (CTIP), and involves cognitive restructuring of maladaptive illness perceptions, drawing on Leventhal’s Self Regulation Model (CSRM) (Leventhal, Brissette et al., 2003). Illness perceptions are the patient’s personal beliefs about their illness, and are reliably shown to influence patient behavior (Leventhal et al., 2003). Socratic dialogues are used to explore the patient’s thoughts about the illness, and to identify and discuss those illness perceptions that are maladaptive (Siemonsma, Schroder et al., 2008). The protocol had a clear structure and included three different phases, which started with mapping the illness perceptions of the patient, subsequently disputing of maladaptive illness perceptions and was finalized by applying alternative illness perceptions in an every day situation. Following through the phase was deemed to be an important...
feature of CTIP (Siemonsma et al., 2008). Previous research showed CTIP to be
effective in increasing physical activity (Siemonsma, Stuive et al., 2009;
Siemonsma, Stuive et al., 2011b). In an RCT (N=156), CTIP was compared to a
waiting list control group (international trial register number: ISRCTN35108886).
Post treatment, a statistically significant (p=0.013) and clinically relevant effect
was found on patient-relevant physical activity. Statistically significant changes
were also found on the majority of illness perception scales. In an embedded
study identifying treatment-specific predictors of the effectiveness of CTIP,
rational problem-solving was found to be a significant (p = 0.014) predictor for the
change in physical activity (Siemonsma, Stuive et al., 2011a).

To obtain a comprehensive description of the treatment, assessment of
fidelity to treatment protocol the structure as well as the content of the treatment
were assessed. The protocol included a clear treatment structure with three
phases (mapping, disputing and applying). The treatment content was verified by
assessing whether illness perceptions were indeed targeted, and not perceptions,
ideas or beliefs on other matters. The files of all patients in the study were
assessed, to provide an assessment which is focused on the permanent products
of treatment, i.e. the notes and protocol forms in the patient files.

Despite the importance of treatment fidelity, relatively little research has
been done documenting the levels of treatment fidelity and establishing the
relationship between fidelity and outcome (Perepletchikova et al., 2009; Borrelli,
Sepinwall et al., 2005; Bellg et al., 2004; Hardeman et al., 2008). In addition to
providing an description of fidelity to the protocol, we explored which factors that
might influence the fidelity to protocol. A priori we hypothesized that several
patient related factors might influence fidelity to treatment protocol for this
intervention. First, duration of the CLBP was hypothesized to influence the fidelity
to treatment protocol. Longer duration of the CLBP was assumed to lead to more
rigid patterns of thinking and behavior and in turn, this might lead to protocol
deviations and to lower fidelity. Second, the complexity of the CLBP was included.
In more complex CLBP, in which psychosocial distress was more prominent, it was
hypothesized that it would be harder to follow the protocol and therefore fidelity
to treatment protocol might be lower. Third, the length of the treatment was
hypothesized to be associated with fidelity to treatment protocol, because fidelity
ratings are known to decrease if the intervention included more treatment
sessions (Perepletchikova et al., 2005). We also investigated whether therapist-
related factors were associated with fidelity to the treatment protocol comparing
their relative performance on several aspects of fidelity.
The main aim of this fidelity study was to describe fidelity to treatment protocol by assessing fidelity to treatment structure and assessing fidelity to treatment content. Our second aim was to investigate which patient-related or therapist-related factors are associated with fidelity, and explore the relationship of fidelity with intervention outcomes. Patient-related characteristics examined were duration of CLBP, complexity of CLBP and treatment length. Therapist-related characteristics were explored by examining the relative performance of therapists on all assessments of fidelity to both structure and content. Finally, the relation between fidelity to treatment protocol and outcome of the intervention was assessed, i.e. the change in outcome on physical activity was compared for treatments with high versus low fidelity to treatment protocol.

METHOD

Design
Observational design embedded in randomized controlled trial.

Participants

Therapists. A total of ten therapists provided CTIP according to protocol: four physiotherapists, three occupational therapists and three psychologists. Two therapists are male, and eight female.

From 2001 to 2004 therapists received initial training (six sessions, 21 hours) to ensure competence in and adequate delivery of the treatment according to protocol. Between 2005 to 2008 five follow-up training sessions (20 hours in total) were provided to address specific implementation problems.

Training focused on the treatment protocol, underlying theories, aims and pitfalls. A prominent place in the training was given to practice of skills such as Socratic dialogue, identifying and disputing illness perceptions. These skills were mostly trained in role play. Training was well received by the therapists, and the training program was geared towards addressing the perceived difficulties of the therapists. Besides training, supervision was given during the trial period to counter and avoid problems in treatment. The protocol required that each patient was discussed at least twice with a supervisor.

Patients

The patients in this study are CLBP patients (n=156) who were treated in a trial on CTIP (Siemonsma et al., 2011b) in our rehabilitation centre. Therapists treated CLBP patients, according to the available protocol, in 10-14 weekly one hour sessions.
Intervention materials used to assess fidelity

To facilitate the use of the treatment structure one form was included for each phase of the treatment structure, i.e. mapping, disputing and applying alternatives. The forms were designed to be used in treatment, i.e. not as a registration form to retrospectively fill in what was done. The form for mapping guided the therapist through the dimensions of illness perceptions, by providing boxes to write down the patient’s personal illness perceptions. The form for disputing provided space to write down the maladaptive illness perception that needed to be addressed and a space to write down the alternative illness perception that was deemed to be helpful in resuming physical activity. The form for applying provided space to write down the experiment that was composed to confirm the utility of the alternative illness perception and to evaluate and adjust the experiment. Besides the primary provision of forms to facilitate therapists is following the treatment structure, therapists were free to write down notes on the blank sheets provided in the patient file. These notes formed a second source of information about the use of the treatment structure.

In addition to fidelity to treatment structure, fidelity to treatment content was assessed using notes and forms. This content should be focused on illness perceptions. Therefore the maladaptive and alternative perceptions that were recorded in the notes or forms were assessed for whether they were beliefs about illness or other issues.

Procedures

Patient files were rated independently by two raters (CS and LvO) after the trial had finished. The raters were blinded for the treatment outcome. Ratings were given only for those patients that had not terminated or finished the treatment at the particular phase being rated. To determine which patient files were eligible for scoring it was first determined which (if any) phases were delivered. The files contained notes and form. A phase was classed as ‘delivered’ if the therapists referred in the notes to working on a specific treatment phase or if a forms was used. For patients that were treated but where no information was found in the notes and files, the phase was rated ‘not delivered’. Treatment was classed as ‘stopped’ if no (further) information was found in the patient notes and forms or if it was mentioned in the notes that treatment was stopped before treatment phase was reached. During the trial, therapists were unaware that their patient files would be rated for this study.
Assessing fidelity to treatment protocol

Measures

*Masures of treatment delivery: treatment structure.* First, the notes were rated either 1 (yes, present) or 0 (no, absent) for the questions: 1) are illness perceptions mapped, 2) are illness perceptions disputed, and 3) is an alternative illness perceptions applied in an everyday situation. Second, the same questions were rated for the forms (i.e. mapping form, disputation form, application form). Thus, two ratings for fidelity to the structure of the treatment protocol were available.

*Masures of treatment delivery: treatment content.* Fidelity to the content of the treatment protocol was assessed by rating whether or not illness perceptions were targeted. Assessments were performed for 1) maladaptive illness perceptions and 2) alternative illness perceptions, that were written down either in the notes or on the disputation form. For each perception it was judged: is this perception related to (chronic) low back pain and to physical activity, and does it fit in (one or more) dimensions of the CSRM. This was rated 1 (yes, illness perception) or 0 (no, not an illness perception).

Hypothesized predictors of fidelity to treatment protocol

Three *patient-related variables* were explored for their associations with fidelity to the treatment protocol. 1) duration of CLBP (months), 2) complexity of CLBP (sum score of SCL90) was measured with the Symptom Check List (SCL90) (Peebles, McWilliams et al., 2001; Arrindell & Akkerman, 2004), 3) treatment length (number of sessions) was available from the trial data-set.

*Therapist-related variables* were investigated by comparing the relative performance of therapists on the different measures of fidelity, to assess the extent to which fidelity is unidimensional, i.e. the therapists perform consistently well or poor on all measures of fidelity?

Outcome measure of the main trial (physical activity).

The primary outcome of the treatment was increased patient-relevant physical activity. This was measured both pre- and post treatment using the Patient Specific Functioning List (PSFL) (Beurskens, Vet de et al., 1999; Pengel, Refshauge et al., 2004), in which patients indicated on a VAS (0=no difficulty; 100=impossible) how difficult it was to perform the physical activity that they have self-selected as most important. Total scores ranged from 0 to 100. The PSFL is valid, reliable and sensitive to change (Beurskens, Vet de et al., 1996). PSFL change scores (for pre to post treatment) were available from the trial data-set.
Changes scores were calculated by subtracting the post treatment scores from the pre treatment scores. A negative changes therewith denotes an improvement in physical activity.

**Analyses**

*Interrater agreement* was estimated using Cohens Kappa’s. These values represent the level of agreement between the two raters. The results were interpreted according to Fleiss: poor (< 0.4), fair to good (0.4-0.75), and excellent (> 0.75) (Fleiss, 1981).

*Fidelity ratings* were calculated for each phase for both notes and forms, using the number of patients and percentage of the total number eligible patients. Fidelity levels were judged according to Perepletchikova and Kazdin who defined fidelity levels as: high (80-100%), medium (<80% - 50%) and low (less than 50%) (Perepletchikova et al., 2005). This illness perception intervention was a complex and time consuming treatment (10 to 14 hours); therefore we set an a priori level of > 50% for satisfactory fidelity to treatment protocol.

*Assessments of variables associated with fidelity to treatment protocol* were made for both structure and content. For *patient-related variables* logistic regression was used to study whether the odds of fidelity were influenced by three predictor variables, 1) complexity of CLBP, 2) duration of CLBP, and 3) treatment length. The three variables were entered in the logistic regression equation simultaneously. Assessments of fidelity derived from the patient notes, forms, maladaptive and alternative perceptions were the dependent variables. Statistical significance was set at p < 0.05.

Therapists (n=10) were ranked four times using their mean scores on the measures of fidelity to structure (notes and forms), and content (maladaptive and alternative perceptions). Mean fidelity scores were calculated for each measure of fidelity by dividing the therapist’s fidelity sum score by the total number of patients treated by the therapist. Thus for each therapists four ranks were calculated which ranged from 1 (highest rank) to 10 (lowest rank). The four rankings were summed in an total score, resulting in a score ranging from 4 (highest overall rank) to 40 (lowest overall rank). The three therapists with the highest fidelity (i.e. lowest sum ranking) were contrasted to the three therapists with the lowest fidelity scores to get an indication of the role of therapist related variables on fidelity to treatment protocol.
Assessing fidelity to treatment protocol

Assessment of the effect of fidelity on treatment outcome was made for both structure and content. Independent samples t-tests were used to compare the results for treatments with high fidelity (score 1) to those with low fidelity (score 0) on the change scores of the outcome variable physical activity. Each of the phases, forms and perceptions were grouped into high or low fidelity. The independent samples t-tests were used to evaluate whether there is a significant difference between the two fidelity groups created from the phases, forms and perceptions respectively.

RESULTS

Files available for analysis
Of the CLBP patients (n=156) who were treated in the trial, fidelity analyses were performed on the files which were available for 125 patients. The remaining 31 patient files were unavailable for scoring. For 8 patient files it is likely that they did not exist, as 5 patients did not show up for treatment and 3 stopped within 2 sessions. Of the remaining 23 unavailable files 16 patients were treated by a psychologist and 7 by a physical or occupational therapist. These differences might have resulted from different filing systems for the professions. The number of patients treated ranged from 2 to 28 per therapist, with an average of 12.5 patients treated per therapist.

The number of patient files eligible for scoring declined over the phases, because for some patients treatment was finished or terminated before completing all sessions in the protocol. The phase mapping included 115 eligible patient files, for disputing this was 105 and for application this was 96.

Description of the fidelity to the structure of the treatment protocol
Kappa coefficients for interrater agreement for the notes on the phases disputing and applying were 0.38 and 0.27. The interrater agreement for mapping could not be calculated for the notes, because one rater (CS) only rated ‘yes’. Kappa coefficients for the forms for the phases mapping, disputing and applying were 0.66, 0.96 and 0.81. According to Fleiss’ criteria the agreement was rated as ‘poor’ for the notes and ‘fair to excellent’ for the forms.

There was a decline in fidelity to the use of both the notes (100% to 71%) and the forms (99% to 21%) across phases (see Table 1). Fidelity was much higher according to the notes on mapping and disputing than was the case for to the application phase. According to Perepletchikova (2005) the fidelity to the structure of the treatment protocol was ‘high’ for the notes on mapping and disputing and for the use of the mapping form. ‘Medium’ fidelity was reached for
the notes on applying and the use of the dispute form. Fidelity was ‘low’ for the use of the application form. Overall, assessed fidelity to treatment phases was higher for notes than for forms.

**Table 1.** Fidelity to the structure of the treatment protocol: frequency of treatment notes and the forms used per treatment phase.

<table>
<thead>
<tr>
<th>Treatment phase</th>
<th>Notes</th>
<th>Delivered</th>
<th>Not delivered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping (n=115)</td>
<td>Notes on mapping</td>
<td>115 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Disputing (n=105)</td>
<td>Notes on disputing</td>
<td>101 (96%)</td>
<td>4 (4%)</td>
</tr>
<tr>
<td>Applying (n=96)</td>
<td>Notes on application</td>
<td>68 (71%)</td>
<td>28 (29%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protocol forms</th>
<th>Used</th>
<th>Not used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mapping (n=115)</td>
<td>114 (99%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Disputing (n=105)</td>
<td>70 (67%)</td>
<td>35 (33%)</td>
</tr>
<tr>
<td>Applying (n=96)</td>
<td>20 (21%)</td>
<td>76 (79%)</td>
</tr>
</tbody>
</table>

Values presented are number of patients (percentage of patients).

**Fidelity to the content of the treatment protocol**

The Kappa coefficients for interrater agreement on maladaptive and on alternative illness perceptions were both 0.7, for both notes and form ‘fair to good’ agreement according to Fleiss.

The phase of disputing illness perceptions was reached by 105 patients. For 59 (56%) the maladaptive perceptions were rated to be indeed illness perceptions whereas in 23 patients (22%) they were classed as non illness perceptions. For the remaining 23 patients (22%), the patient files did not contain explicit information about what maladaptive illness perceptions were targeted. Alternative illness perceptions were rated for 99 patients, because for six patients no further treatment was given. Of those 99 patients, for 44 (44%) the alternative perceptions were indeed illness perceptions, in 21 patients (21%) they were classed as non illness perceptions. For the remaining 34 patients (34%), the patient files did not contain explicit information about what perceptions were targeted.

**Assessment of variables associated with fidelity to treatment protocol**

*Patient-related variables.* None of the patient-related variables (duration of CLBP, complexity and treatment length) were significant predictors in logistic regressions, where the dependent variables were fidelity assessment derived.
from the patient notes and forms for the phases disputing and applying. Due to a lack of variance the two variables corresponding to mapping could not be used as dependent variables. Results are presented in Table 2.

**Table 2.** Logistic regression models for variables associated with fidelity to treatment protocol.

<table>
<thead>
<tr>
<th></th>
<th>Beta (SE)</th>
<th>Wald</th>
<th>Sig</th>
<th>Exp(B) (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Disputing illness perceptions: notes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of CLBP</td>
<td>0.00 (0.00)</td>
<td>0.00</td>
<td>0.95</td>
<td>1.00 (0.99;1.01)</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.02 (0.02)</td>
<td>0.57</td>
<td>0.45</td>
<td>1.02 (0.97;1.07)</td>
</tr>
<tr>
<td>Treatment length</td>
<td>-0.01 (0.02)</td>
<td>0.43</td>
<td>0.51</td>
<td>0.99 (0.95;1.03)</td>
</tr>
<tr>
<td><strong>Application of illness perceptions: notes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of CLBP</td>
<td>-0.00 (0.00)</td>
<td>0.72</td>
<td>0.40</td>
<td>1.00 (1.00;1.00)</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.00 (0.01)</td>
<td>0.31</td>
<td>0.58</td>
<td>1.00 (0.99;1.02)</td>
</tr>
<tr>
<td>Treatment length</td>
<td>-0.01 (0.01)</td>
<td>1.54</td>
<td>0.21</td>
<td>0.99 (0.97;1.01)</td>
</tr>
<tr>
<td><strong>Disputing illness perceptions: forms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of CLBP</td>
<td>0.00 (0.00)</td>
<td>0.04</td>
<td>0.84</td>
<td>1.00 (0.99;1.00)</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.00 (0.01)</td>
<td>0.01</td>
<td>0.93</td>
<td>1.00 (0.99;1.01)</td>
</tr>
<tr>
<td>Treatment length</td>
<td>0.01 (0.01)</td>
<td>1.14</td>
<td>0.29</td>
<td>1.01 (0.99;1.03)</td>
</tr>
<tr>
<td><strong>Application of illness perceptions: forms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration of CLBP</td>
<td>0.00 (0.00)</td>
<td>0.01</td>
<td>0.91</td>
<td>1.00 (1.00;1.00)</td>
</tr>
<tr>
<td>Complexity</td>
<td>0.00 (0.01)</td>
<td>0.07</td>
<td>0.80</td>
<td>1.00 (0.99;1.02)</td>
</tr>
<tr>
<td>Treatment length</td>
<td>0.02 (0.01)</td>
<td>2.33</td>
<td>0.13</td>
<td>1.02 (1.00;1.04)</td>
</tr>
</tbody>
</table>

CLBP = chronic low back pain

**Therapist-related variables.** The mean fidelity scores per therapists (n=10) were calculated for structure (notes and forms) and for content (maladaptive perceptions and alternative perceptions). This resulted in the following mean scores (range) for 1) treatment structure a) notes 0.75 (0.50-1.00) and b) forms 0.55 (0.42-0.80), and 2) treatment content a) maladaptive illness perceptions 0.47 (0.20–1.00) and b) alternative illness perceptions 0.41 (0.10–0.90). Therapists were ranked for their mean scores on notes, forms, maladaptive perceptions and alternative perceptions. The four rankings were summed in an total score. This resulted in scores ranging from 4 to 32. The score of 4 thus denotes that one therapist scored rank 1 on all four categories.

Three issues arise from a comparison of the three therapists with the highest rankings (range 4–14) with the three therapists with the lowest rankings
(range 28-32). First, the three best therapists score consistently high on all four categories. Second, the three therapists with the lowest scores have more inconsistent scores which ranged from 3 to 10. The therapist scoring 32 had three bottom rankings (ranks 9 and 10) and one top ranking (rank 3) for fidelity to structure as assessed from the notes. Details of the rankings are not shown. Third, no apparent influence was seen for professional background in the ranking. All three professions were found in the bottom ranking and in the middle ranking, and physical and occupational therapists were found in the top ranking. Therefore there are no indications that profession played a decisive role in fidelity to the protocol.

**Table 3.** Independent samples t-test for fidelity (high versus low) on treatment outcome (physical activity).

<table>
<thead>
<tr>
<th></th>
<th>Mean difference on change in physical activity</th>
<th>SE dif</th>
<th>T (df)</th>
<th>p</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Treatment structure (high vs. low fidelity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disputing Notes</td>
<td>5.1</td>
<td>12.9</td>
<td>0.40 (96)</td>
<td>0.70</td>
<td>-20.6; 30.8</td>
</tr>
<tr>
<td>Applying Notes</td>
<td>8.8</td>
<td>5.9</td>
<td>1.53 (89)</td>
<td>0.14</td>
<td>-2.8; 20.4</td>
</tr>
<tr>
<td>Dispute Form</td>
<td>-2.5</td>
<td>5.4</td>
<td>-0.46 (96)</td>
<td>0.65</td>
<td>-13.2; 8.3</td>
</tr>
<tr>
<td>Application Form</td>
<td>-7.3</td>
<td>6.6</td>
<td>-1.10 (91)</td>
<td>0.27</td>
<td>-20.4; 5.9</td>
</tr>
<tr>
<td><strong>Treatment content (high vs. low fidelity)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maladaptive illness perceptions</td>
<td>-6.0</td>
<td>4.6</td>
<td>-1.3 (108)</td>
<td>0.20</td>
<td>-15.2; 3.2</td>
</tr>
<tr>
<td>Alternative illness perceptions</td>
<td>-0.1</td>
<td>4.9</td>
<td>-0.30 (108)</td>
<td>0.97</td>
<td>-9.8; 9.6</td>
</tr>
</tbody>
</table>

Values resulted from t-test for Equality of Means (equal variances are assumed); physical activity was measured with Patient Specific Functioning List (PSFL).

**Assessment of the effect of fidelity on treatment outcome**

Independent samples t-tests were used to compare the results for treatments with high fidelity to treatment with low fidelity on the change scores of the outcome variables physical activity. For each of the phases, forms and illness perceptions the results were grouped in high fidelity (score 1) and low fidelity
Assessing fidelity to treatment protocol

The primary aim of this paper was to describe to what degree the treatment, as planned in the protocol, was actually delivered. We have put great effort into establishing and maintaining therapists' fidelity to the treatment protocol of an intervention aimed at changing illness perceptions of patients with chronic low back pain (called CTIP). Treatments that are multifaceted, multi-session treatments and which are applied over a period of time are expected to yield lower ratings than uncomplicated treatments (Perepletchikova et al., 2005). Therefore we are very satisfied with the high fidelity ratings for the treatment structure: mapping (100%) and disputing (96%) and the medium fidelity to the application phase (71%). For forms the scores were high (mapping 99%), medium (disputing 67%) and low (application 21%). The percentages are higher than our a priori set level of at least 50%, except for the score on application form. This latter finding warrants further investigation into the causes of this low use. A likely explanation is that for some patients structured practice did take place although the application form was not used.

We assessed all treatment sessions for all available patients' files, in contrast to many fidelity studies which generally rely on a sample of patients or sessions. Another strength of this approach is also that by focusing on permanent products of interventions, we avoided influencing therapists' behavior, as would have been the case with sessions being recorded or videotaped. This has been reported to potentially inflate ratings (Perepletchikova et al., 2005). The lower scores on fidelity on to the forms illustrated on one hand that therapists were unaware that their patient files would be assessed, and on the other hand this illustrates that results may well depend on the sources used for such assessment (i.e. written information from notes or forms, or other sources). The difference in interrater agreement for forms (fair to excellent) and notes (poor) probably reflects that the structured nature of the forms was clearer to rate than the...
unstructured notes. Here also the results seem to have depended on the sources for such assessments. In sum, we conclude that the fidelity to the treatment protocol for CTIP is certainly satisfactory for the use of the treatment structure, except for the use of the application form as assessed in the patient notes.

The results for fidelity to the treatment content showed lower percentages than those for structure. Generally assessment of fidelity is centered around the structure of treatments. Such assessment answer questions such as: did goals setting take place, and was information provided? However this leaves open the content of what was actually done. This would require questions such as: were the goals in line with the treatment aims and was the information appropriately tailored to the patient’s needs. The description of the fidelity to the treatment content in this study is therefore a strength of this paper and useful addition to the ratings for structure. In 56% of the patients the maladaptive perception could be identified as an illness perception, and in 44% of patients the alternative perception was an illness perception. In other words a substantial percentage of the perceptions described in the patient notes or forms were perceptions however, not always illness perceptions about CLBP. These percentages are judged as medium and low fidelity (Perepletchikova et al., 2005). Our self-set limit for satisfactory fidelity of > 50% was only met for the identification of maladaptive illness perceptions. This might indicate that identifying maladaptive illness perceptions and formulating alternative illness perceptions is a challenging task for the health professional and may warrant careful consideration in future implementation. A further reason for careful consideration of this issue is that the therapist rankings for content show that all professions share this difficulty. Therewith the assessment of fidelity to the treatment content provided insights that would not have surfaced if only the fidelity to structure was assessed.

So, how do these ratings compare to other studies on fidelity? In general this treatment is classed as multifaceted (including several phases, forms and skills) and multi session (10 to 14 sessions of one hour, average actual delivery 8,5 hours). Compared to other rehabilitation interventions for CLBP (16-26 hours) (Leeuw, Goossens et al., 2008) this treatment is relatively light. Rating for an intervention (4 sessions) to increase physical activity among sedentary adults with a family history of Type 2 diabetes (ProActive trial) (Hardeman et al., 2008) resulted in an overall mean of 45% fidelity to the treatment structure. The percentages for fidelity reached in our study are somewhat higher (except for the application form) than a fidelity study on CLBP rehabilitation by Leeuw et al., reporting 90-86%, 67-65%, and 67-58% for the three phases of treatment (Leeuw
Assessing fidelity to treatment protocol

et al., 2008), which already yielded high scores. Overall we reached very high percentages. We think that this is due to the training of the therapists over a prolonged period of time, the strong focus on perceived implementation problems in the training, the possibility to discuss individual treatment problems in supervision, and the lay-out of the protocol as a session-by-session patient file.

Did this description of the fidelity result in gaining a better understanding of the results of the randomized controlled trial of CTIP? (Siemonsma et al., 2011b; Siemonsma et al., 2009) The study showed that indeed the actual treatment provided to the patients was a close resemblance of the planned treatment, i.e. the treatment protocol. It is therefore likely that the results found in the RCT (i.e. statistically significant changes in both physical activity and illness perceptions) indeed can be attributed to CTIP. However, fidelity was not perfect, and therefore this leaves potential for improvement of the treatment delivery.

And what do the study results add to the interpretation of the predictor study, which aim was to identify best candidates for CTIP? (Siemonsma et al., 2011a) The main result for the predictor study was that rational problem-solving skills was found the only predictor of physical change in CTIP, explaining 3.9% of the total variance on physical change. Indeed a small proportion of the total variance, but a remarkable one given the a priori nature of the study hypotheses and the complexity of both the chronic low back pain problem and the treatment. The 56% identification of maladaptive illness perceptions and the 44% formulation of alternative illness perceptions, may indeed indicate that the treatment did not call upon the patients’ rational problem-solving skills as much as was expected. The predictor rational problem-solving may have been proven to be a stronger predictor if this skill was called upon much more. Possibly other predictors have not reached significance for the same reason.

The secondary aim of this paper was to explore the relation of fidelity with outcome and to explore which patient-related or therapist-related factors are associate with fidelity. The relationship of fidelity with the outcome of the trial (change in physical activity) was explored, and was found to be statistically non-significant. A number of factors are likely to contribute to explaining this finding: 1) patient files as the main source of information may reflect only a part of the treatment that was delivered, 2) in treatments as complex as the one studied here, many factors influence to outcome, therewith reducing the impact of treatment fidelity on the outcome, 3) the uniformly high levels of fidelity reduced the variation in delivery, and therewith the (statistical) possibility to find an
association, 4) with a sample size of effectively n=125, the study is likely to suffer from lack of power to detect the limited effect size that fidelity may produce.

None of the patient-related variables (duration of CLBP, complexity and treatment length) had a significant association with the dependent variables indicating fidelity. Therewith the influence of these factors on fidelity seems small. The role of therapist-related variables seem larger, because the top three ranked therapists consistently scored high on fidelity to both structure and content of the treatment. This study does not provide insight into which therapist-related factors are responsible for high or low fidelity. Future research should investigate which factors are responsible. Given that therapist fidelity is a series of behaviours, general models of human behaviour, e.g. the Theory of Planned Behaviour (Ajzen, 1991) should be appropriate for studying this further.

For the general discussion on the relationship between fidelity and outcome, in which higher fidelity is assumed to lead to better outcomes, the statistically non-significant results from this study add to the available equivocal results. It should be noted though, that the relationship between fidelity and outcome is an interaction: it relies on both an effect of the treatment, as well as variation in the delivery of treatment moderating this effect. Finding statistically significant interactions is notoriously difficult in field settings, for a variety of statistical reasons (McClelland & Judd, 1993).

However, for this specific discussion on fidelity to treatment protocol of this particular treatment, the results add detail to the understanding of the trial results. The low use of the application forms, and lower scores on maladaptive and alternative illness perceptions give rise to new questions about the treatment’s working mechanism, such as: is the phase of applying needed in all patients; can treatment be effective in some patients only by mapping their illness perceptions, or by highlighting possible maladaptive illness perceptions, and is it therefore possible that not in all patients it is necessary to formulate alternatives? In the further development of this treatment and in similar illness perception focused treatments these questions warrant further attention.

In conclusion, this fidelity study generally showed satisfactory fidelity to the structure of the treatment protocol, a very positive finding. Improvement of CTIP’s fidelity is best sought in the content of the treatment protocol, in other words, in making sure that the treatment focus is on illness perceptions instead of perceptions that are more generally related to problems with their CLBP. The assessment fidelity to content was therewith shown to be a valuable addition to the commonly used assessment of the fidelity to structure. No significant effects
were found for the effect of fidelity on treatment outcome. Therefore this study cannot confirm the direct relevance of treatment fidelity for outcome. The study was shown to be of value for the interpretation of both the trial on CTIP effectiveness and the predictor study: by highlighting how potentially higher fidelity and stronger predictors may be established in future. Therewith we call for the wider report on fidelity as a means to interpret and understand research results as well as to gain better understanding of the treatment processes.

**REFERENCE LIST**


