A randomized controlled trial on the effectiveness of a classification-based system for sub-acute and chronic low back pain in primary care.

Submitted as:
Apeldoorn AT, Ostelo RW, Van Helvoirt H, Fritz JM, Knol DL, Van Tulder MW, De Vet HCW.
A randomized controlled trial on the effectiveness of a classification-based system for sub-acute and chronic low back pain in primary care.
Abstract

**Study design:** A randomized controlled trial.

**Objective:** To assess the effectiveness of a classification-based treatment approach compared to usual physical therapy care in patients with sub-acute or chronic low back pain (LBP) in primary care.

**Summary and background:** There has been considerable interest in developing classification systems in patients with LBP. Delitto’s classification-based system showed promising results in patients with acute LBP compared to usual physical therapy care, but no trial has evaluated this approach in patients with sub-acute and chronic LBP.

**Methods:** All patients were classified by research physical therapists according to a modified version of Delitto’s classification-based system. Patients were randomly assigned to receive either treatment based on this classification or usual physical therapy care. Follow-up assessments were completed at 8, 26, and 52 weeks. The primary analysis was performed according to the intention-to-treat principle, using multilevel analysis. Outcome measures included global perceived effect, disability, pain intensity, quality of life, fear-avoidance beliefs and psychosocial status.

**Results:** A total of 156 patients were included. There were no statistically significant differences between the treatment groups for any of the outcomes at any of the follow-up time points.

**Conclusion:** The classification-based system as used in this study was not effective for improving physical therapy care outcomes in a population of patients with sub-acute and chronic low back pain in primary care.
Introduction

For primary care providers it is often not possible to make a specific patho-anatomical diagnosis for patients with low back pain (LBP). Consequently, 85-95% of patients with LBP are diagnosed by their general practitioner as having non-specific LBP.\(^1\) Given this absence of diagnostic precision, most clinicians use pattern recognition and patient profiling in an attempt to optimize treatment outcomes.\(^2\)

One of the attempts is a classification-based approach initially developed by Delitto et al.\(^3\) and subsequently updated using more recent research.\(^4\) This approach is based on the identification of specific findings from a patient's history, clinical presentation, and physical examination that are integrated into a classification algorithm. This algorithm is thought to help clinicians identify the treatment most likely to benefit the patient. Two randomized controlled trials (RCTs) have investigated the effectiveness of this approach. Both studies reported slightly more favorable outcomes for the classification-based approach as compared with other physical therapy management strategies in patients with acute and sub-acute LBP and mild to severe disability.\(^5,6\)

The aim of the present study was to determine whether the effectiveness of this classification-based approach generalized to another health care system, other clinicians and another population. In the present study we compared classification-based treatment with usual physical therapy care in patients with sub-acute (6-12 weeks) and chronic (> 12 weeks) LBP in primary care.
Methods

The methods have been described in detail elsewhere.\textsuperscript{7} The Medical Ethics Committee of the VU University Medical Centre in Amsterdam approved this study (registration number: 2008/5).

Patients and setting
Participants were recruited by physical therapists from private physical therapy clinics. Inclusion criteria were: LBP as primary complaint, with or without associated leg pain, age between 18 and 65 years, current episode longer than six weeks, and able to read and write Dutch. Exclusion criteria were: known or suspected specific LBP (e.g. cauda equina compression, fractures), severe radiculopathy, spondylolisthesis (grade 2 or more), serious co-morbidity, psychopathology, lumbar spinal surgery in the previous year, more than one low back operation, more than one year of absence from work due to LBP, currently pregnant or given birth in the past three months, inability to attend six or more regular physical therapy appointments, moderate complaints about one or more items of the Urogenital Distress Inventory (UDI 6, short form),\textsuperscript{8,9} or inability to demonstrate any LBP symptoms during mechanical examination.

Examination and classification algorithm
Potentially eligible patients were referred to one of four research physical therapists, who evaluated eligibility, obtained written informed consent, collected baseline questionnaires and conducted a clinical examination. The research physical therapists classified the patients into one of the following subgroups: direction-specific exercises, manipulation, or stabilization exercises, using the individual criteria of the classification algorithm (Figure 2, Chapter 5). This algorithm is based on an updated version of Fritz et al,\textsuperscript{4} and modified to fit into the Dutch health care system and the typical practice patterns of Dutch physical therapists. Each patient was classified into one subgroup. Following the suggestion of Stanton et al,\textsuperscript{10} patients that could be placed in one single subgroup using only the first part of the classification algorithm received the label ‘clear classification’. Patients were given the label ‘unclear classification’ if the additional criteria in the second part of the classification algorithm were needed.

Randomization procedure
After the examination an independent researcher randomized each patient to the intervention or control group using computer generated pre-prepared randomization lists.
Physical therapists in the participating clinics
In each participating clinic, at least one physical therapist was instructed about the classification system, and at least one physical therapist was instructed about the usual physical therapy care procedure, depending on preference and expertise. The physical therapists providing treatment based on the classification algorithm were not allowed to provide usual physical therapy care for patients included in this trial, and vice versa. For both treatment groups, decisions about the frequency and the number of sessions were left to the discretion of the physical therapist.

Usual physical therapy group
Patients assigned to usual physical therapy care received individually tailored treatment according to the current Dutch physical therapy guidelines.11,12

Classification-based treatment group
Patients assigned to the classification-based group were treated according to their primary classification category (direction-specific exercises, spinal manipulation, or stabilization exercises) for a minimum of 4 weeks. Full details of the interventions are provided in our design article.7 After this period, the physical therapist was allowed to change treatment strategy according to the current Dutch LBP guidelines.11,12

Contrast between interventions
The key distinction between the two treatment groups was the use of a standardized examination protocol and related treatment protocols in the classification-based group as compared to the usual physical therapy group. It was expected that in the latter group examination techniques varied,13,14 and a wide range of treatment goals and strategies was used.

Blinding
The statistician, research physical therapists and the treating physical therapists were aware of the treatment allocation, but were not involved in the data collection. The administrative assistants who collected all follow-up questionnaires were blinded for the patient’s group assignment. Every attempt was made to keep patients blinded to treatment allocation.
Measurements
The baseline questionnaire included socio-demographic characteristics (Table 1) and all primary and secondary outcome measures. The primary outcomes were assessed at 8, 26, and 52 weeks after the start of treatment and the secondary outcomes only at 52 weeks, all by means of postal questionnaires.

Primary outcome measures
Pain intensity over the previous week was measured on an 11-point numerical rating scale (NRS, 0 = ‘no pain’ to 10 = ‘worst imaginable pain’). Functional status was measured with the 10-item Oswestry Disability Index (ODI), version 2.1a,15,16 (score range 0-100) with higher scores indicating lower functional status. Global perceived effect (GPE) was measured by self-assessment on a 7-point Likert scale ranging from completely recovered to worse than ever. This was dichotomized a priori into success (completely recovered and much recovered) and non-success (slightly recovered to worse than ever).

Secondary outcome measures
Fear-avoidance beliefs were measured with the Fear-Avoidance Beliefs Questionnaire Activity section (FABQA, score range 0-24) and Work section (FABQW, score range 0-42),17 with higher scores indicating an increased level of fear-avoidance beliefs. Psychosocial status was assessed with the Örebro Musculoskeletal Pain Screening Questionnaire (ÖMPSQ, score range 0-210),18 with higher scores indicating greater risk of prolonged recovery from LBP. General health status was evaluated with the Short-Form 36 (SF-36).19,20 Scores can be aggregated into two higher-order scores; the Physical Component Summary and the Mental Component Summary (score range 0-100), with higher scores indicating better functioning. Responders to treatment were evaluated for the NRS and the ODI according to an a priori defined cut-off value (> 30% improvement from baseline).21

Sample size
Sample size calculations were based on detecting a clinically relevant difference using results of the studies of Brennan et al6 and Fritz et al,5 and were performed for the 3 primary outcomes (for all: power 0.90; two-sided alpha 0.05). We calculated that 68 patients were needed per group. Anticipating a potential drop-out of 10%, we aimed at 75 patients per group.
Data analysis
The primary analysis was an intention-to-treat analysis. All continuous responses
were analyzed in a linear mixed model, with responses at 0 (baseline), 8, 26 and 52
weeks. In these multilevel analyses therapists clustered under practices, patients
under therapists, and repeated measurement clustered within a patient. Baseline
scores were taken into account and the effect of interest for the present study was the
time by treatment interaction. Regression coefficients with 95% confidence intervals
(CI) between baseline and follow-up measurements were calculated. If appropriate,
analyses were adjusted for patient or therapist characteristics that differed between
the two groups.
For the dichotomous outcomes we used a generalized linear mixed model (logit link)
with the same multilevel structure, and the effect of interest was the difference between
groups at each time point. Odds ratios (OR’s) with 95% CI between intervention and
usual physical therapy care were calculated.
A first per-protocol analysis was performed for all outcome measures to estimate the
extent to which protocol deviations might have influenced the results. A deviation from
the protocol was defined as not receiving treatment after allocation, withdrawal from
therapy after one to three visits and not being treated according to patient’s subgroup
for patients allocated to the classification-based group. Additionally, an exploratory
second per-protocol analysis was performed including only patients with a ‘clear
classification’ label.
If necessary, missing items within a questionnaire were imputed using the expectation
maximization (EM) method, which estimates missing values by an iterative process.
For all tests p < 0.05 (2-tailed) was considered significant. Continuous data were
analyzed using SPSS 18.0 and dichotomized data using STATA 11.1 (StataCorp LP)
software.
Results

From June 2008 to October 2009, 158 patients were randomized. Two patients dropped out immediately after randomization and before treatment started: one patient moved to another country and withdrew consent and another was diagnosed with a spondylolisthesis > grade 1, resulting in a sample of 156 patients. Baseline characteristics of these patients are described in Table 1 and were largely similar in both groups. Figure 2 shows the flow of the participants during the study. The attempt to blind patients to treatment allocation was reasonably successful. Eight weeks after the intervention started, 46% of patients had no idea of the treatment group they had been allocated to and of the 54% of patients that did have some idea, 66% were right. The characteristics of the physical therapists were largely similar in both groups, although 93% of the physical therapists in the classification-based group were also a manual therapist compared to 45% in the usual physical therapy group.

Classification procedure
Of the 156 patients, 74% were examined by the first author (AA), 4% by one of the three other research physical therapists and 22% by the first author together with one of the three other research physical therapists. In this latter situation, there was almost perfect agreement (kappa value 0.95, 95% CI 0.85 to 1) about the final classification. The percentages of patients meeting the criteria of 0, 1, 2 or 3 subgroups using only the first part of the algorithm were 24%, 60%, 14% and 2% respectively. The percentage of patients that meet the criteria of a ‘clear classification’ was 74.4% (n = 116).

Contrast between the two interventions
In the usual physical therapy group, 16% of the patients who were classified into the manipulation subgroup, were treated with manipulations and 10% who were classified into the direction-specific exercise subgroup were treated by a physical therapist who had completed two or more McKenzie courses that focus on these exercises. However, no information was collected whether these patients were actually treated with direction-specific exercises. In the classification-based group, 53% of the patients received follow-up interventions after classification-based treatment.
### Table 1  Baseline characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Classification-based group (n=74)</th>
<th>Usual physical therapy group (n=82)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years</td>
<td>43.2 (11.7)</td>
<td>42.0 (10.9)</td>
</tr>
<tr>
<td>Female, %</td>
<td>54.1</td>
<td>59.8</td>
</tr>
<tr>
<td>Dutch nationality, %</td>
<td>89.2</td>
<td>87.8</td>
</tr>
<tr>
<td><strong>History of LBP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First experience of LBP ever in months†</td>
<td>114 (48-192)</td>
<td>96 (24-216)</td>
</tr>
<tr>
<td>Previous episodes of LBP, %</td>
<td>91.9</td>
<td>80.5</td>
</tr>
<tr>
<td>Lower back surgery, %</td>
<td>1.4</td>
<td>1.2</td>
</tr>
<tr>
<td>Duration of current LBP in months†</td>
<td>5 (2–12)</td>
<td>3 (2–12)</td>
</tr>
<tr>
<td>Current LBP (6-12 weeks), %</td>
<td>17.6</td>
<td>23.2</td>
</tr>
<tr>
<td>Current LBP (&gt;12 weeks), %</td>
<td>82.4</td>
<td>76.8</td>
</tr>
<tr>
<td>Pain intensity in the past week (NRS 0-10)</td>
<td>6.0 (1.7)</td>
<td>6.2 (1.8)</td>
</tr>
<tr>
<td>Pain radiated into the leg, %</td>
<td>43.2</td>
<td>32.9</td>
</tr>
<tr>
<td>ODI (0-100)</td>
<td>18.1 (11.5)</td>
<td>21.9 (14.5)</td>
</tr>
<tr>
<td>ODI (0-100) (≥25%), %</td>
<td>27.0</td>
<td>31.7</td>
</tr>
<tr>
<td>Currently taking medication for LBP, %</td>
<td>8.1</td>
<td>14.6</td>
</tr>
<tr>
<td>Smoker, %</td>
<td>29.2</td>
<td>30.8</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/living with a partner,%</td>
<td>75.7</td>
<td>76.8</td>
</tr>
<tr>
<td>Single/divorced, %</td>
<td>24.3</td>
<td>23.2</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low, n (%)</td>
<td>14 (19)</td>
<td>13 (16)</td>
</tr>
<tr>
<td>Middle, n (%)</td>
<td>23 (31)</td>
<td>38 (46)</td>
</tr>
<tr>
<td>High, n (%)</td>
<td>37 (50)</td>
<td>31 (38)</td>
</tr>
<tr>
<td>Employed, %</td>
<td>83.8</td>
<td>85.4</td>
</tr>
<tr>
<td>Employed and currently working, %</td>
<td>85.2</td>
<td>80.0</td>
</tr>
<tr>
<td>Employed, but currently on sick leave, %</td>
<td>14.8</td>
<td>20.0</td>
</tr>
</tbody>
</table>
### Characteristics Classification-based group (n=74)  Usual physical therapy group (n=82)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Classification-based group (n=74)</th>
<th>Usual physical therapy group (n=82)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FABQ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activity (0-24)</td>
<td>11.6 (5.4)</td>
<td>12.7 (5.2)</td>
</tr>
<tr>
<td>Work (0-42)</td>
<td>11.5 (9.5)</td>
<td>15.1 (11.9)</td>
</tr>
<tr>
<td>ÖMPSQ (0-210)</td>
<td>80.0 (20.5)</td>
<td>87.2 (27.8)</td>
</tr>
<tr>
<td><strong>SF-36</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCS (0-100)</td>
<td>43.7 (8.3)</td>
<td>40.2 (8.7)</td>
</tr>
<tr>
<td>MCS (0-100)</td>
<td>52.3 (8.5)</td>
<td>51.1 (10.6)</td>
</tr>
</tbody>
</table>

### Classification outcome

<table>
<thead>
<tr>
<th>Classification outcome</th>
<th>Classification-based group (n=74)</th>
<th>Usual physical therapy group (n=82)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction-specific exercises, n (%)</td>
<td>45 (61)</td>
<td>40 (49)</td>
</tr>
<tr>
<td>Manipulation, n (%)</td>
<td>22 (30)</td>
<td>20 (24)</td>
</tr>
<tr>
<td>Stabilization exercises, n (%)</td>
<td>7 (9)</td>
<td>22 (27)</td>
</tr>
</tbody>
</table>

Values are the mean (standard deviation) unless otherwise indicated. † Median and inter-quartile range
LBP, Low Back Pain; NRS, Numerical Rating Scale; ODI, Oswestry Disability Index; FABQ, Fear-avoidance Beliefs Questionnaire; ÖMPSQ, Örebro Musculoskeletal Pain Screening Questionnaire; SF-36, Short-Form 36; PCS, Physical Component Summary; MCS, Mental Component Summary
 Eligible patients recruited by physical therapists (n=243)

Completed clinical examination (n=172)

Included and randomized (n=158)

Classification-based group (n=74)

Usual physical therapy care group (n=82)

Follow-up:
8 weeks  (68 / 71)
26 weeks  (64 / 73)
52 weeks  (67 / 76)

Excluded before treatment started:
- Withdrew consent (n=1)
- Spondylolisthesis >grade 1 (n=1)

Not fulfilling inclusion criteria:
- Not insured (n=3)
- Moderate or severe urogenital distress complaints (n=6)
- >1 year of absence from work due to LBP (n=7)
- Others (n=55)

Not fulfilling inclusion criteria:
- Red flags (n=3)
- Severe radiculopathy (n=1)
- Spondylolisthesis >grade 1 (n=1)
- Moderate or severe urogenital distress complaints (n=2)
- Others (n=7)

LBP, Low Back Pain
**Intention-to-treat analysis**

Multilevel analyses adjusted for type of therapist (physical therapist or manual therapist) showed no significant differences between the two groups for any primary or secondary outcome measure (Table 2). For all outcome measures, the intraclass correlation coefficients within physical therapy practices were almost zero and varied within physical therapists between almost zero and 0.14, indicating no practice-specific effect and a minor therapist-specific effect.

Patients allocated to the classification-based group attended fewer treatment sessions compared to the usual physical therapy group after one year (mean difference 4.6, 95% CI 0.7 to 8.5, p = 0.02), but this difference was not significant after the initial 8-week study period (mean difference 1.3, 95% CI -0.2 to 2.8, p = 0.08).

For the intention-to-treat and the following analyses, only the missing items of the ÖMPSQ were imputed (2.2%) provided that at least 75% of the questionnaire was completed.
**Per-protocol analysis**

In the first per-protocol analysis 142 patients were evaluated. The characteristics at baseline did not differ between the two intervention groups. Multilevel analyses adjusted for type of therapist (physical therapist or manual therapist) showed statistically significant differences in GPE with an OR of 4.0 (95% CI 1.2 to 13.5) in favor of the classification-based method at one year. Statistically non-significant differences were seen in the other outcome measures at all time points (data not shown). In the second per-protocol analysis 104 patients with a ‘clear classification’ were evaluated. Baseline characteristics of the patients were largely similar in both groups, although more patients allocated to usual physical therapy care had significantly higher scores on the ÖMPSQ than the patients allocated to classification-based treatment, 90 versus 78 respectively. Multilevel analyses adjusted for type of therapist (physical therapist or manual therapist) and for scores on the ÖMPSQ showed favorable outcomes for classification-based treatment on the GPE with an OR of 5.9 (95% CI 1.5 to 24.0) and on the dichotomized ODI scores with an OR of 7.0 (95% CI 1.0 to 48.2) at one year (Figure 3). No statistically significant differences were seen in the other outcome measures at any time point (data not shown).
Table 2  Mean scores, regression coefficients and odds ratios for primary and secondary outcome measures (ITT, n=156)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classification-based group (n=74)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcome measures</strong></td>
<td></td>
</tr>
<tr>
<td>NRS (0-10)</td>
<td></td>
</tr>
<tr>
<td>overall effect: $F_{3,136} = 0.67$ (p=0.57)</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>6.06</td>
</tr>
<tr>
<td>8 wks</td>
<td>4.04</td>
</tr>
<tr>
<td>26 wks</td>
<td>3.14</td>
</tr>
<tr>
<td>52 wks</td>
<td>3.24</td>
</tr>
<tr>
<td>ODI (0-100)</td>
<td></td>
</tr>
<tr>
<td>overall effect: $F_{3,136} = 0.80$ (p=0.50)</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>18.77</td>
</tr>
<tr>
<td>8 wks</td>
<td>13.66</td>
</tr>
<tr>
<td>26 wks</td>
<td>12.09</td>
</tr>
<tr>
<td>52 wks</td>
<td>10.25</td>
</tr>
<tr>
<td>GPE (success, %)</td>
<td></td>
</tr>
<tr>
<td>overall effect: Wald= 5.46 (p=0.14)</td>
<td></td>
</tr>
<tr>
<td>8 wks</td>
<td>52</td>
</tr>
<tr>
<td>26 wks</td>
<td>63</td>
</tr>
<tr>
<td>52 wks</td>
<td>72</td>
</tr>
<tr>
<td><strong>Secondary outcome measures</strong></td>
<td></td>
</tr>
<tr>
<td>FABQA (0-24)</td>
<td></td>
</tr>
<tr>
<td>overall effect: $F_{1,140} = 1.67$ (p=0.20)</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>11.29</td>
</tr>
<tr>
<td>52 wks</td>
<td>7.90</td>
</tr>
<tr>
<td>FABQW (0-42)</td>
<td></td>
</tr>
<tr>
<td>overall effect: $F_{1,99} = 0.00$ (p=0.99)</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>12.79</td>
</tr>
<tr>
<td>52 wks</td>
<td>9.90</td>
</tr>
</tbody>
</table>
### Table 2: Mean scores, regression coefficients and odds ratios for primary and secondary outcome measures (ITT, n=156)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classification-based group (n=74)</th>
<th>Usual physical therapy group (n=82)</th>
<th>Regression coefficients (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary outcomes measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRS (0-10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall effect: $F_{3,136} = 0.67$ (p=0.57)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline 6.06</td>
<td></td>
<td>6.11</td>
<td></td>
</tr>
<tr>
<td>8 wks 4.04</td>
<td></td>
<td>-0.49 (-1.34, 0.37)</td>
<td></td>
</tr>
<tr>
<td>26 wks 3.14</td>
<td></td>
<td>0.01 (-0.95, 0.98)</td>
<td></td>
</tr>
<tr>
<td>52 wks 3.24</td>
<td></td>
<td>0.13 (-0.83, 1.10)</td>
<td></td>
</tr>
<tr>
<td>ODI (0-100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall effect: $F_{3,136} = 0.80$ (p=0.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline 18.77</td>
<td></td>
<td>22.25</td>
<td></td>
</tr>
<tr>
<td>8 wks 13.66</td>
<td></td>
<td>-0.48 (-4.59, 3.63)</td>
<td></td>
</tr>
<tr>
<td>26 wks 12.09</td>
<td></td>
<td>-2.13 (-6.92, 2.66)</td>
<td></td>
</tr>
<tr>
<td>52 wks 10.25</td>
<td></td>
<td>0.86 (-3.61, 5.33)</td>
<td></td>
</tr>
<tr>
<td><strong>Odds ratio (95%CI), p value</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPE (success, %)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall effect: Wald= 5.46 (p=0.14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 wks 52</td>
<td></td>
<td>1.01 (0.31, 3.28)</td>
<td>0.99</td>
</tr>
<tr>
<td>72</td>
<td></td>
<td>0.66 (0.20, 2.21)</td>
<td>0.50</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td>2.75 (0.84, 9.07)</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Secondary outcome measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FABQA (0-24)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect: $F_{1,140} = 1.67$ (p=0.20)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline 11.29</td>
<td></td>
<td>12.68</td>
<td></td>
</tr>
<tr>
<td>52 wks 7.90</td>
<td></td>
<td>1.38 (-0.73, 3.50)</td>
<td></td>
</tr>
<tr>
<td>FABQW (0-42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect: $F_{1,99} = 0.00$ (p=0.99)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline 12.79</td>
<td></td>
<td>10.67</td>
<td></td>
</tr>
<tr>
<td>52 wks 9.90</td>
<td></td>
<td>1.38 (-0.73, 3.50)</td>
<td></td>
</tr>
<tr>
<td>Variables</td>
<td>Classification-based group (n=74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>ÖMPSQ (0-210)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>effect: $F_{1,140} = 0.00$ (p=0.97)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>80.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52 wks</td>
<td>60.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36, PCS (0-100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>effect: $F_{1,146} = 0.26$ (p=0.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>43.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52 wks</td>
<td>50.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36, MCS (0-100)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>effect: $F_{1,141} = 0.37$ (p=0.54)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>52.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52 wks</td>
<td>53.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRS (&gt;30% improvement from baseline,%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overall effect: Wald = 1.74 (p=0.63)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 wks</td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 wks</td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52 wks</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODI (&gt;30% improvement from baseline,%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overall effect: Wald = 2.61 (p=0.46)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 wks</td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 wks</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52 wks</td>
<td>81</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and type of care provider (manual therapist or physical therapist). Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline. Positive values favour the classification-based group. ITT, Intention-To-Treat; CI, Confidence Interval; NRS, Numerical Rating Scale; ODI, Oswestry Disability Index; GPE, Global Perceived Effect; FABQA and FABQW, Fear-avoidance Beliefs Questionnaire subscales Activity and Work; ÖMPSQ, Örebro Musculoskeletal Pain Screening Questionnaire; SF-36, Short-Form 36, PCS, Physical Component Summary; MCS, Mental Component Summary.
### Usual physical therapy group (n=82) vs. Regression coefficients (95% CI)

<table>
<thead>
<tr>
<th>Usual physical therapy group (n=82)</th>
<th>Regression coefficients (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>87.31</td>
<td></td>
</tr>
<tr>
<td>67.30</td>
<td>-0.15 (-8.68, 8.38)</td>
</tr>
<tr>
<td>40.16</td>
<td></td>
</tr>
<tr>
<td>47.90</td>
<td>-0.81 (-3.99, 2.36)</td>
</tr>
<tr>
<td>50.73</td>
<td></td>
</tr>
<tr>
<td>50.96</td>
<td>1.00 (-2.25, 4.25)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Odds ratio (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>73</td>
<td>0.54 (0.17, 1.70)</td>
</tr>
<tr>
<td>74</td>
<td>0.98 (0.30, 3.16)</td>
</tr>
<tr>
<td>74</td>
<td>1.16 (0.37, 3.69)</td>
</tr>
<tr>
<td>45</td>
<td>0.70 (0.19, 2.57)</td>
</tr>
<tr>
<td>71</td>
<td>1.43 (0.37, 5.47)</td>
</tr>
<tr>
<td>68</td>
<td>1.95 (0.51, 7.40)</td>
</tr>
</tbody>
</table>
Figure 3  Results of global perceived effect and disability for patients with a ‘clear classification’ (per-protocol analysis, n=104)

Percentage recovered (Global Perceived Effect) and percentage >30% improvement on disability (ODI, Oswestry Disability Index) compared to baseline for the classification-based and the usual physical therapy group at 8, 26, and 52 weeks.
Discussion

The classification-based approach was originally developed for patients with acute symptoms and ODI scores above 30. There is preliminary evidence that this approach is more effective compared to usual physical therapy care.5 Our aim was to evaluate if this approach could be generalized to patients with sub-acute and chronic LBP and mainly low disability scores in the Dutch health care system. We found no significant differences between the classification-based approach and usual physical therapy care for any primary or secondary outcome measure. The different outcomes found in the present study probably reflect differences in patient selection criteria, the comparison treatment and the used classification algorithm.

Interestingly, the results of a secondary per-protocol analysis for patients with a ‘clear classification’ (patients that could be classified using only the first part of the classification algorithm) were slightly in favour of the classification procedure compared to usual physical therapy care on the outcome measures GPE and the dichotomized ODI score (> 30% improvement) at 52 weeks. This could be an indication that the classification-based approach might be useful for ‘clear classification’ patients with sub-acute and chronic LBP in the long term. However, these results must be interpreted very cautiously. We did not design or power our study for this subgroup analysis in advance, and outcomes of subgroup analysis have proven to be particularly unreliable.22 Another finding of interest is that for GPE and the dichotomized ODI scores patients allocated to the classification-based approach showed an improvement over the total period of 52 weeks, whereas patients treated according to usual physical therapy care deteriorated in the last 26 weeks. We do not have a clear explanation for these findings. One potentially relevant factor is that most patients in the classification-based group received instruction booklets. This written instruction may have improved the patient’s knowledge about their LBP problem,23 their adherence to the clearly defined protocol-led exercises,24,25 and/or their feeling of control. The use of exercise instruction booklets is not common practice in The Netherlands; therefore we did not administrate this in the usual physical therapy group.

Ideally, a classification algorithm should classify patients into one subgroup only. In the present study, using only the first part of the algorithm 24% of all patients did not meet any of the subgroups and 16% met more than one subgroup. These findings are consistent with Stanton et al.10 They found that 25% of all patients did not meet any subgroup, and 25% met more than one subgroup, although they used a different version of the classification algorithm and recruited only patients with acute and sub-
acute LBP with ODI scores above 20. Therefore, we agree with Stanton et al\textsuperscript{10} that the classification algorithm needs further refinement, e.g. by generating additional criteria, to be able to classify other relevant subgroups and to explore the hierarchical ordering of the algorithm.

**Limitations**
Contrary to the original algorithm\textsuperscript{3} and the algorithm as used in the study of Fritz et al,\textsuperscript{5} we did not include the subgroup traction. However, we do not think this has influenced our results significantly, because only three patients (2\%) had signs of nerve root compression and thus met the modified criteria for patients who might benefit from traction.\textsuperscript{26}

We attempted to provide optimal training and guidance for our treating physical therapists in the classification-based group, however this support may have been insufficient for optimal competence and may therefore have caused an underestimation of the effectiveness of the direction-specific exercises and consequently also of the classification-based approach.

**Conclusions**
The classification-based treatment approach as used in this study did not improve primary care outcome in a population of patients with sub-acute and chronic low back pain. To assess whether the classification-based treatment approach is effective, we suggest further refinement of the classification algorithm or evaluation of the approach in a subgroup of patients who clearly meet the classification algorithm criteria.
References